Arthritis & Rheumatology

An Official Journal of the American College of Rheumatology
www.arthritisrheum.org and wileyonlinelibrary.com

Cover design: Todd Machen

This journal is printed on acid-free paper.

Arthritis & Rheumatology

An Official Journal of the American College of Rheumatology
www.arthritisrheum.org and wileyonlinelibrary.com

VOLUME 78 • March 2026 • NO. 3

Cover image: The image on the cover (Fu et al; pages 672-683) shows immunofluorescence staining of a kidney from a mouse that received IL-23 minicircle injection. In the glomerulus, pSTAT3 expression (green) is increased and colocalizes with the podocyte marker synaptopodin (red). Nuclei are counterstained with DAPI (blue).

In this Issue

Highlights from this issue of *A&R* | By Lara C. Pullen, PhD

Anti-IL-5/R Therapy Effect on OGC Use in Eosinophilic Granulomatosis With Polyangiitis

Eosinophilic inflammation is a prominent feature of eosinophilic granulomatosis with polyangiitis (EGPA), and rheumatologists consider long-term use of oral glucocorticoids (OGCs) the standard of care for its treatment. Studies in both vasculitis and asthma, however, have demonstrated that cumulative exposure to OGCs is significantly correlated with OGC-related toxicity. The phase 3 MANDARA study previously found that benralizumab was noninferior to mepolizumab for the induction of remission in patients with relapsing or refractory EGPA. **In this issue, Nair et al. (p. 695)** report results of post hoc exploratory analyses of the MANDARA study to further elucidate the timing and sustainability of reducing OGCs as well as the cumulative use of OGCs.

p. 695

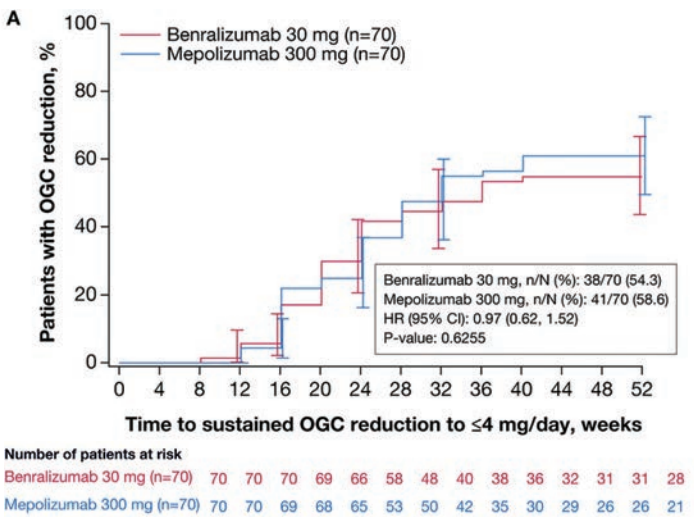


Figure 2A. Time to sustained oral glucocorticoid (OGC) reduction to ≤4 mg/day for patients receiving benralizumab or mepolizumab.

The investigators found that the time to first or sustained OGC reduction to ≤4 mg/day, reduction by ≥50%, and accrued OGC-free duration were similar between treatment groups; however, more patients in the benralizumab group than in the mepolizumab group completely withdrew from OGCs.

In each treatment arm, complete OGC withdrawal was similar across patient subgroups, including those defined by antineutrophil cytoplasmic antibody status and use of immunosuppressives. Patients with a baseline OGC dose <12 mg/day were more likely to discontinue OGCs than those with higher baseline doses, though the authors note that longer study duration might enable more patients with higher starting doses to further reduce their OGC use. The length of accrued OGC withdrawal was not associated with any deterioration of lung function or loss of asthma control. The authors conclude that treatment with the biologics benralizumab or mepolizumab may reduce the need for prolonged OGC therapy and the associated burden and toxicity.

Predictors of Incident Left Ventricular Systolic Dysfunction and Cardiac Recovery in SSc

A lack of tools for predicting the development and course of left ventricular systolic dysfunction (LVSD) means that patients with systemic sclerosis (SSc) myocardial disease often do not receive timely intervention. **In this issue, Kim et al. (p. 713)** evaluated a range of both baseline and time-varying factors and report that distinct demographic, SSc-specific, and cardiac characteristics are associated with an increased risk of incident LVSD in SSc. In particular, skeletal myopathy and anti-Ku antibodies emerged as important risk factors for severe disease. In contrast, among patients with SSc myocardial disease, anti-RNA polymerase III (anti–RNAP III) was associated with higher odds of cardiac recovery.

p. 713

The investigators analyzed data from 2,303 patients with SSc (13,209 echocardiograms) in the Johns Hopkins Scleroderma Center Research Registry. They identified 247 patients (11%) with incident LVSD and 114 patients (5%) with severe LVSD. The researchers reviewed the medical records of patients with reduced ejection fraction (EF) to identify factors associated with severe LVSD and used random forest models to select the most relevant variables for inclusion as predictors of cardiac recovery.

Male sex, Black race, diffuse skin disease, higher modified Rodnan skin score, echocardiographic evidence of pulmonary hypertension, kidney disease, and atrial fibrillation were associated with increased odds of incident LVSD (EF <50%), whereas anticentromere and anti–topoisomerase-1 were protective. For previous EF <50%, tendon friction rubs were associated with lower odds of cardiac recovery and anti–RNAP III was associated with higher odds.

Effect of Time to Start of Biologic Therapy on JIA Treatment Response

Biologic therapies are highly effective in juvenile idiopathic arthritis (JIA). Because research has suggested a potential window of opportunity for treatment in patients with rheumatoid arthritis, rheumatologists have questioned whether a similar window may exist in patients with JIA. **In this issue, de Jonge et al. (p. 743)** report real-life evidence that delaying the start of biologic treatment negatively affects treatment response in children with non-systemic JIA. Each month of delay increased the adjusted odds of having active arthritis after 6 months of treatment by a factor of 1.09 (interquartile range: 1.02-1.17, P=0.009).

p. 743

The prospective study included 130 children (66% female) with JIA whose median age at symptom onset was 11 years. The proportion of patients who reached inactive arthritis status within 6 months of starting biologics was significantly higher among those who initiated biologic therapy within the first 6 months of symptom onset than among those who started biologics more than 12 months after symptoms (83% versus 57%). The authors note in their discussion that, although the analysis adjusted for prior use of conventional synthetic disease-modifying antirheumatic drugs (csDMARD), it did not account for the exact duration of csDMARD therapy or the interval between initiation of csDMARD therapy and biologic treatment.

Although the study cannot rule out the possibility that more severe cases of JIA were over-represented in the early biologic treatment group, the authors conclude that the findings indicate a window of opportunity to control disease activity. They thus call for a paradigm shift in the approach to the treatment of JIA, suggesting that patients likely to require biologics should initiate biologic therapy early without first spending time with other treatment strategies.

Journal Club

A monthly feature designed to facilitate discussion on research methods in rheumatology.

GLP-1RA Use and Risk of Adverse Cardiac and Kidney Outcomes Among Patients With SLE and Lupus Nephritis

Jorge et al. *Arthritis Rheumatol.* 2026;78:665–671.

There is an unmet need in systemic lupus erythematosus (SLE) care to reduce the risks of cardiovascular events and end-stage kidney disease (ESKD). GLP-1 receptor agonists (GLP-1RA) have been previously found to prevent these adverse events in people with type 2 diabetes. This study tested whether this class of medications would have similar cardioprotective and nephroprotective benefits for people with lupus.

As it would be challenging to conduct a randomized clinical trial (RCT) of sufficient size and follow-up duration to address this study question, we used the target trial emulation framework. This method involves the use of observational data to address causal questions by specifying the detailed protocol of a hypothetical RCT that would address the study question, then specifying how to emulate each component of that detailed protocol using observational data. This approach helps avoid common mistakes that introduce bias into observational studies. A key component is emulation of randomization, where all patients included in the study population must have been eligible to receive the treatment strategy or comparator, with the same chance of receiving either one. We compared the initiation of a GLP-1RA or a comparator hypoglycemic medication, DPP4 inhibitors.

All patients included in this study had lupus and type 2 diabetes and had not previously tried either treatment. The propensity score incorporated various factors that influence the use of GLP-1RA or DPP4i, so that the propensity score–weighted populations should have a similar chance of receiving either treatment strategy. We also identified study outcomes to emulate those of an ideal RCT, including the composite kidney disease progression outcome identified as worsening estimated glomerular filtration rate or new-onset ESKD. We also conducted a secondary analysis among patients with lupus nephritis at baseline, as this subgroup of patients with SLE has heightened risks for the outcomes of interest.

Questions

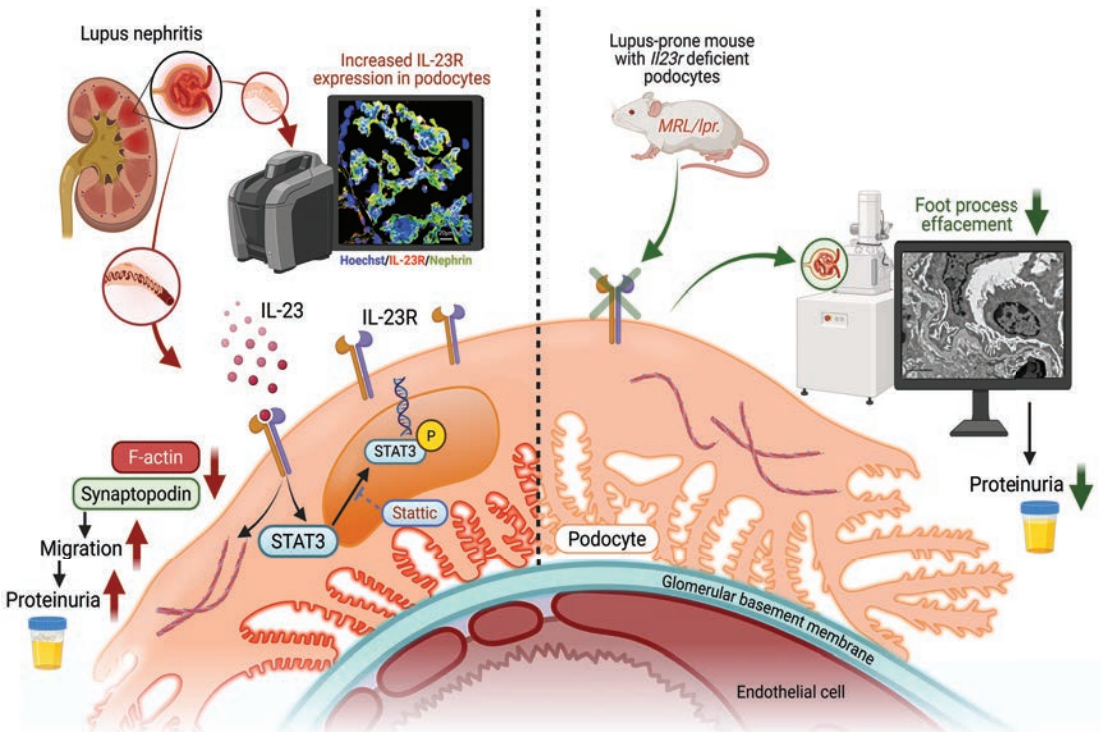
1. **What is known about the effects of GLP-1RA in patients with lupus or other systemic rheumatic diseases?**
2. **Were the methods appropriate for the study question, and what other methods could have been considered?**
3. **How did the strategy to emulate a target trial influence the risk of bias and confidence in the study findings?**
4. **How does the target trial emulation approach compare to other methods of observational studies?**

Clinical Connections

IL-23 Receptor Deficiency in Podocytes Averts Development of Lupus Nephritis

Fu et al. *Arthritis Rheumatol.* 2026;78:672–683.

CORRESPONDENCE
George C. Tsokos, MD: gtsokos@bidmc.harvard.edu



KEY POINTS

- IL-23R expression is elevated in podocytes from patients and lupus-prone mice.
- IL-23 directly injures podocytes by activating STAT3, disrupting the actin cytoskeleton, and increasing podocyte motility and loss.
- Podocyte-specific *Il23r* deletion in lupus-prone mice preserves glomerular architecture and prevents proteinuria.
- Podocyte IL-23R represents a potential therapeutic target.

SUMMARY

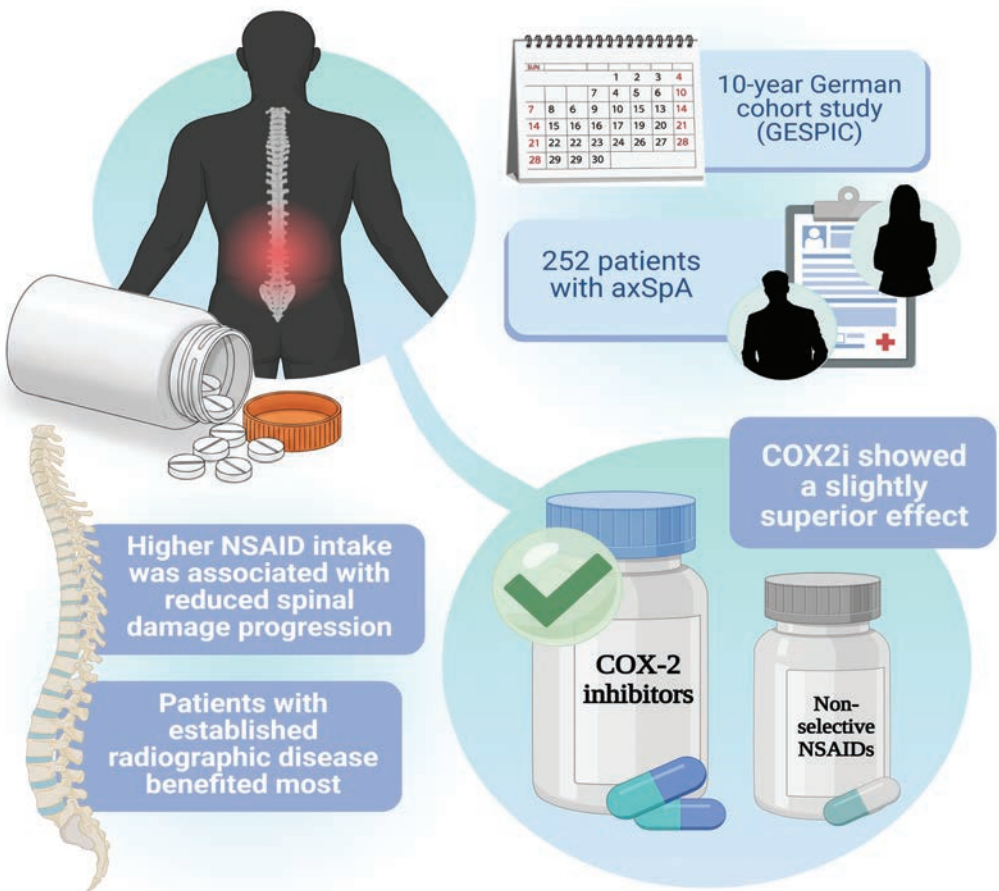
Lupus nephritis is a severe manifestation of systemic lupus erythematosus (SLE) in which podocyte injury plays a central role in disrupting the glomerular filtration barrier and driving renal inflammation and damage. Fu et al. show that interleukin-23 (IL-23), which is elevated in the peripheral blood of patients with SLE, directly compromises podocyte integrity by upregulating its receptor, IL-23R, and activating the pSTAT3/STAT3 signaling pathway. This cascade destabilizes the actin cytoskeleton, increases podocyte motility, and promotes proteinuria and glomerular injury.

Using MRL/*lpr* lupus-prone mice, the researchers demonstrated that podocyte-specific deletion of IL-23R conferred significant renal protection—reducing proteinuria, inflammatory pathology, immune complex deposition, and ultrastructural damage—without altering systemic autoantibody levels, indicating a local kidney-intrinsic mechanism. These findings establish podocyte IL-23R as a central mediator of kidney injury in SLE and a promising therapeutic target for lupus nephritis, particularly in patients with elevated podocyte IL-23R expression.

NSAID Intake and Radiographic Progression in AxSpA

Torgutalp et al. *Arthritis Rheumatol*. 2026;78:582–591.

CORRESPONDENCE
Murat Torgutalp, MD, MSc: murat.torgutalp@charite.de



KEY POINTS

- Higher NSAID intake is associated with a modest but consistent reduction in radiographic spinal progression in patients with axSpA, particularly those with r-axSpA.
- COX2i showed a greater effect on slowing radiographic progression compared to non-selective NSAIDs.
- The effect of NSAIDs on radiographic spinal progression may need to be considered when choosing an optimal treatment approach, especially in r-axSpA.

SUMMARY

Using data from the German Spondyloarthritis Inception Cohort, Torgutalp et al. investigated the potential disease-modifying role of NSAIDs on spinal structural damage over 10-year follow-up in 252 patients with axial spondyloarthritis (axSpA). The analysis used longitudinal statistical models to adjust for complex factors that change over time, ensuring a robust estimation of NSAID effects. Results consistently showed that higher NSAID use was associated with a modest but significant retardation of radiographic progression as measured by the modified Stoke Ankylosing Spondylitis Spine Score.

This protective effect was most notable in patients with radiographic axSpA (r-axSpA), suggesting that NSAIDs may help slow bone fusion in this specific subgroup. When different types of NSAIDs were compared, COX2 inhibitors (COX2i) appeared to have a numerically stronger effect on slowing progression across the total axSpA population compared to non-selective NSAIDs. These findings reinforce the importance of NSAIDs—beyond their role in symptom control—as a therapy for potentially slowing long-term disease progression in patients with axSpA, particularly those with r-axSpA.

Effect of Time to Start of Biologic Therapy on JIA Treatment Response

Biologic therapies are highly effective in juvenile idiopathic arthritis (JIA). Because research has suggested a potential window of opportunity for treatment in patients with rheumatoid arthritis, rheumatologists have questioned whether a similar window may exist in patients with JIA. **In this issue, de Jonge et al. (p. 743)** report real-life evidence that delaying the start of biologic treatment negatively affects treatment response in children with non-systemic JIA. Each month of delay increased the adjusted odds of having active arthritis after 6 months of treatment by a factor of 1.09 (interquartile range: 1.02-1.17, P=0.009).

p. 743

The prospective study included 130 children (66% female) with JIA whose median age at symptom onset was 11 years. The proportion of patients who reached inactive arthritis status within 6 months of starting biologics was significantly higher among those who initiated biologic therapy within the first 6 months of symptom onset than among those who started biologics more than 12 months after symptoms (83% versus 57%). The authors note in their discussion that, although the analysis adjusted for prior use of conventional synthetic disease-modifying antirheumatic drugs (csDMARD), it did not account for the exact duration of csDMARD therapy or the interval between initiation of csDMARD therapy and biologic treatment.

Although the study cannot rule out the possibility that more severe cases of JIA were over-represented in the early biologic treatment group, the authors conclude that the findings indicate a window of opportunity to control disease activity. They thus call for a paradigm shift in the approach to the treatment of JIA, suggesting that patients likely to require biologics should initiate biologic therapy early without first spending time with other treatment strategies.

Journal Club

A monthly feature designed to facilitate discussion on research methods in rheumatology.

GLP-1RA Use and Risk of Adverse Cardiac and Kidney Outcomes Among Patients With SLE and Lupus Nephritis

Jorge et al. *Arthritis Rheumatol.* 2026;78:665–671.

There is an unmet need in systemic lupus erythematosus (SLE) care to reduce the risks of cardiovascular events and end-stage kidney disease (ESKD). GLP-1 receptor agonists (GLP-1RA) have been previously found to prevent these adverse events in people with type 2 diabetes. This study tested whether this class of medications would have similar cardioprotective and nephroprotective benefits for people with lupus.

As it would be challenging to conduct a randomized clinical trial (RCT) of sufficient size and follow-up duration to address this study question, we used the target trial emulation framework. This method involves the use of observational data to address causal questions by specifying the detailed protocol of a hypothetical RCT that would address the study question, then specifying how to emulate each component of that detailed protocol using observational data. This approach helps avoid common mistakes that introduce bias into observational studies. A key component is emulation of randomization, where all patients included in the study population must have been eligible to receive the treatment strategy or comparator, with the same chance of receiving either one. We compared the initiation of a GLP-1RA or a comparator hypoglycemic medication, DPP4 inhibitors.

All patients included in this study had lupus and type 2 diabetes and had not previously tried either treatment. The propensity score incorporated various factors that influence the use of GLP-1RA or DPP4i, so that the propensity score–weighted populations should have a similar chance of receiving either treatment strategy. We also identified study outcomes to emulate those of an ideal RCT, including the composite kidney disease progression outcome identified as worsening estimated glomerular filtration rate or new-onset ESKD. We also conducted a secondary analysis among patients with lupus nephritis at baseline, as this subgroup of patients with SLE has heightened risks for the outcomes of interest.

Questions

1. **What is known about the effects of GLP-1RA in patients with lupus or other systemic rheumatic diseases?**
2. **Were the methods appropriate for the study question, and what other methods could have been considered?**
3. **How did the strategy to emulate a target trial influence the risk of bias and confidence in the study findings?**
4. **How does the target trial emulation approach compare to other methods of observational studies?**

Arthritis & Rheumatology
Vol. 78, No. 3, March 2026, pp 509–522
DOI 10.1002/art.43287

American College of Rheumatology Guidance Statement for Diagnosis and Management of VEXAS Developed by the International VEXAS Working Group Expert Panel

Arsene Mekinian,[1] Sophie Georgin-Lavialle,[2] Marcela A. Ferrada,[3] Sinisa Savic,[4,5] Matthew J. Koster,[6] Olivier Kosmider,[7,8] Thibault Comont,[9] Mael Heiblig,[10] Juan I. Arostegui,[11,12,13] Annmarie Bosco,[14,15,16] Rim Bourguiba,[17,18] Katherine R. Calvo,[19] Catherine Cargo,[20] Chiara Cattaneo,[21] François Chasset,[22] Henrique Coelho,[23] Corrado Campochiaro,[24] Francesca Crisafulli,[21] Stephanie Ducharme-Benard,[25] Raquel Faria,[26,27,28] Franco Franceschini,[29] Micol Frassi,[29] Emma M. Groarke,[30] Carmelo Gurnari,[31,32] Yervand Hakobyan,[33] Yvan Jamilloux,[34] Ciprian Jurcut,[35] Yohei Kirino,[36] Austin Kulasekararaj,[37] Hiroyoshi Kunimoto,[36] Lauren M. Madigan,[38] Heřman F. Mann,[39] Chiara Marvisi,[40,41] Marcin Milchert,[42] Sara Morais,[43] Katja Sockel,[44,45] Francesco Muratore,[41,46] Hideaki Nakajima,[36] Mrinal M. Patnaik,[47] Luísa Regadas,[48] Marie Robin,[49] Abraham Rutgers,[50] Carlo Salvarani,[41,46] Anthony M. Sammel,[16,51] Joerg Seebach,[52] Pierre Sujobert,[53] Alessandro Tomelleri,[24] Geoffrey Urbanski,[54,55] Frédéric Vandergheynst,[56] Romana Vieira,[57] David S. Viswanatha,[58] Ewa Więsik-Szewczyk,[59] Elisa Diral,[60] Benjamin Terrier,[61] Bhavisha A. Patel,[30] Pierre Fenaux,[62] Peter C. Grayson,[63] and David B. Beck,[64,65] on behalf of the International VEXAS working group, and with endorsement of EuroBloodNet, the European Reference Network in Rare Hematological Diseases

Objective. Vacuoles E1 enzyme X-linked autoinflammatory somatic syndrome (VEXAS) is a recently identified rare genetic disorder associated with somatic mutations in the *UBA1* gene. VEXAS presents with a combination of inflammatory and hematologic manifestations, leading to increased morbidity and mortality.

Methods. Given the variability in disease presentation and the limited number of studies to date, no clinical documents currently exist to provide guidance to health care providers about the management of VEXAS. To address this gap, we formed an international multidisciplinary panel of VEXAS experts.

Results. Through formalized meetings and a voting process, the group developed consensus clinical guidance considerations for the management of VEXAS. These considerations offer practical advice on several key topics: (1) clinical features of VEXAS, (2) *UBA1* screening methods, (3) the diagnosis of myelodysplastic syndromes (MDSs) in patients with VEXAS, and (4) prognosis and management. The aim is to provide expert guidance on which patients to test, how to test for VEXAS, how to approach MDS in the context of VEXAS, and considerations for management.

Conclusion. This work marks the first formal international consensus guidance for VEXAS and is intended to be used as a resource for clinicians seeking to understand the disease and its management.

INTRODUCTION

Vacuoles E1 X-linked autoinflammatory somatic syndrome (VEXAS) is a newly recognized rare disease associated with pathogenic, somatic mutations in the *UBA1* gene, resulting in a combination of inflammatory and hematologic manifestations.[1] Since its first description in 2020, more than 450 cases have been documented, with clinical profiling showing heterogeneous manifestations.

[Correction added on 11 March 2026, after first online publication: The last name of Mael Heiblig was initially misspelled and has been corrected in this version.]

[1]Service de médecine interne, Hôpital Saint Antoine, Assistance Publique Hopitaux de Paris, Sorbonne Université, Imagine Institute, CEREMAIAA reference center, Paris, France; [2]Internal Medicine Department, Hopital Tenon, Assistance Publique Hopitaux de Paris, Sorbonne University, CEREMAIAA reference center, Paris, France; [3]Division of Rheumatology and Immunology, Department of Medicine, Baltimore, University of Maryland, USA; [4]Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, Leeds, United Kingdom; [5]National Institute for Health and Care Research (NIHR)-Leeds Biomedical Research Centre, Leeds, United Kingdom; [6]Division of Rheumatology, Mayo Clinic, Rochester, MN, USA; [7]Université de Paris Cité, Institut Cochin, CNRS UMR8104, INSERM U1016, Paris, France; [8]Hematology Laboratory, Cochin Hopital, Assistance Publique-Hôpitaux de Paris, Université de Paris Cité, Paris, France; [9]Service de Médecine interne, Centre Hospitalier Universitaire de Toulouse, Université Toulouse III-Paul Sabatier Faculté de santé, Pole IUC de Toulouse Oncopole CHU, Toulouse, France; [10]Service d'Hématologie Clinique, Hospices Civils de Lyon, Hôpital Lyon Sud, Pierre Bénite, France; [11]Department of Immunology, Hospital Clínic, Barcelona,

The prevalence of VEXAS is estimated to be nearly one million individuals worldwide, and *UBA1* pathogenic variants are present in 1 in 13,591 individuals, although comprehensive global data are still lacking.[2] VEXAS presents a significant diagnostic challenge, as its clinical and laboratory features often mimic those of other established rheumatologic and hematologic diseases. Genetic identification of patients with VEXAS is crucial to guide appropriate management and promote optimal outcomes. Because of the broad spectrum of nonspecific symptoms and limited awareness of the disease, it is likely that many patients remain undiagnosed or misdiagnosed.

This guidance summary was developed by a multidisciplinary panel of experts, including rheumatologists, hematologists, dermatologists, internists, immunologists, pathologists, geneticists, and infectious disease specialists, selected for their experience with VEXAS. The primary purpose of this document is to assist health care providers in identifying the clinical manifestations of VEXAS, promote early confirmatory molecular diagnosis, and provide evidence-based management recommendations. The ultimate goal is to reduce morbidity and mortality associated with this disease. Given the recent identification of VEXAS and the lack of extensive prospective and randomized studies, each statement is primarily based on expert consensus and currently available literature.

The guidance summary presented here was developed using systematic voting and iterative modifications and by including a diverse group of clinicians. The goal is to provide succinct guidance for physicians treating patients with suspected or confirmed VEXAS and support this guidance with available evidence.

METHODS

Composition of the panel and item generation

Consensus guidance was developed by a multidisciplinary expert panel composed of 57 internationally recognized members

Spain; [12]Institut d'Investigacions Biomèdiques August Pi i Sunyer, Barcelona, Spain; [13]School of Medicine, Universitat de Barcelona, Barcelona, Spain; [14]Hematology Department, Prince of Wales Hospital, Sydney, New South Wales, Australia; [15]Hematology Department, NSW Health Pathology Randwick, Sydney, New South Wales, Australia; [16]Prince of Wales Clinical School, University of New South Wales, Sydney, New South Wales, Australia; [17]Department of Internal Medicine, Hôpital des forces de sécurité de l'intérieur, Tunis, Tunisia; [18]Université Tunis el Manar, Faculté de médecine de Tunis, Tunis, Tunisia; [19]Department of Laboratory Medicine, National Institutes of Health, Bethesda, MD, USA; [20]Hematological Malignancy Diagnostic Service, Leeds Cancer Centre, St James's University Hospital, Leeds, United Kingdom; [21]ASST Spedali Civili di Brescia, University of Brescia, Brescia, Italy; [22]Dermatology and Allergology Department, Assistance Publique Hopitaux de Paris, Sorbonne University, Paris, France; [23]Department of Hematology, Centro Hospitalar Vila Nova de Gaia, Porto, Portugal; [24]Unit of Immunology, Rheumatology, Allergy and Rare Diseases (UnIRAR), IRCCS Ospedale San Raffaele & Vita-Salute San Raffaele University, Milan, Italy; [25]Service de médecine interne, Hôpital du Sacré-Cœur de Montréal, Montréal, Québec, Canada; [26]Unidade de Imunologia Clínica, Unidade Local de Saúde Santo António (ULSSA), Porto, Portugal; [27]Autoimmunity and Neurosciences Group, UMIB - Unit for Multidisciplinary Research in Biomedicine, ICBAS - School of Medicine and Biomedical Sciences, University of Porto, Porto, Portugal; [28]ITR - Laboratory for Integrative and Translational Research in Population Health, Porto, Portugal; [29]Rheumatology and Clinical Immunology, ASST Spedali Civili of Brescia and Department of Clinical and Experimental Sciences, University of Brescia, Brescia, Italy; [30]Hematology Branch, National Heart, Lung, Blood Institute, National Institute of Health, Bethesda, MD, USA; [31]Department of Translational Hematology and Oncology Research, Taussig Cancer Institute, Cleveland Clinic, Cleveland, OH, USA; [32]Department of Biomedicine and Prevention, University of Rome Tor Vergata, Rome, Italy; [33]Department of Adult Hematology, Yeolyan Hematology Center, Yerevan, Armenia; Department of Hematology and Transfusion Medicine, National Institute of Health, Yerevan, Armenia; [34]Service de médecine interne, Hôpital de la Croix Rousse, Hospices Civils de Lyon, Lyon, France; [35]Department of Internal Medicine, Dr. Carol Davila Central Military Emergency University Hospital, Bucharest 010825, Romania; [36]Department of Stem Cell and Immune Regulation, Yokohama City University Graduate School of Medicine, Yokohama, Japan; [37]Department of Hematology, Kings College Hospital and Kings College London, London, UK [38]Department of Dermatology, University of Utah, Salt Lake City, Utah; [39]Institute of Rheumatology and Department of Rheumatology, First Faculty of Medicine, Charles University, Prague, Czech Republic; [40]Rheumatology Unit, Azienda Unità Sanitaria Locale-IRCCS di Reggio Emilia, Reggio Emilia, Italy; [41]University of Modena and Reggio Emilia, Modena, Italy [42]Department of Internal Diseases Rheumatology Diabetology Geriatrics and Clinical Immunology with the Sub-Department of Gastroenterology, Pomeranian Medical University in Szczecin, Szczecin, Poland.D43:N68 [43]Centro Hospitalar Universitário de Santo António, Porto, Portugal; [44]Medical Clinic and Policlinic I, University Hospital Dresden, TU Dresden, Dresden, Germany; [45]German Cancer Consortium (DKTK), CHOICE Consortium, Partner Sites, MunichDresden, Germany; [46]Azienda Unità Sanitaria Locale-IRCCS di Reggio Emilia, Reggio Emilia, Italy; [47]Division of Hematology, Department of Internal Medicine, Mayo Clinic, Rochester, MN, USA; [48]Centro Hospitalar e Universitário do Porto, Porto, Portugal; [49]Service d'hématologie-greffe, Hôpital Saint-Louis, Assistance Publique Hopitaux de Paris, Université de Paris Cité, Paris, France; [50]Department of Rheumatology and Clinical Immunology, University Medical Center Groningen, University of Groningen, 9713 GZ Groningen, The Netherlands; [51]Rheumatology Department, Prince of Wales Hospital, Sydney, New South Wales, Australia; [52]Division of Clinical Immunology and Allergy, Department of Medical Specialties, University Hospital and Faculty of Medicine, Geneva, Switzerland; [53]Laboratory of Hematology, Hôpital Lyon Sud, Hospices Civils de Lyon, Pierre Bénite, Lyon, France; [54]Department of Orofacial Sciences, School of Dentistry, University of California, San Francisco, CA, USA; [55]Service de Médecine Interne et Immunologie Clinique, Centre Hospitalier Universitaire Anger, Angers, France; [56]Department of Internal Medicine, Hôpitaux Universitaires de Bruxelles, Hôpital Erasme, Université Libre de Bruxelles, Brussels, Belgium; [57]Rheumatology Department, Centro Hospitalar de Vila Nova de Gaia e Espinho, Vila Nova de Gaia, Portugal; [58]Division of Hematopathology, Department of Pathology and Laboratory Medicine, Mayo Clinic, Rochester, MN, USA; [59]Department of Internal Medicine, Pneumonology, Allergology and Clinical Immunology, Central Clinical Hospital of the Ministry of National Defense, Military Institute of Medicine, National Health Institute, Warsaw, Poland; [60]Unit of Hematology and Bone Marrow Transplant, IRCCS San Raffaele Scientific Institute, Milan, Italy; [61]Service de Médecine interne, Hôpital Cochin, Assistance Publique Hopitaux de Paris, Université Paris Cité, Paris, France; [62]Service de Hématologie, Hôpital Saint-Louis, Assistance Publique-Hôpitaux de Paris, Université Paris Cité, Paris, France; [63]Vasculitis Translational Research Program, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, MD, USA; [64]Center for Human Genetics and Genomics, New York University School of Medicine, New York, NY, USA; [65]Division of Rheumatology, Department of Medicine, New York University School of Medicine, New York, NY, USA.

Drs Mekinian, Georgin-Lavaille, Ferrada, Savic, Koster, Kosmider, Comont, and Heilblig are co-first authors and contributed equally to this work. Drs Terrier, Patel, Fenaux, Grayson, and Beck are co-last authors and contributed equally to this work.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/art.43287.

[Correction added on 21 October 2025, after first online publication: the names of Arsene Mekinian and Sophie Georgin-Lavialle have been corrected.]

[Correction added on 27 August 2025, after first online publication: Affiliation number 41 was added to Francesco Muratore and Carlo Salvarani.]

Address correspondence via email to Arsene Mekinian, MD, PhD, at arsene.mekinian@aphp.fr; or to David B. Beck, MD, PhD, at David.Beck@nyulangone.org.

Submitted for publication December 13, 2024; accepted in revised form April 25, 2025.

from 18 countries including Armenia, Australia, Belgium, Canada, Czech Republic, France, Germany, Italy, Japan, The Netherlands, Poland, Portugal, Romania, Spain, Switzerland, Tunisia, the United Kingdom, and the United States. Each panelist was selected based on clinical expertise in the treatment of patients with VEXAS, with representation from dermatology, genetics, hematology, immunology, internal medicine, rheumatology, infectious diseases, critical care, and pathology specialists. Patient representatives were not included in these efforts.

To ensure comprehensive coverage of the topic, the panel conducted a series of structured virtual meetings. During these discussions, the panel identified and prioritized the critical topics that needed to be addressed. The identified items were organized into four major groups, each reflecting a core aspect of VEXAS.

- Group 1: Clinical and laboratory features of VEXAS. This category focused on identifying the clinical and laboratory features that should prompt consideration of genetic testing for VEXAS.
- Group 2: *UBA1* screening. This section addressed the modalities and best practices for screening *UBA1* mutations, which are pivotal for the diagnosis of VEXAS.
- Group 3: Diagnosis of myelodysplastic syndrome (MDS) in patients with VEXAS. This group explored the intersection of VEXAS with MDS, providing guidance for accurate diagnosis and differentiation.
- Group 4: Outcome, prognostic factors, and management. This category covered the management strategies for VEXAS, including prognostic factors influencing patient outcomes.

Each category was assigned to a subgroup of experts who were responsible for conducting a detailed review of the relevant literature and generating a draft document (please see Appendix A for members of each group). The subgroups worked independently, holding multiple virtual meetings to discuss their findings, synthesize data, and formulate initial statements. A comprehensive virtual and in-person meeting was convened, bringing together all panel members to discuss the proposed items from each group. The panel engaged in an in-depth review and debate of each item, followed by an electronic voting process to determine whether the item should be included in the final statements. This rigorous process ensured that the statements were built on a foundation of consensus and expert validation, resulting in the generation of the actual initial guidance document.

Literature review and level of evidence

The foundation of the statements was an extensive review of the current literature related to VEXAS from January 2020 to August 2024. Each subgroup undertook a thorough literature search using PubMed and Medline, focusing on articles published in English and abstracts from leading international rheumatology and hematology conferences. Given that VEXAS is a newly discovered disease with limited evidence, a formal systematic review was not performed by each subgroup. For cases in which prospective data were unavailable, the statements were based on the best available evidence and expert opinion. All statements are conditional, due in part to a lack of randomized controlled trials in VEXAS. Additionally, the Grading of Recommendations Assessment, Development, and Evaluation methodology was used to rate the quality of evidence supporting each statement.[3,4] Supportive references are provided for each consensus statement but are nonexhaustive and meant to highlight studies with more participants to provide stronger evidence.

Final item generation and voting

The process of finalizing the statements took place at an international, two-day VEXAS workshop held in Paris in May 2024, where the majority of the panel gathered to review, discuss, and refine the proposed items. This workshop provided a critical opportunity for face-to-face deliberation, allowing the panelists to address any outstanding issues, reconcile differing opinions, and achieve a stronger consensus. Following the workshop, an electronic tool was used to conduct a final consensus vote on all items. This tool allowed panel members to provide their feedback anonymously, ensuring that the final statements reflected the true consensus of the group. There was a single round of voting in which participants were asked to vote "yes" or "no" on their agreement to each statement and to provide reasoning in cases of disagreement. Consensus for item inclusion in the statements was defined by a threshold of 80% agreement among the panelists. All proposed statements achieved consensus, and all final voting results are included with each statement (Table 1). Consensus statements are bulleted within the text with expert panel voting results in parentheses, followed by an explanation and supporting data.

RESULTS/RECOMMENDATIONS

The following results summarize the agreed-upon clinical indicators, diagnostic strategies, and therapeutic recommendations developed through structured expert review and formal consensus statements for the four key domains.

Clinical features of VEXAS consensus statements (Figure 1)

VEXAS is a hematoinflammatory disease characterized by a spectrum of rheumatologic, dermatologic, and hematologic features.

Persistently elevated markers of inflammation with a combination of skin, ocular, lung, or cartilage manifestations and/or cytopenia are suggestive of VEXAS.

Table 1. Statements and level of agreements of different guidance statements of groups 1 through 4*

Statements	Agree, n (%)	Disagree, n (%)	Total, n (%)
Group 1: clinical and laboratory features			
VEXAS is a hematoinflammatory disease characterized by a spectrum of rheumatologic, dermatologic, and hematologic features	52 (98)	1 (2)	53 (100)
Persistently elevated markers of inflammation with a combination of skin, ocular, lung, or cartilage manifestations and/or cytopenia are suggestive of VEXAS	50 (94)	3 (6)	53 (100)
Macrocytic anemia is a common hematologic abnormality in VEXAS but may not always be present	52 (98)	1 (2)	53 (100)
Vacuoles present in early erythroid or myeloid precursors from bone marrow aspirates are suggestive of, but not diagnostic for, VEXAS	53 (100)	0 (0)	53 (100)
VEXAS is not a heritable disease	53 (100)	0 (0)	53 (100)
Most cases of VEXAS have been reported in men older than 50 years	52 (98)	1 (2)	53 (100)
Group 2: *UBA1* mutation screening			
The majority of patients with typical VEXAS have missense or splice site mutations at exon 3 of *UBA1*	52 (100)	0 (0)	52 (98)
A variety of sequencing methods can be used to diagnose VEXAS depending on resources and expertise	51 (98)	1 (2)	52 (98)
NGS has increased sensitivity to detect low-level somatic variants with coverage of all coding or splice site regions	50 (96)	2 (4)	52 (98)
Targeted testing for exon 3 variants can be used as an alternative to NGS with the understanding that such modalities may miss somatic variants with low VAF (Sanger sequencing) or variants outside of exon 3 (Sanger sequencing and ddPCR)	50 (96)	2 (4)	52 (98)
Active treatment may affect the accuracy of genetic testing by decreasing mutation level	45 (87)	7 (13)	52 (98)
Somatic mutations defining VEXAS are detected in peripheral blood or bone marrow–derived DNA	51 (98)	1 (2)	52 (98)
If clinical suspicion is high and genetic testing for mutations in *UBA1* at exon 3 is negative on the peripheral blood, clinicians should ensure that testing covers the complete gene and consider testing a bone marrow sample	47 (90)	5 (10)	52 (98)
Group 3: diagnosis of MDS in patients with VEXAS			
Bone marrow examination is recommended in patients with VEXAS with associated cytopenia to exclude an associated hematologic neoplasm	51 (98)	1 (2)	52 (98)
Karyotype and NGS for coexisting somatic mutations should be performed with bone marrow examination	50 (96)	2 (4)	52 (98)
Interpretation of bone marrow is challenging in VEXAS due to the frequent presence of signs of dysplasia, without meeting existing criteria for myelodysplastic syndrome diagnosis based on the WHO and ICC 2022 classifications	50 (96)	2 (4)	52 (98)
The somatic clonal landscape in VEXAS is dominated by mutations in *DNMT3A* and *TET2*, generally without other driver mutations	49 (94)	3 (6)	52 (98)
Group 4: outcomes, prognostic factors, and management			
Patients with VEXAS should be managed by a multidisciplinary team with consideration of referral to an expert center when available evidence	52 (100)	0 (0)	52 (98)
Disease activity is defined by inflammation or worsening bone marrow failure	48 (92)	4 (8)	52 (98)
The goals of treatment include controlling inflammation; preventing and treating bone marrow failure; preventing and treating secondary medical complications arising from implemented treatments; and improving quality of life	50 (96)	2 (4)	52 (98)
Infections, thromboembolic embolic disease, and treatment-related comorbidities are common in VEXAS and can mimic disease activity	49 (94)	3 (6)	52 (98)
Flare of disease can be defined as recurrence of clinical or laboratory signs or symptoms of VEXAS	47 (90)	5 (10)	52 (98)
Glucocorticoids are usually needed to control inflammation in patients with VEXAS and should be tapered slowly to achieve the minimal dose required to control disease activity	52 (100)	0 (0)	52 (98)
Medications that target inflammatory pathways (eg, JAK inhibitors, IL-6 inhibitors) are more effective than conventional DMARDs (eg, methotrexate, azathioprine) or B cell–directed therapies (eg, rituximab)	48 (92)	4 (8)	52 (98)
Medications that reduce or eliminate the clonal burden of disease (eg, azacytidine) can be an effective treatment in some patients with VEXAS syndrome	51 (98)	1 (2)	52 (98)
Allogenic hematopoietic stem cell transplantation may be a curative treatment but should be reserved for select patients with VEXAS after a careful hematologic evaluation	50 (96)	2 (4)	52 (98)
Prophylaxis against opportunistic infections, prevention of thromboembolic disease, and minimization of side effects related to chronic glucocorticoid therapy should be considered in all patients with VEXAS	51 (98)	1 (2)	52 (98)
Future clinical trials will be needed to identify the best treatment for patients with VEXAS	52 (100)	0 (0)	52 (98)

* ddPCR, digital droplet polymerase chain reaction; DMARD, disease-modifying antirheumatic drug; ICC, International Consensus Classification; IL-6, interleukin-6; MDS, myelodysplastic syndrome; NGS, next-generation sequencing; WHO, World Health Organization; VAF, variant allele fraction; VEXAS, vacuoles E1 enzyme X-linked autoinflammatory somatic syndrome.

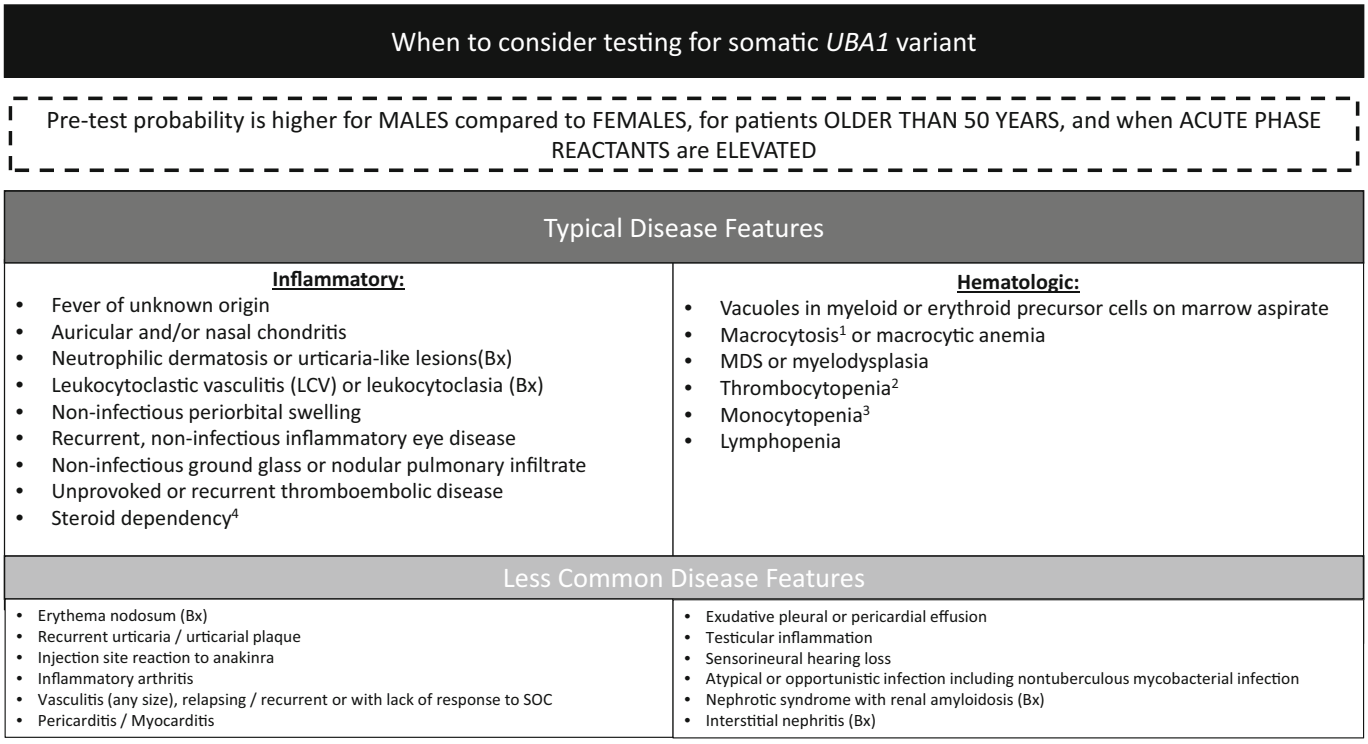


Figure 1. When to consider testing for VEXAS. A list of typical and less common disease features of VEXAS are presented to help guide clinicians regarding when to consider testing for acquired mutations in *UBA1*. Bx, biopsy confirmation required; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; LCV, leukocytoclastic vasculitis; MCV, mean corpuscular volume; MDS, myelodysplastic syndrome; SOC, standard of care; VEXAS, vacuoles E1 enzyme X-linked autoinflammatory somatic syndrome.

Inflammatory features. Inflammatory manifestations are ubiquitous in patients with VEXAS, and elevated acute phase reactants are present at some point in the majority of patients.[1,5] The spectrum of clinical features of VEXAS is wide and highly variable and overlaps several domains and organ systems. Although rare, mild disease forms can occur with minimal to no inflammation (sometimes detectable only in laboratory results).

It is crucial for clinicians and other health providers to be familiar with the most common features present in confirmed patients with VEXAS, including recurrent noninfectious fever (75–100%); various skin involvements (erythema nodosum, maculopapular lesions, panniculitis, urticarial, and possibly others; 85%); lung involvement (interstitial infiltrates, pleuropericarditis; 55–95%); auricular, nasal, and tracheal chondritis (14–64%); venous thrombosis (50%); and various inflammatory ocular involvements (20–50%). Less common organ involvement could also be present, such as peripheral nervous system, heart, and kidneys impairments. A multidisciplinary approach is crucial for both diagnosis and treatment, as patients can require treatment by various medical specialists.

Erythrocyte sedimentation rate (ESR), ferritin, and C-reactive protein (CRP) are increased in most untreated patients and can trend with acute episodes of inflammation. Patients often carry specific rheumatologic diagnoses, but in most cases, their clinical course and initial manifestations are atypical for these diseases with respect to disease severity, pattern of organ system involvement, and response to treatment. For example, a patient with VEXAS may be initially diagnosed with polymyalgia rheumatica or seronegative rheumatoid arthritis, but over time they may develop treatment-refractory systemic inflammation, progressive bone marrow failure, and additional clinical symptoms that phenocopy other rare diseases, such as Sweet syndrome or urticarial vasculitis. Inflammation can occur in various tissues, most commonly affecting the skin, lungs, cartilage, and eyes.[1,5–7] Histologic characteristics are mainly available from skin biopsies, showing the presence of inflammatory infiltrates consisting of immature myeloid cells and variable proportions of mature neutrophils, lymphocytes, and histiocytes.

Most cases of VEXAS have been reported in men older than 50 years

Demographics. In a large cohort of VEXAS comprising 116 patients, the median age at symptom onset was 71 years, and the lower limit of the interquartile range was 66 years.[5] Similar

age at onset has been shown in other cohorts, although the age range extended below 50 years of age in multiple reports.[8–10] Additionally, there is a single case report of a patient presenting with VEXAS-like symptoms at age 23, with a confirmed pathogenic *UBA1* mutation.[11] Many factors likely influence the reported age at disease onset, including heterogeneity of patient and physician reporting, inconsistent recognition of mild symptoms caused by the disease, and disease-related variability such as sex, variant allele fraction (VAF) levels, and mutation type, each of which is currently being studied. Future research on younger populations will help clarify the natural history of this disease, including the timing of somatic mutation emergence and the development of clinical and hematologic manifestations. Testing for VEXAS can be considered in individuals under 50 years of age when clinical suspicion is high; however, most cases to date involve disease onset after age 60. The vast majority of patients identified are male, due to *UBA1* being an X-chromosomal gene, although VEXAS has been reported in women, particularly those with monosomy X.[12]

Macrocytic anemia is a common hematologic abnormality in VEXAS but may not always be present

Hematologic features. Macrocytosis and macrocytic anemia are the most common peripheral blood findings in patients with VEXAS,[13,14] and macrocytosis should be recognized early because it could be initially an isolated presenting feature. Additionally, patients may present with lymphopenia, thrombocytopenia, and monocytopenia. The degree and instances of cytopenia often increase as the disease progresses and may not be evident in early presentations. The bone marrow is typically hypercellular with an increased myeloid-to-erythroid ratio.[13,14] Concurrent conditions such as thalassemia or iron deficiency may obscure red cell macrocytosis by causing pseudonormalization of the mean corpuscular volume. MDS is present at diagnosis in one-third to one-half of patients with VEXAS, although diagnosing MDS in the setting of VEXAS can be challenging[1,5,13,14] (see group 3). Association with plasma cell dyscrasias is found in up to one-quarter of patients, and macrophage activation syndrome associated with VEXAS has also been reported.[1]

Vacuoles present in early erythroid–myeloid precursors from bone marrow aspirates are suggestive of but not diagnostic for VEXAS

Histopathology. Vacuoles in bone marrow histopathology are neither specific nor sensitive for VEXAS; in fact, they may also be found in individuals with nutritional deficiencies (especially copper) and chemotherapy recipients.[13,15,16] Although vacuoles in isolation are not specific to VEXAS, vacuoles in combination with relevant clinical symptoms have provided an 87.5% positive predictive value for pathogenic *UBA1* mutations.[17] Genetic testing for pathogenic *UBA1* variants remains necessary to diagnose VEXAS, regardless of vacuole presence. Absence of vacuoles should not prevent consideration of *UBA1* testing if clinical suspicion remains high.[18]

VEXAS is not a heritable disease

Genetics. To date, all reported cases of VEXAS have been attributed to somatic or acquired mutations in *UBA1*, with no evidence of heritable VEXAS. There has yet to be a documented case of a *UBA1* mutation being transmitted to offspring or inherited, consistent with the mutation being restricted to the blood lineage.[1] However, no standardized assessment of children has been performed, and there are reports of pathogenic *UBA1* mutations being found outside of the blood, in both skin and nail samples.[19–21] In both of these nonhematologic tissues, the *UBA1* mutation may be present due to peripheral blood contamination because the culturing of patient fibroblasts did not reveal any pathogenic mutations.[1]

VEXAS diagnosis: *UBA1* screening modalities consensus statements (Figure 2)

The majority of patients with typical VEXAS have missense or splice site mutations at exon 3 of UBA1.

A variety of sequencing methods can be used to diagnose VEXAS depending on resources and expertise.

Next-generation sequencing (NGS) has increased sensitivity to detect low-level somatic variants with coverage of all coding or splice site regions.

Targeted testing for exon 3 variants can be used as an alternative to NGS with the understanding that such modalities may miss somatic variants with low VAF (Sanger sequencing) or variants outside of exon 3 (Sanger sequencing and digital droplet polymerase chain reaction [ddPCR]).

Active treatment may affect the accuracy of genetic testing by decreasing mutation level.

The diagnosis of VEXAS requires detection of pathogenic somatic mutations in the *UBA1* gene, which typically result in loss of the UBA1b isoform expression or reduction of E1 enzymatic activity.[22] The choice of genetic testing modality depends on availability and cost, with the goal of sequencing all genomic regions where pathogenic mutations have been identified in *UBA1*, ensuring sufficient depth to detect lower VAF somatic mutations. As an initial diagnostic approach, low-cost genetic testing, such as exon 3 analysis of the *UBA1* gene by Sanger sequencing, can be used. This method targets the canonical mutations responsible for VEXAS, which involve p. Met41 (eg, p. Met41Thr, p.Met41Val, p.Met41Leu), and surrounding splice site mutations[1,5,8,20] (eg, c.118-1G>C). Sanger sequencing of

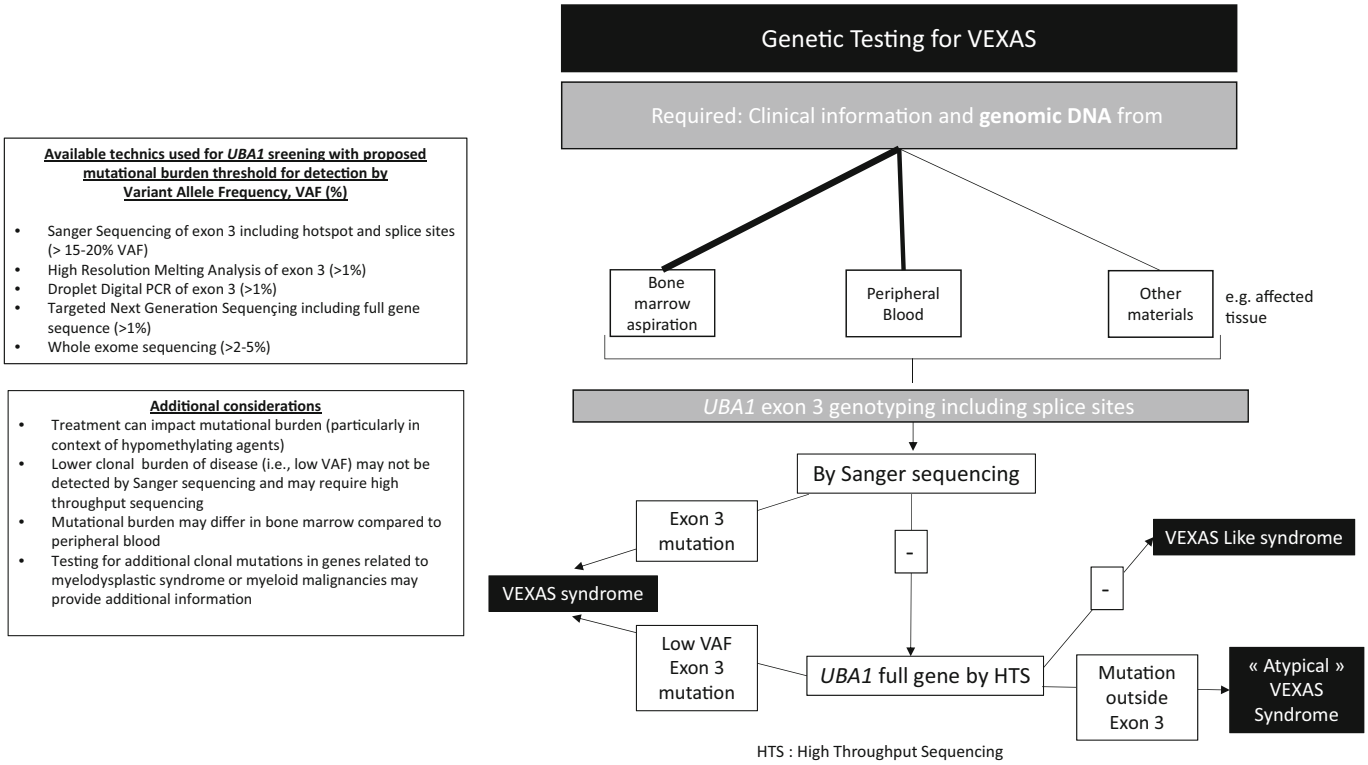


Figure 2. Genetic testing strategies for VEXAS. A practical approach to genetic testing is shown. Testing can be performed on bone marrow aspirate or peripheral blood. In rare instances when a bone marrow or peripheral blood sample is not available, testing could be performed on affected tissue. Different genetic testing strategies have different sensitivities to detect mutational burden of disease. Choice of genetic testing may be guided by access to resources. HTS, high throughout sequencing; PCR, polymerase chain reaction; VEXAS, vacuoles E1 enzyme X-linked autoinflammatory somatic syndrome.

exon 3 also covers the residue p. Ser56; mutations at this position cause reduced E1 enzyme function and are associated with a milder inflammatory phenotype of VEXAS.[23,24] Sanger sequencing can identify mutations with VAFs roughly above 20% but lacks the sensitivity to detect lower mutational burden. Although most *UBA1* mutations can be detected by Sanger at diagnosis, a subset of patients has lower VAFs that may cause clinically significant VEXAS.

More sensitive methods for genetic testing, including targeted NGS are required to detect low-level pathogenic mutations in *UBA1*. The choice of targeted NGS or Sanger sequencing will depend on factors such as test availability, costs, and clinical need for additional testing. Because VEXAS frequently overlaps with MDS and clonal hematopoiesis of uncertain significance, many patients may already be undergoing a myeloid malignancy–targeted NGS panel, which should ideally include *UBA1* along with standard myeloid malignancy genes.

Many variants of uncertain significance throughout the *UBA1* gene have recently been reported by diagnostic centers, which now routinely include *UBA1* as part of their genetic panel for investigation of hematologic malignancies.[25] The functional relevance of these variants for the pathogenesis of VEXAS and other hematologic disorders remains to be fully elucidated. Recently, mutations located outside of exon 3 have been reported to modify the catalytic function of E1 enzyme and lead to a "VEXAS-like" clinical presentation, suggesting that they might play a role in the disease pathogenesis.[15,22] The diagnosis of VEXAS will depend on identifying pathogenic or likely pathogenic mutations in *UBA1*, as determined by ACMGG/AMP American College of Medical Genetics and Genomics / Association of Molecular Pathology, which may require functional biochemical validation of such suspected mutations[26,27] in *UBA1*.

DNA source and screening methods

Somatic mutations defining VEXAS are detected in peripheral blood or bone marrow–derived DNA.

If clinical suspicion is high and genetic testing for mutations in UBA1 at exon 3 is negative in the peripheral blood, clinicians should ensure that testing covers the complete gene and consider testing a bone marrow sample.

Typically, DNA for genetic testing for VEXAS is obtained from peripheral blood or bone marrow aspirates. *UBA1* mutations originate in the bone marrow progenitors and become lineage restricted to myeloid lineages and are primarily absent in lymphocytes. Occasionally, DNA from affected skin or lymph nodes may be used, although mutation levels tend to be lower due to the

smaller number of myeloid cells infiltrating tissue. The detection of *UBA1* mutations in the blood may be compromised by use of certain treatments, particularly hypomethylating agents, such as azacitidine. Although the mutation is usually identified in the peripheral blood and mutation levels are concordant with mutation burden in the bone marrow, there are isolated cases in which mutations are detected at much higher VAFs in the bone marrow. Thus, the testing site, detection method, and previous treatments should be considered when evaluating results. If clinical suspicion for VEXAS remains high despite negative testing in peripheral blood or skin biopsy, we suggest targeted NGS of bone marrow–derived samples to ensure testing tissue most likely to be enriched with pathogenic mutations, using the most sensitive sequencing approach for low VAF mutations.

In addition to Sanger sequencing, additional methods can be used to detect mutations on exon 3 with very low VAFs. These methods include high-resolution melting analysis, ddPCR, which can target specific mutations, targeted NGS, and whole-exome sequencing to analyze the complete coding sequence of the *UBA1* gene. Additionally, in conjunction with genetic testing, flow cytometry from bone marrow samples can help to identify characteristic vacuolated myeloid cells associated with VEXAS.[28]

Diagnosis of MDS in patients with VEXAS consensus statements (Figure 3)

We recommend bone marrow examination in patients with VEXAS with associated cytopenia to exclude an associated hematologic neoplasm.

Karyotype and NGS for coexisting somatic mutations should be performed with bone marrow examination.

Although most patients with VEXAS predominantly exhibit inflammatory manifestations, a subset, especially later in the disease course, develop progressive bone marrow failure. Assessment by a hematologist is recommended to diagnose and manage various hematologic conditions including bone marrow failure, MDS, monoclonal gammopathies, and thrombosis.

Cytopenia, particularly macrocytic anemia, is a common finding in VEXAS, even in the absence of associated MDS. The need for bone marrow examination in a patient with a molecularly confirmed VEXAS diagnosis is debated due to lack of data, especially in patients without cytopenia. Macrocytosis is the most common peripheral blood finding in patients with VEXAS, followed by anemia, absolute lymphopenia (80%), moderate thrombocytopenia (30–50%), and monocytopenia (30–50%), whereas neutrophils are often within normal values. We recommend a baseline bone marrow examination in all patients with cytopenic VEXAS, especially before initiating disease-modifying treatments. When a bone marrow examination is performed, cytogenetic analysis should be included to identify additional cytogenetics abnormalities consistent with MDS, and loss of an X chromosome, which can predict VEXAS in women.[12] An NGS panel targeting somatic mutations for myeloid malignancy, including *UBA1*, is strongly recommended. As with the diagnosis of VEXAS, Sanger sequencing can be used but has limitations in detecting low VAF mutations. If a pathogenic *UBA1* mutation is identified via Sanger sequencing or targeted *UBA1* testing, patients should still undergo NGS myeloid malignancy genetic testing. This testing can be performed from bone marrow or peripheral blood, regardless of clinical phenotype, to assess for coexisting mutations. VEXAS is a chronic and progressive disease. Cytopenia may worsen over time and periodic re-evaluation with bone marrow examination may be needed to assess disease progression or treatment-related toxicity.[14,29] We recommend that patients with blood count abnormalities are followed closely by a hematologist.

Interpretation of bone marrow is challenging in VEXAS due to the frequent presence of signs of dysplasia, without meeting existing criteria for myelodysplastic syndrome diagnosis based on the World Health Organization (WHO) and International Consensus Classification (ICC) 2022 classifications.

Many patients with undiagnosed VEXAS before the discovery of the disease underwent bone marrow examinations due to varying degree of cytopenia, predominantly macrocytic anemia. Several of these patients were diagnosed with MDS due to presence of dysplasia. However, with a better understanding of the histologic findings in VEXAS, it is important to consider the diagnosis of VEXAS in the differential diagnosis of MDS when reviewing bone marrow for dysplasia.[14,30] Atypia in one or more lineages, in addition to cytoplasmic vacuolation in erythroid and myeloid precursor cells, is often found in bone marrow examinations of patients with VEXAS, even when they do not meet the diagnostic criteria for MDS.[16] Cellularity is typically high with an increased myeloid-to-erythroid ratio. Marrow flow cytometry analysis commonly shows absence of precursor B cells. In patients with a molecularly confirmed VEXAS diagnosis or with high suspicion, assessment of dysplasia should require morphologic evidence other than the presence of vacuolization alone, given this disease feature's ubiquitous presence in VEXAS.[13,14] Albeit rare, the presence of increased blasts (5–20%) or MDS-defining cytogenetic abnormalities would also indicate MDS. Importantly, marrow morphologic findings need to be evaluated in the clinical context, including the presence of cytopenia's due to other causes such as nutritional deficiencies, infections, medications, or uncontrolled inflammation. Given these diagnostic complexities, authors of a recent article employed a more stringent approach to diagnosing MDS. They required the presence of significant and persistent cytopenia, defined as transfusion-dependent anemia, and thrombocytopenia (<100K per microliter) along with morphologic or genetic (cytogenetic or molecular) findings consistent with MDS. Using this approach, the incidence of MDS in

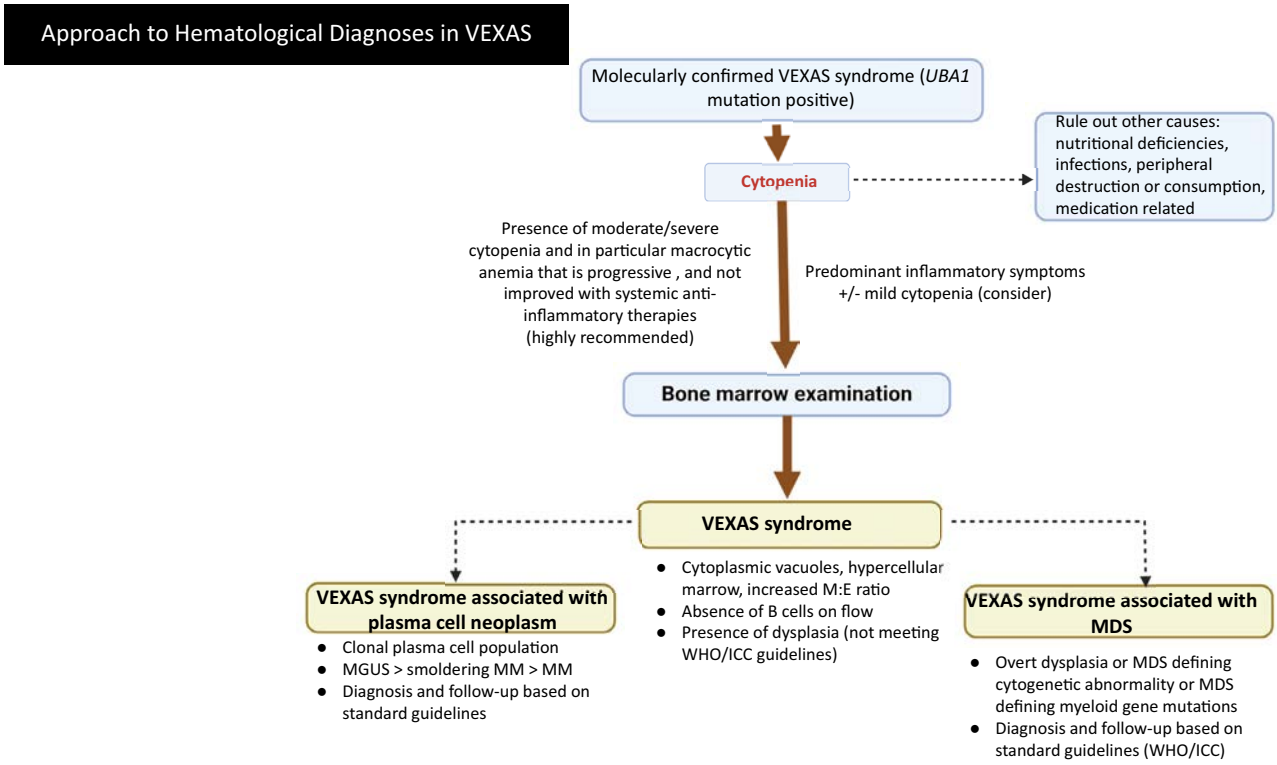


Figure 3. Approach to hematological diagnoses in VEXAS. A flow diagram is provided regarding a practical approach to hematologic evaluation of patients with a genetically confirmed diagnosis of VEXAS. ICC, International Consensus Classification; MDS, myelodysplastic syndrome; M:E, myeloid-to-erythroid; MGUS, Monoclonal Gammopathy of Undetermined Significancy; MM, multiple Myeloma; WHO, World Health Organization; VEXAS, vacuoles E1 enzyme X-linked autoinflammatory somatic syndrome.

VEXAS was 18%, lower than reported in earlier literature.[29] The importance of diagnosing MDS in a patient with an established diagnosis of VEXAS is crucial because it helps better stratify the risk of progression to acute myeloid leukemia and can inform decisions about when to consider more aggressive treatment strategies.

The somatic clonal landscape in VEXAS is dominated by mutations in DNMT3A and TET2, generally without other driver mutations, but is to date difficult to discriminate from usual age-related to more VEXAS-specific signature.

Most reported cases of MDS in VEXAS are categorized as lower risk according to International Prognostic Scoring System categorization, with a low incidence of progression to acute myeloid leukemia.[29,31–34] The karyotype is normal in more than 50% of cases. Somatic mutations in myeloid driver genes are present in nearly 60% of patients with VEXAS irrespective of a diagnosis of MDS and likely reflect the underlying inflammatory environment with predominance of *DNMT3A* and *TET2* mutations.[29,35] Most patients with VEXAS also have additional mutations of clonal hematopoiesis in conjunction with *UBA1* mutations (with *DNMT3A* and *TET2* being most frequent).[29] Splicing factor mutations have also been identified in cases with VEXAS, but their clinical significance is yet to be determined. For the previously mentioned reasons, prognostic tools generally used in MDS such as the International Prognostic Scoring System-Molecular have identified mostly lower-risk cases in the MDS-VEXAS overlap.[36] Other myeloid neoplasms such as MDS–myeloproliferative neoplasm overlap syndromes including chronic myelomonocytic leukemia, have also been reported in the literature.[37] We recommend that final diagnostic categorization for myeloid neoplasms in patients with VEXAS should follow WHO 2022 ICC guidelines.[38,39] Similarly, plasma cell neoplasms should be categorized and assessed using the 2014 International Myeloma Working Group diagnostic criteria and treated per standard therapeutic interventions.[40]

Do patients with vacuoles and no other WHO-defined dysplastic features, normal karyotype, and no other somatic mutations have MDS? For these patients, the expert consensus is not to consider these patients as having an MDS diagnosis, as the dysplastic findings may represent inherent VEXAS, but to recommend closer follow-up in case of further evolution. Based on the available data, in many cases, cytopenia appears to be related to bone marrow failure and ineffective hematopoiesis in VEXAS, distinct from MDS.

Outcome, prognosis, and management consensus statements (Figure 4)

Patients with VEXAS should be treated by a multidisciplinary team with consideration of referral to an expert center when available.

Disease activity is defined by inflammation or worsening bone marrow failure.

The goals of treatment include controlling inflammation; preventing, and treating bone marrow failure; preventing and treating secondary medical complications arising from implemented treatments; and improving quality of life.

VEXAS is a complex and potentially life-threatening disease characterized by multisystem organ involvement.[5,41] The median survival from symptom onset is approximately 10 years.[8] All patients diagnosed with VEXAS should be evaluated by a hematologist to determine the extent of hematologic involvement and manage the sequalae of progressive bone marrow failure. Rheumatologists, clinical immunologists, or internists play a key role in treating the inflammatory aspects of the disease and mitigating complications arising from immunosuppressive therapies. Dermatologists, pulmonologists, cardiologists, infectious disease specialists, and other health care providers are often needed to address organ-specific and infectious complications. Given the clinical heterogeneity and death risk associated with VEXAS, referral to an expert center or consultation with a provider who has experience treating these patients is advisable whenever possible.

Infections, thromboembolic embolic disease, and treatment-related comorbidities are common in VEXAS and can mimic disease activity.

Disease-associated and treatment-related complications contribute to the increased morbidity and mortality risk in patients with VEXAS. A wide differential diagnosis should be considered when assessing new or recurrent clinical symptoms in patients with VEXAS, as not all complaints may be directly related to disease activity. Factors such as cytopenia, advanced age, use of immunosuppressive therapy, defects in ubiquitylation, and organ-related damage to skin and mucosal barriers increase susceptibility to typical, atypical, and opportunistic infections, including *Staphylococcus aureus*, *Streptococcus pyogenes*, *Legionella species*, *nontuberculous Mycobacterium*, *Pneumocystis jirovecii* pneumonia, and *alphaherpesviruses*.[42,43] Inflammation and infection can co-occur in the same patient, often in the skin and lungs. Thromboembolic disease, primarily venous but also arterial, is prevalent in approximately 50% of patients with

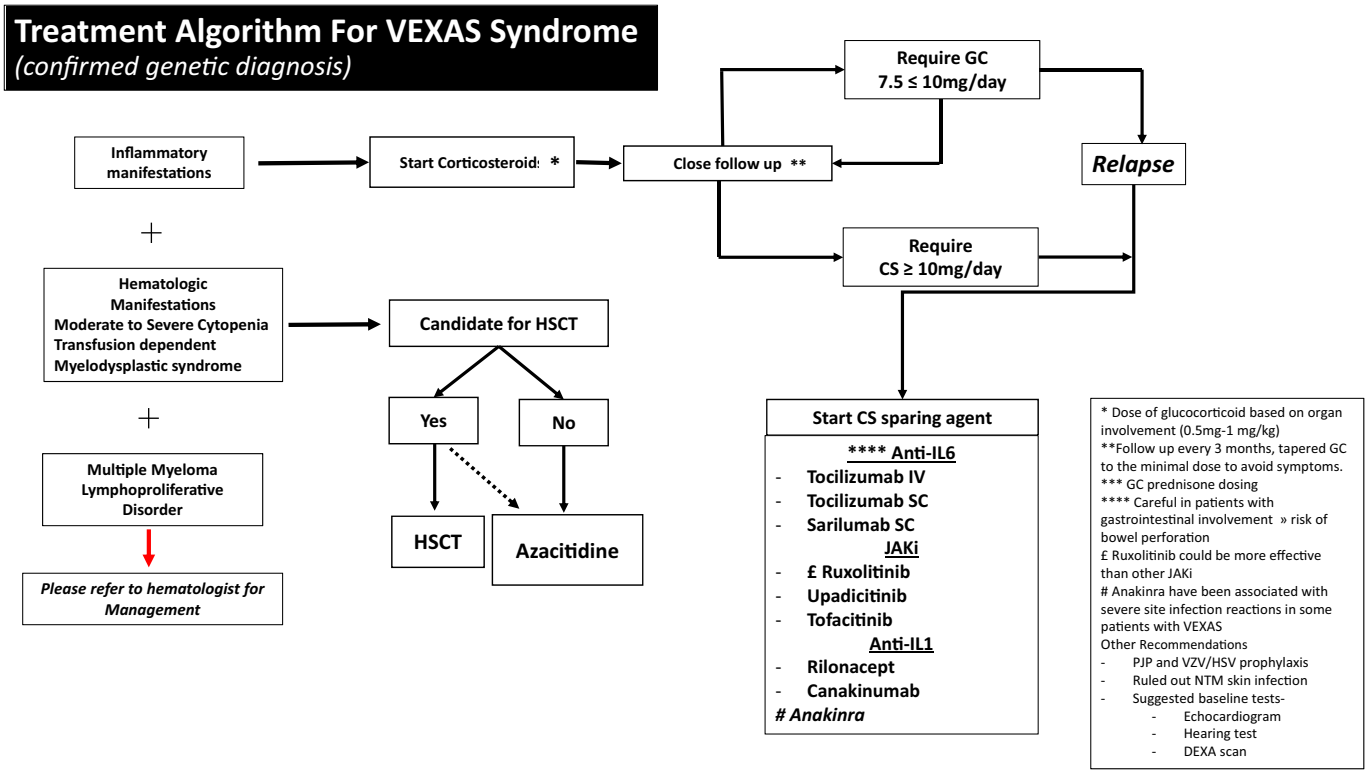


Figure 4. Treatment of VEXAS. A practical treatment algorithm is provided for clinical management of VEXAS syndrome based on current available evidence. Treatment decisions should be customized to each individual patient. CS, corticosteroid; DEXA, dual-energy x-ray absorptiometry; GC, glucocorticoid; HSCT, hematopoietic stem cell transplantation; HSV, Virus Herpes Simplex; IL-1, interleukin-1; IV, intravenous; JAKi, JAK inhibitor; NTM, nontuberculous mycobacteria; PJP, *Pneumocystis jiroveci* pneumonia; VEXAS, vacuoles E1 enzyme X-linked autoinflammatory somatic syndrome; VZV, varicella-zoster virus. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.43287/abstract.

VEXAS, typically in association with ongoing disease activity.[44] Complications from chronic immunosuppressive treatments, particularly glucocorticoids, can be related to both dose and duration of therapy.

Flare of disease can be defined as recurrence of clinical or laboratory signs or symptoms of VEXAS.

Standardized definitions of disease states and disease activity are essential for clinical and research purposes. In VEXAS, disease activity is related to inflammation and bone marrow failure. Systemic inflammation can affect nearly every organ system, with a predilection for the skin, eyes, cartilage, lungs, and blood vessels. Limited prospective data are currently available regarding the nature and frequency of flares; however, clinical manifestations tend to reoccur during flare episodes. Elevated acute phase reactants are detectable in most patients with VEXAS and can serve as potential biomarkers of disease activity. CRP may be a more reliable marker of inflammation compared to ESR in VEXAS because hematologic abnormalities can elevate ESR independent of inflammation. Progressive bone marrow failure typically manifests as cytopenia, including macrocytic anemia, thrombocytopenia, lymphopenia, and monocytopenia.[14] A subset of patients become transfusion-dependent, a factor associated with increased death.[8] Hematologic and inflammatory manifestations are often correlated but can worsen independently.

Glucocorticoids are usually needed to control inflammation in patients with VEXAS and should be tapered slowly to achieve the minimal dose required to control disease activity.

Glucocorticoids are highly effective to treat patients with VEXAS, mostly dosed orally. Moderate to high doses of daily glucocorticoids (eg, prednisone 10–25 mg/day or higher) are often needed to control inflammation. Patients may experience flares of disease when glucocorticoids are tapered below specific daily thresholds. Unlike many rheumatologic diseases, glucocorticoid-free remission is unusual in VEXAS, even with the concomitant use of steroid-sparing therapies. Caution should be exercised when tapering glucocorticoids, as small reductions in dosage can precipitate recurrence of inflammation. The goal of steroid therapy is to find the minimal glucocorticoid dose required to control disease activity while minimizing the complications of long-term therapy.

Medications that target inflammatory pathways (eg, JAK inhibitors, interleukin-6 [IL-6] inhibitors) are more effective than conventional disease-modifying antirheumatic drugs (DMARDs) (eg, methotrexate, azathioprine) or B cell–directed therapies (eg, rituximab).

Mutations in VEXAS are lineage restricted to myeloid cells (eg, monocytes, neutrophils) and produce a state of hyperinflammation involving the activation of a myriad of inflammatory pathways. Therapies that target pathways of innate immunity are more effective than medications targeted against the adaptive immune system.[1,45] Specific anticytokine therapies may be effective but are often insufficient to control inflammation. Of these, IL-6 inhibitors are the most widely used. IL-1 inhibitors can be considered, but severe injection site reactions have been reported from anakinra,[7] and few data highlighted the benefit of canakinumab. JAK inhibitors are useful to target a broader spectrum of disease-relevant cytokines (eg, IL-6, type 1 and 2 interferons). Of the various JAK inhibitors, ruxolitinib may be preferred and can be titrated across a greater dose range as needed, typically dosed at 10 mg twice a day to 20 mg twice a day.[46] Worsening cytopenia related to use of JAK inhibitors should be closely monitored and managed. Given the association of JAK inhibitors and increased risk for viral infections, thromboembolic disease, and cardiovascular events, careful selection of patients to receive this type of therapy is advisable and prophylaxis for alpha herpesviruses should be considered. Conventional DMARDs may have limited efficacy for some patients, potentially related to activation of nonmutant lymphocytes. B cell–depleting therapies are contraindicated because they are not efficacious in VEXAS and can worsen disease-associated B cell lymphopenia. Prospective clinical trials to assess comparative effectiveness and dosing among existing treatment options are needed, and the initial choice of treatment should be individualized to each patient.

Medications that reduce or eliminate the clonal burden of disease (eg, azacytidine) can be an effective treatment in some patients with VEXAS.

The elimination of clonal burden of disease is an aspirational goal for future medical therapies in VEXAS. Preliminary data on the use of azacytidine, a hypomethylating agent commonly used to treat MDS, has demonstrated reduction or elimination of *UBA1* mutation burden in peripheral blood or bone marrow in some patients with VEXAS in conjunction with improvement in inflammation.[32,33,47] Successful therapy cessation with sustained remission has been reported in small case reports, but long-term data and prospective comparisons are lacking.[47,48] Identification of specific patients with VEXAS who are most likely to benefit from medical therapies that are cytotoxic to cells harboring *UBA1* mutations remains to be determined. For example, it is currently unknown whether patients with VEXAS and concomitant MDS should be the most appropriate candidates for treatment with hypomethylating agents or whether these therapies may benefit a broader spectrum of patients.

Allogenic hematopoietic stem cell transplantation may be a curative treatment but should be reserved for select patients with VEXAS after a careful hematologic evaluation.

Because VEXAS is associated with a genetic mutation restricted to the hematopoietic compartment, allogeneic hematopoietic stem cell transplantation is a curative option for this disease.[49] Benefits of this procedure should be weighed against the morbidity

and mortality associated risks for the procedure, especially infectious complications, and graft-versus-host disease.[50] Patients with VEXAS who have progressive and severe bone marrow failure (eg, transfusion-dependent) or treatment-refractory inflammation are candidates for allogeneic stem cell transplantation. The timing of bone marrow transplantation and identification of patients who are acceptable candidates is the most challenging because of older patient age, comorbidities, and often suboptimal performance status because of disease and treatment, particularly glucocorticoids.[51] So far, 33 patients have been reported to have undergone allogenic hematopoietic cell transplantation, with 27 individuals (81.8%) alive and 11 individuals with posttransplant molecular data showing a complete eradication of *UBA1* mutation mostly over a less-than-12-month follow-up period.[52]

Prophylaxis against opportunistic infections, prevention of thromboembolic disease, and minimization of side effects related to chronic glucocorticoid therapy should be considered in all patients with VEXAS.

Strategies to mitigate disease-associated and treatment-related factors that contribute to morbidity and mortality are essential in the clinical management of patients with VEXAS. Prophylaxis against infectious complications of disease (eg, *Pneumocystis jirovecii* pneumonia or *alphaherpesviruses*) should be considered in all patients, with particular consideration for older patients and any patient with severe cytopenia who are receiving immunosuppressive therapies.[42,43] For skin lesions that do not respond to treatment, a skin biopsy should be considered to rule out nontuberculous *Mycobacterium* infections before increasing immunosuppression. Given the high risk of thromboembolic disease, patients with VEXAS should receive thromboprophylaxis in high-risk settings (eg, hospitalization) unless contraindicated.[44] In patients who develop thromboembolic complications of disease, duration and choice of anticoagulation therapy has not been defined. Given the substantial exposure to glucocorticoids experienced by most patients with VEXAS, active management of steroid-related complications, with particular focus on strategies to promote bone health and prevent cardiovascular risk, is critical.

We need future clinical trials to identify the best treatment for patients with VEXAS.

Several recent reports have detailed clinical response to various treatments for VEXAS. Most of these studies provide case report– or case series–level data. Data from prospective observational cohort studies are lacking, and randomized controlled trials have not been performed to date. The quality of data to support treatment considerations in VEXAS is therefore low primarily due to selection bias, insufficient follow-up time and precision in the diagnosis of associated MDS, and a lack of standardized definitions of treatment response. Efficacy of any medication beyond glucocorticoids has not been definitively established. Furthermore, because VEXAS is typically a progressive clonal disease of the bone marrow, the timing of therapy initiation, in addition to the specific therapeutic strategy, may have an important impact on clinical response and will need to be considered in future therapeutic trials.

VEXAS is a newly recognized disease, and these clinical guidance statements are intended to help clinicians better manage this condition. The statements are based on expert opinion, serving as a guide for current clinical practice and highlighting areas of ongoing investigation. Prospective studies are necessary to validate the different strategies outlined. As evidence accumulates related to the management of VEXAS, future efforts to provide evidence-based clinical guidelines, with input from patient partners, will be important to improve clinical care and promote further research to better define the disease.

ACKNOWLEDGMENTS

We would like to thank FAI2R for their support of the 1st international VEXAS workshop. We would like to thank patients and family members for participating in research. Biorender was used for figure design. This guidance has been endorsed by ISSAID society. HM: VEXAS research supported by: Ministry of Health of the Czech Republic, Agency for Health Research ČR (AZV) grant Nr. NU23-10-00160. [Correction added on 27 August 2025, after first online publication: Grant support "HM: VEXAS research supported by: Ministry of Health of the Czech Republic, Agency for Health Research ČR (AZV) grant Nr. NU23-10-00160" was added to the Acknowledgments section.]

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding authors, Drs Mekinian and Beck confirm that all authors have provided the final approval of the version to be published and take responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. Beck DB, Ferrada MA, Sikora KA, et al. Somatic mutations in *UBA1* and severe adult-onset autoinflammatory disease. N Engl J Med 2020;383(27):2628–2638.
2. Beck DB, Bodian DL, Shah V, et al. Estimated prevalence and clinical manifestations of UBA1 variants associated with VEXAS syndrome in a clinical population. JAMA 2023;329(4):318–324.
3. Guyatt GH, Oxman AD, Vist GE, et al; GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008;336(7650):924–926.
4. Andrews J, Guyatt G, Oxman AD, et al. GRADE guidelines: 14. Going from evidence to recommendations: the significance and presentation of recommendations. J Clin Epidemiol 2013;66(7):719–725.
5. Georgin-Lavialle S, Terrier B, Guedon AF, et al; French VEXAS group; GFEV, GFM, CEREMAIA, MINHEMON. Further characterization of clinical and laboratory features in VEXAS syndrome: large-scale analysis of a multicentre case series of 116 French patients. Br J Dermatol 2022;186(3):564–574.
6. Vitale A, Caggiano V, Martín-Nares E, et al. Orbital/ocular inflammatory involvement in VEXAS syndrome: data from the international AIDA network VEXAS registry. Semin Arthritis Rheum 2024;66:152430.

7. Tan IJ, Ferrada MA, Ahmad S, et al. Skin Manifestations of VEXAS syndrome and associated genotypes. JAMA Dermatol 2024;160(8): 822–829.
8. Ferrada MA, Savic S, Cardona DO, et al. Translation of cytoplasmic UBA1 contributes to VEXAS syndrome pathogenesis. Blood 2022; 140(13):1496–1506.
9. van der Made CI, Potjewijd J, Hoogstins A, et al. Adult-onset autoinflammation caused by somatic mutations in UBA1: a Dutch case series of patients with VEXAS. J Allergy Clin Immunol 2022;149(1): 432–439e4.
10. Tsuchida N, Kunishita Y, Uchiyama Y, et al. Pathogenic UBA1 variants associated with VEXAS syndrome in Japanese patients with relapsing polychondritis. Ann Rheum Dis 2021;80(8):1057–1061.
11. Sánchez-Hernández BE, Calderón-Espinoza I, Martín-Nares E. Challenging the paradigm: a case of early-onset VEXAS syndrome. Rheumatology (Oxford) 2024;63(3):e99–e100.
12. Echerbault R, Bourguiba R, Georgin-Lavialle S, et al. Comparing clinical features between males and females with VEXAS syndrome: data from literature analysis of patient reports. Rheumatology (Oxford) 2024;63(10):2694–2700.
13. Olteanu H, Patnaik M, Koster MJ, et al. Comprehensive morphologic characterization of bone marrow biopsy findings in a large cohort of patients with VEXAS syndrome: a single-institution longitudinal study of 111 bone marrow samples from 52 patients. Am J Clin Pathol 2024;161(6):609–624.
14. Obiorah IE, Patel BA, Groarke EM, et al. Benign and malignant hematologic manifestations in patients with VEXAS syndrome due to somatic mutations in UBA1. Blood Adv 2021;5(16):3203–3215.
15. Gurnari C, Pagliuca S, Durkin L, et al. Vacuolization of hematopoietic precursors: an enigma with multiple etiologies. Blood 2021;137(26): 3685–3689.
16. Rabut A, Jasserand L, Richard C, et al. Quantitative assessment of vacuolization of myeloid precursors in VEXAS syndrome. HemaSphere 2023;7(2):e828.
17. Hines AS, Koster MJ, Bock AR, et al. Targeted testing of bone marrow specimens with cytoplasmic vacuolization to identify previously undiagnosed cases of VEXAS syndrome. Rheumatology (Oxford) 2023;62(12):3947–3951.
18. Lacombe V, Hadjadj J, Georgin-Lavialle S, et al. Vacuoles in bone marrow progenitors: VEXAS syndrome and beyond. Lancet Haematol 2024;11(2):e160–e167.
19. Zakine E, Schell B, Battistella M, et al. UBA1 Variations in neutrophilic dermatosis skin lesions of patients with VEXAS syndrome. JAMA Dermatol 2021;157(11):1349–1354.
20. Mascaro JM, Rodriguez-Pinto I, Poza G, et al. Spanish cohort of VEXAS syndrome: clinical manifestations, outcome of treatments and novel evidences about *UBA1* mosaicism. Ann Rheum Dis 2023; 82(12):1594–1605.
21. Ospina Cardona D, Rodriguez-Pinto I, Iosim S, et al. Description of a novel splice site variant in UBA1 gene causing VEXAS syndrome. Rheumatology (Oxford) 2024;63(10):2897–2902.
22. Collins JC, Magaziner SJ, English M, et al. Shared and distinct mechanisms of UBA1 inactivation across different diseases. EMBO J 2024; 43(10):1919–1946.
23. Poulter JA, Collins JC, Cargo C, et al. Novel somatic mutations in UBA1 as a cause of VEXAS syndrome. Blood 2021;137(26): 3676–3681.
24. Al-Hakim A, Kulasekararaj A, Norouzi M, et al. S56F UBA1 variant is associated with haematological predominant subtype of VEXAS. Br J Haematol 2023;203(2):331–335.
25. Sakuma M, Blombery P, Meggendorfer M, et al. Novel causative variants of VEXAS in UBA1 detected through whole genome transcriptome sequencing in a large cohort of hematological malignancies. Leukemia 2023;37(5):1080–1091.
26. Richards S, Aziz N, Bale S, et al; ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 2015;17(5):405–424.
27. Stiburkova B, Pavelcova K, Belickova M, et al. Novel Somatic UBA1 variant in a patient with VEXAS syndrome. Arthritis Rheumatol 2023; 75(7):1285–1290.
28. Ding Y, Dulau-Florea AE, Groarke EM, et al. Use of flow cytometric light scattering to recognize the characteristic vacuolated marrow cells in VEXAS syndrome. Blood Adv 2023;7(20):6151–6155.
29. Gutierrez-Rodrigues F, Kusne Y, Fernandez J, et al. Spectrum of clonal hematopoiesis in VEXAS syndrome. Blood 2023;142(3): 244–259.
30. Patel N, Dulau-Florea A, Calvo KR. Characteristic bone marrow findings in patients with UBA1 somatic mutations and VEXAS syndrome. Semin Hematol 2021;58(4):204–211.
31. Bourbon E, Heiblig M, Gerfaud Valentin M, et al. Therapeutic options in VEXAS syndrome: insights from a retrospective series. Blood 2021;137(26):3682–3684.
32. Comont T, Heiblig M, Rivière E, et al; French VEXAS study group, Groupe Francophone des Myélodysplasies (GFM) and MedecineINterne, HEmato et ONco (MINHEMON) group. Azacitidine for patients with vacuoles, E1 enzyme, X-linked, autoinflammatory, somatic syndrome (VEXAS) and myelodysplastic syndrome: data from the French VEXAS registry. Br J Haematol 2022;196(4):969–974.
33. Mekinian A, Zhao LP, Chevret S, et al. A phase II prospective trial of azacitidine in steroid-dependent or refractory systemic autoimmune/inflammatory disorders and VEXAS syndrome associated with MDS and CMML. Leukemia 2022;36(11):2739–2742.
34. Battipaglia G, Vincenzi A, Falconi G, et al. New scenarios in vacuoles, E1 enzyme, X linked, autoinflammatory, somatic (VEXAS) syndrome: evolution from myelodysplastic syndrome to acute myeloid leukemia. Curr Res Transl Med 2023;71(2):103386.
35. Sirenko M, Bernard E, Creignou M, et al. Molecular and clinical presentation of UBA1-mutated myelodysplastic syndromes. Blood 2024;144(11):1221–1229.
36. Gurnari C, Pascale MR, Vitale A, et al. Diagnostic capabilities, clinical features, and longitudinal UBA1 clonal dynamics of a nationwide VEXAS cohort. Am J Hematol 2024;99(2):254–262.
37. Zhao LP, Schell B, Sébert M, et al. Prevalence of UBA1 mutations in MDS/CMML patients with systemic inflammatory and auto-immune disease. Leukemia 2021;35(9):2731–2733.
38. Arber DA, Orazi A, Hasserjian RP, et al. International Consensus Classification of myeloid neoplasms and acute leukemias: integrating morphologic, clinical, and genomic data. Blood 2022;140(11): 1200–1228.
39. Lee WH, Lin CC, Tsai CH, et al. Comparison of the 2022 World Health Organization classification and International Consensus Classification in myelodysplastic syndromes/neoplasms. Blood Cancer J 2024; 14(1):57.
40. Rajkumar SV, Dimopoulos MA, Palumbo A, et al. International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma. Lancet Oncol 2014;15(12):e538–e548.
41. Grayson PC, Patel BA, Young NS. VEXAS syndrome. Blood 2021; 137(26):3591–3594.
42. Czech M, Cuellar-Rodriguez J, Patel BA, et al. Opportunistic infections, mortality risk, and prevention strategies in patients with vacuoles, E1 enzyme, X-linked, autoinflammatory, somatic (VEXAS) syndrome. *Open Forum* Infect Dis 2024;11(7):ofae405.

43. de Valence B, Delaune M, Nguyen Y, et al; French VEXAS Group. Serious infections in patients with VEXAS syndrome: data from the French VEXAS registry. Ann Rheum Dis 2024;83(3):372–381.
44. Kusne Y, Ghorbanzadeh A, Dulau-Florea A, et al. Venous and arterial thrombosis in patients with VEXAS syndrome. Blood 2024;143(21): 2190–2200.
45. Hadjadj J, Nguyen Y, Mouloudj D, et al; FRENVEX. Efficacy and safety of targeted therapies in VEXAS syndrome: retrospective study from the FRENVEX. Ann Rheum Dis 2024;83(10):1358–1367.
46. Heiblig M, Ferrada MA, Koster MJ, et al. Ruxolitinib is more effective than other JAK inhibitors to treat VEXAS syndrome: a retrospective multicenter study. Blood 2022;140(8):927–931.
47. Aalbers AM, van Daele PLA, Dalm VASH, et al. Long-term genetic and clinical remissions after cessation of azacitidine treatment in patients with VEXAS syndrome. HemaSphere 2024;8(8):e129.
48. Sockel K, Götze K, Ganster C, et al. VEXAS syndrome: complete molecular remission after hypomethylating therapy. Ann Hematol 2024;103(3):993–997.
49. Mohty R, Reljic T, Abdel-Razeq N, et al. Assessing the efficacy of allogeneic hematopoietic cell transplantation in VEXAS syndrome: results of a systematic review and meta-analysis. Bone Marrow Transplant 2024;59(10):1423–1427.
50. Gurnari C, Koster L, Baaij L, et al. Allogeneic hematopoietic cell transplantation for VEXAS syndrome: results of a multicenter study of the EBMT. Blood Adv 2024;8(6):1444–1448.
51. Dias AL, Groarke EM, Hickstein D, Patel B. Role of allogeneic hematopoietic cell transplantation in VEXAS syndrome. Ann Hematol 2024; 103(11):4427–4436.
52. Ali SB, Gurnari C. Allogenic haematopoietic stem cell transplantation in VEXAS: a review of 33 patients. Clin Rheumatol 2024;43(11): 3565–3575.

APPENDIX A: MEMBERS OF VEXAS WORKING GROUPS

The members of the VEXAS working groups are as follows: group 1: Heřman Mann, Benjamin Terrier, François Chasset, Sophie Georgin Lavialle, Alessandro Tomelleri, Campochiaro Corrado, Carlo Salvarani, Francesca Crisafulli, Franco Franceschini, Micol Frassi, Yohei Kirino, Ewa Więsik-Szewczyk, Marcin Milchert, Raquel Faria, Ciprian Jurcut, Joerg Seebach, Sinisa Savic, David Beck, Lauren Madigan, Matthew Koster, Patnaik Mrinal; group 2: Olivier Kosmider, Pierre Sujobert, Juan I. Arostegui, Nakajima, Catherine Cargo, David Viswanatha; group 3: Yervand Hagopian, Mael Heilblig, Pierre Fenaux, Thibault Comont, Bruno Alessandro, Chiara Cattaneo, Elisa Diral, Sara Morais, Austin Kulasekarara, Emma Groarke, Katherine Calvo, Patel Bhavisha; group 4: Anthony Sammel, Arsene Mekinian, Benjamin Terrier, Marie Robin, Sophie Georgin Lavialle, Katja Sockel, Yvan Jamilloux, Carmelo Gurnari, Henrique Coelho, Romana Vieira, Rim Bourguiba, Marcela Ferrada, Peter Grayson.

Arthritis & Rheumatology
Vol. 78, No. 3, March 2026, pp 523–536
DOI 10.1002/art.43218

American College of Rheumatology
Empowering Rheumatology Professionals

EULAR/American College of Rheumatology Risk Stratification Criteria for Development of Rheumatoid Arthritis in the Risk Stage of Arthralgia

Hanna W. van Steenbergen,[1] Frank Doornkamp,[2] Stefano Alivernini,[3] John Backlund,[4] Catalin Codreanu,[5] Stanley B. Cohen,[6] Bernard Combe,[7] Andrew P. Cope,[8] Kevin D. Deane,[9] Bryant R. England,[10] Marie Falahee,[11] Pascal H. P. de Jong,[12] Arnd Kleyer,[13] Diane Lacaille,[14] Bertha Maat,[15] Kulveer Mankia,[16] Elise van Mulligen,[17] György Nagy,[18] Liam J. O'Neil,[19] Linda Rodamaker,[20] Ilfita Sahbudin,[21] Dirkjan van Schaardenburg,[22] Alexandre Sepriano,[23] Jose A. P. da Silva,[24] Lukas De Smet,[1] Jeffrey A. Sparks,[25] Ewout W. Steyerberg,[2] Paul Studenic,[26] Elisabeth Wethington,[27] Robert L. Landewé,[28] Karim Raza,[29] and Annette H. M. van der Helm-van Mil[30]

This criteria set has been approved by the American College of Rheumatology (ACR) Board of Directors and the EULAR Executive Committee. This signifies that the criteria set has been quantitatively validated using patient data, and it has undergone validation based on an independent data set. All ACR/EULAR-approved criteria sets are expected to undergo intermittent updates.

Classification criteria are essential in clinical and basic science research because they allow investigators to study relatively homogeneous populations of patients recruited from a single or multiple research sites. In clinical settings, diagnoses are made by health care professionals evaluating an individual patient's symptoms, signs, and results of laboratory and imaging studies in order to guide therapeutic recommendations. Patients diagnosed with a particular disease may or may not fulfill classification criteria for that disease. Classification criteria, in the hands of an experienced clinician with expertise in rheumatology, may inform a diagnostic evaluation, but improperly applied classification criteria may lead to misdiagnosis.

The ACR is an independent, professional, medical, and scientific society that does not guarantee, warrant, or endorse any commercial product or service.

Objective. The field of rheumatoid arthritis (RA) is moving towards identification of and intervention in people at risk of RA, but a validated risk stratification method is lacking. This work was undertaken to develop a risk stratification method for persons presenting with arthralgia considered to be at risk of RA.

Methods. A joint EULAR/American College of Rheumatology (ACR) expert committee was established. Risk factor and outcome data from 10 arthralgia cohorts (including clinically suspect arthralgia and autoantibody-positive arthralgia) were studied. The work focused on differentiating the risk of progression to clinically apparent inflammatory arthritis (IA) within 1 year, using clinical and serologic variables, without and with subclinical joint inflammation detected by ultrasound (US) or magnetic resonance imaging (MRI). Developing RA according to the 2010 EULAR/ACR criteria within 1 year was a secondary outcome. A set of validated risk stratification criteria was developed.

Results. Using data from 2,293 symptomatic at-risk individuals, a stratification method was derived consisting of 6 clinical and serologic variables (morning stiffness, patient-reported joint swelling, difficulty making a fist, C-reactive protein, rheumatoid factor, and anti-citrullinated peptide antibody) yielding an area under the curve (AUC) of 0.80 (95% confidence interval [CI] 0.77–0.83) for IA development. The inclusion of US variables did not increase the discriminative ability. When MRI-detected subclinical inflammation variables were included, the AUC was 0.87 (95% CI 0.82–0.90).

In the presence of clinical, serologic, and MRI variables, a sensitivity and specificity of >75% was achieved. For RA development, the AUC of the criteria with MRI was 0.93 (95% CI 0.90–0.97).

Conclusions. EULAR/ACR risk stratification criteria have been developed for people with arthralgia in secondary care who are considered at risk for RA. The criteria can be applied in the absence or presence of imaging data and have been developed to define homogeneous risk groups for future prevention trials.

INTRODUCTION

Rheumatoid arthritis (RA) is among the most common chronic autoimmune diseases. The presence of clinically apparent synovitis (swollen joints) is critical to diagnose and classify RA.[1] It has been established that early therapeutic interventions for RA improve clinical outcomes and reduce the progression of joint damage and disability.[2] Furthermore, it is well recognized that autoimmune processes may be aberrant long before clinical arthritis first develops.[3–5] This has led to a focus on the prearthritis stages of RA, under the hypothesis that these early stages are more amenable to disease-modifying interventions.

The transition from 'health' to RA has been divided into several stages: genetic and environmental risk factors, evidence of systemic autoimmunity, and symptoms, followed by clinically apparent inflammatory arthritis (IA) onset.[6] The symptomatic risk stage was defined using a data-driven approach with clinical expertise for reference, resulting in the EULAR definition of arthralgia suspicious for progression to RA.[7,8] This definition (consisting of symptoms and signs) is useful to distinguish arthralgia suspicious for progression to RA from other kinds of arthralgia. Many groups have initiated observational cohort studies of arthralgia at risk for RA, either in persons with clinically suspected arthralgia (CSA) or in persons with autoantibodies and arthralgia/musculoskeletal (MSK) symptoms. A recent EULAR taskforce summarized existing biomarker data and concluded that there was not yet consensus on which combination of predictors was most informative and that proper validation was lacking.[9,10] A uniform, accurate, and validated method for risk stratification in individuals with arthralgia is therefore warranted.

This article is simultaneously published in *Annals of the Rheumatic Diseases*.

Supported by EULAR and the American College of Rheumatology (ACR) (QoC 017).

Patient involvement: patient research partners from EULAR and ACR were part of the taskforce and included as authors on this work.

[1]Hanna W. van Steenbergen, Lukas De Smet: Department of Rheumatology, Leiden University Medical Centre, Leiden, The Netherlands; [2]Frank Doornkamp, Ewout W. Steyerberg: Department of Biomedical Data Sciences, Leiden University Medical Centre, Leiden, The Netherlands; [3]Stefano Alivernini: Division of Rheumatology, Fondazione Policlinico Universita'rio A. Gemelli Istituto di Ricovero e Cura a Carattere; [4]John Backlund: At-risk participant; [5]Catalin Codreanu: Center for Rheumatic Diseases, University of Medicine and Pharmacy, Bucharest, Romania; [6]Stanley B. Cohen: University of Texas Southwestern Medical School, Dallas; [7]Bernard Combe: Montpellier University, Montpellier, France; [8]Andrew P. Cope: Centre for Rheumatic Diseases, School of Immunology and Microbial Sciences, King's College London, London, UK; [9]Kevin D. Deane University of Colorado Anschutz Medical Campus, Aurora; [10]Bryant R. England: Veterans Affairs Nebraska-Western Iowa Health Care System & University of Nebraska Medical Center, Omaha; [11]Marie Falahee: Department of Inflammation and Ageing, School of Infection, Inflammation and Immunology, College of Medicine and Health, University of Birmingham, Birmingham, UK; [12]Pascal H. P. de Jong: Department of Rheumatology, Erasmus Medical Centre, Rotterdam, The Netherlands; [13]Arnd Kleyer: Department of Rheumatology and Clinical Immunology, Charite'-Universitatsmedizin Berlin, Berlin, Germany, and Department of Internal Medicine 3, Rheumatology and Immunology University Hospital Erlangen and Friedrich Alexander University (FAU), Erlangen, Germany; [14]Diane Lacaille: Arthritis Research Canada, Division of Rheumatology, Department of Medicine, University of British Columbia, Vancouver, British Columbia, Canada; [15]Bertha Maat: Patient Research Partner European Alliance of Associations for Rheumatology People with Arthritis and Rheumatism in Europe and National Association ReumaZorg Nederland, The Netherlands; [16]Kulveer Mankia: Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds and National Institute for Health and Care Research Leeds Biomedical Research Centre, Leeds Teaching Hospitals National Health Services Trust, Leeds, UK; [17]Elise van Mulligen: Department of Rheumatology, Leiden University Medical Centre, Leiden, and Department of Rheumatology, Erasmus Medical Centre, Rotterdam, The Netherlands; [18]György Nagy: Department of Rheumatology and Clinical Immunology, Department of Internal Medicine and Oncology, Heart and Vascular Center, Department of Genetics, Cell- and Immunobiology, Semmelweis University, and National Institute of Locomotor Diseases and Disabilities, Budapest, Hungary; [19]Liam J. O'Neil: Department of Internal Medicine, University of Manitoba, Winnipeg, Manitoba, Canada; [20]Linda Rodamaker: Department of Rheumatology, University of Colorado School of Medicine, Aurora; [21]Ilfita Sahbudin: Department of Inflammation and Ageing, School of Infection, Inflammation and Immunology, College of Medicine and Health, University of Birmingham, Birmingham, and National Institute for Health and Care Research Birmingham Biomedical Research Center and Clinical Research Facility, University of Birmingham, Queen Elizabeth Hospital, Birmingham, UK; [22]Dirkjan van Schaardenburg: Amsterdam Rheumatology Center, Reade, Amsterdam, The Netherlands; [23]Alexandre Sepriano: Department of Rheumatology, Leiden University Medical Centre, Leiden, The Netherlands, and Nova Medical School, Universidade Nova de Lisboa, Lisboa, Portugal; [24]Jose A. P. da Silva: Department of Rheumatology, Centro Hospitalar e Universita'rio de Coimbra, Institute for Clinical and Biomedical Research, Faculty of Medicine, University of Coimbra, Portugal; [25]Jeffrey A. Sparks: Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts; [26]Paul Studenic: Department of Internal Medicine 3, Division of Rheumatology, Medical University of Vienna, Vienna, Austria, and Department of Medicine (Solna), Division of Rheumatology, Karolinska Institute, Stockholm, Sweden; [27]Elisabeth Wethington: Patient research partner, Department of Information Services and Technology, National Jewish Health, Denver, Colorado; [28]Robert L. Landewé: Rheumatology & Clinical Immunology, Amsterdam University Medical Center, Amsterdam, and Department of Rheumatology, Zuyderland Medical Center, Heerlen, The Netherlands; [29]Karim Raza: Department of Inflammation and Ageing, School of Infection, Inflammation and Immunology, College of Medicine and Health, University of Birmingham, Birmingham, and Department of Rheumatology, Bronglais Hospital, Hywel Dda University Health Board, Aberystwyth, UK; [30]Annette H. M. van der Helm-van Mil: Department of Rheumatology, Leiden University Medical Centre, Leiden, and Department of Rheumatology, Erasmus Medical Centre, Rotterdam, The Netherlands.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/art.43218).

Address correspondence via email to Annette H. M. van der Helm, at a.h.m.van_der_helm@lumc.nl.

Submitted for publication May 1, 2025; accepted May 1, 2025.

This need is fueled by the results of recent randomized placebo-controlled trials that suggest that therapeutic interventions in CSA or autoantibody-positive arthralgia can modify the disease course.[11–13] To help design future trials efficiently, effective risk stratification is important to identify homogeneous groups when evaluating the effectiveness of interventions. In addition, at-risk individuals themselves have indicated that reliable risk information is pivotal for their interpretation of symptoms and decision-making.[14–16] Finally, accurate and accepted risk stratification is required by regulatory agencies. Therefore, to support and homogenize further studies in the pre-RA field, a EULAR and American College of Rheumatology (ACR) expert committee was formed to develop an accurate and validated risk stratification algorithm. An important consideration was that the method should be applicable regardless of the method used to identify symptomatic persons at risk (clinical grounds [CSA] or autoantibody-positive arthralgia) and regardless of the availability of imaging modalities for detecting subclinical joint inflammation. Second, it was considered important to ensure that the generated risk stratification criteria would not be misinterpreted as indicating the presence of a disease or as classification criteria. With these overarching principles, an expert committee developed EULAR/ACR risk stratification criteria.

METHODS

Expert committee. In 2020, a EULAR committee of experts in the field of 'arthralgia at risk of RA' was established. To pool existing cohort data and develop consensus, members were invited because they were the principal investigator of an arthralgia cohort or because of their expertise in this area. In 2023, it was decided by ACR and EULAR to include ACR experts and to continue as a joint EULAR/ACR initiative. The final expert committee consisted of 20 rheumatologists, 2 fellows, 1 EMEUNET representative, 4 patient research partners, 2 allied health professionals, 1 methodologist, and 2 statisticians, recruited from 10 European countries and North America.

Data sharing procedure. An inventory of available data in at-risk cohorts was made (including inclusion criteria, endpoints, and risk variables). At the first meeting (June 2021), consensus was reached on the variables to be shared and on the primary endpoint. At the second meeting (June 2022), data transfer agreements were signed and some data were shared. The final datasets were shared in October 2022. The principal investigators of the cohorts shared their cohort data for this project, but data will not be openly accessible to the general public.

Study population. Some cohorts included persons defined as being at risk for RA because of autoantibody positivity with any arthralgia/MSK symptoms; others were selected based on the rheumatologists' expert opinion that the person had arthralgia characteristics that put them at high risk of progressing to RA (ie, CSA). The expert committee agreed to include both approaches. Additionally, it was decided to exclude cohorts in which at least some individuals were treated with disease-modifying antirheumatic drugs in the phase of arthralgia/MSK symptoms. The cohorts are described in Supplemental Tables S1 and S2. By definition, all at-risk individuals studied did not have clinically apparent IA at inclusion in the cohorts. Ethical approval was ensured by the individual cohorts.

Endpoint. The primary endpoint was the development of clinically apparent IA, identified by physical examination by rheumatologists, at 1-year follow-up. This was collected across all cohorts. Secondary endpoints were fulfilling of the 2010 EULAR/ACR criteria for RA[1] at 1 year and developing IA at 2 years follow-up.

Risk variables. Three categories of predictor variables were studied: clinical and serologic (available in all cohorts), and, where available, subclinical inflammation detected by ultrasound (US) or magnetic resonance imaging (MRI) (Supplemental Table S1). Clinical variables were studied as available (31 variables, Supplemental Figure S1A). The presence of CSA, based on clinical opinion or EULAR definition,[7] was not included as a variable because it was missing from autoantibody-positive arthralgia cohorts. However, data on several clinical items from the EULAR definition were available and examined. Imaging protocols differed between the cohorts (Supplemental Table S3 for US, Supplemental File S1 for MRI). Anatomical sites and characteristics that were reasonably similar across cohorts were included (92 variables for US and 62 for MRI, Supplemental Figure S1B and S1C). The presence of subclinical inflammation by US was defined as grey scale (GS) ≥2 (except for metatarsophalangeal [MTP] joints 1–3, where a threshold of ≥3 was chosen because of the prevalence of GS ≤2 in healthy persons[17]) or power Doppler ≥1. For MRI, subclinical inflammation was considered present if the Rheumatoid Arthritis Magnetic Resonance Imaging Score at a joint/bone/tendon sheath was present in <5% of age-matched symptom-free persons at the same location.[18,19]

Statistics. A statistical analysis plan was developed and approved.

Derivation. Missing variables were imputed using the total dataset (including available clinical, serologic, imaging, and outcome data). Twenty completed datasets were created using multiple imputation by chained equation.

Because of a high number of variables in relation to the number of events, the number of variables was reduced prior to the main analyses to prevent overfitting. Thirty-one clinical items were present. Some were not included, eg, because of absence of

association in univariable logistic analyses of the primary endpoint, leaving 24 clinical and serologic variables for the main analyses (Supplemental Figure S1A). For US, data from 92 variables were present across the cohorts (Supplemental Figure S1B). Summation reduced this to 10 US variables without loss of discriminatory capacity (area under the receiver operating characteristic curve [AUC]) (Supplemental Figure S1B). Similarly, 62 MRI variables were summed into 12 variables (Supplemental Figure S1C).

For the main analyses, least absolute shrinkage and selection operator (Lasso) penalized regression was used. First, a Lasso logistic regression model was built with 24 clinical and serologic variables and the primary endpoint as outcome in the total dataset of all cohorts for application in the absence of imaging data. The regression coefficients of the selected variables were fixed into a linear predictor (ie, clinical and serologic variables with fixed coefficients). Subsequently, this linear predictor was included in 2 other analyses, including 10 baseline US or 12 MRI variables. These analyses assessed the incremental value of US-/MRI-detected inflammation added to the information from the clinical and serologic variables.

Cohort heterogeneity was evaluated by adding a cohort variable as an adjustment variable. Two cohort variables were studied: grouped in 2 categories (identification of being at risk based on autoantibody-positive arthralgia/MSK symptoms or CSA) or 5 categories (grouping by geography/region and identification method).

A risk stratification algorithm was derived by summing the coefficients of the identified items with scaling and rounding for better applicability. This consisted of a section with clinical and serologic variables, and if US or MRI data were available, these additional variables could be used. The predicted risks were plotted against the risk scores. Test characteristics (sensitivity and specificity) and predictive values (positive and negative predictive values) were determined for different cutoffs of risk scores.

Discriminatory capacity of the 3 models (clinical and serologic, +US, and +MRI) was assessed primarily by the AUC, with higher AUCs indicating better performance. We evaluated whether a sensitivity and specificity of ~80% (a priori selected criterion) was achieved.

We assessed whether risk stratification could be simplified without loss of performance (AUC). Additionally, we compared the percentage of participants with low, intermediate, or high risk where risk categories were arbitrarily defined (low risk <25% development of clinically apparent IA; intermediate risk 25%–75%; high risk ≥75%). The goal was to have as few people as possible in the intermediate risk group. Statistical methods are further described in Supplemental File S2.

Validation. We applied a cross-validation procedure with 200 bootstrap replications. In addition, the total dataset was split into 2 sets comprising two-thirds and one-third; the split was performed at the cohort and outcome levels. Finally, the performance per each of 5 groups of cohorts was evaluated.

Final consensus. The results were presented at the third meeting (in-person, at EULAR congress 2023) and at the final online meeting (June 2023). Consensus was then assessed by voting according to the EULAR standardized operating procedures.[20]

RESULTS

Population. Of the 14 cohorts in the inventory phase, 10 cohorts generated data from a total of 2,583 symptomatic at-risk persons (Supplemental Figure S2, Supplemental Table S1). These data came from observational cohorts from Amsterdam (The Netherlands), Birmingham (UK), Erlangen (Germany), Leeds (UK), Leiden (The Netherlands), Rotterdam (The Netherlands), Rome (Italy), and Vienna (Austria) and the placebo arm of the Dutch multicenter TREAT EARLIER trial. Two cohorts were excluded because of disease-modifying antirheumatic drug treatment, leaving 2,293 persons for analysis (Table 1, Supplemental File S2).

Primary endpoint. Clinically apparent IA developed in 389 persons (17%) (282 anti-citrullinated peptide antibody [ACPA]-positive, 107 ACPA-negative persons) within 1 year. Clinical, serologic, and imaging factors were studied in relation to this endpoint.

Clinical and serologic risk factors. From the 24 variables, 7 were strongly associated with the primary endpoint (Table 2). These were difficulty to make a fist, patient-reported joint swelling, increased C-reactive protein (CRP), rheumatoid factor (RF), and ACPA, and the cohort variable that was included to adjust for cohort heterogeneity (2-category variable, being at risk based on autoantibody-positive arthralgia/CSA). The AUC was 0.80. When the 5-category cohort variable (5 groups based on geography and autoantibody-positive arthralgia/CSA) was used to adjust for cohort heterogeneity, the results were similar (Supplemental Table S4). The 2-category cohort variable was used as an adjustment factor in further analyses.

US-detected subclinical joint inflammation. A model with only the 10 aggregated US variables (n = 835 with US data) yielded an AUC of 0.63 (Supplemental Table S5). Including these to the panel of fixed clinical and serologic variables identified 5 US variables that were associated with the endpoint (Table 2). This model had an AUC of 0.80.

MRI-detected subclinical joint inflammation. MRI data were analyzed in a similar way as US data. MRI data alone resulted in a model with an AUC of 0.76 (Supplemental

Table 1. Characteristics at baseline and frequency of the primary endpoint*

	Total population (n = 2293)	Population with US (n = 835)	Population with MRI (n = 730)	Population without imaging (n = 728)
Baseline characteristics				
Female, n (%)	1,733 (76)	630 (75)	558 (76)	545 (75)
Age, y, mean (SD)	47 (13)	49 (13)	44 (12)	49 (12)
Symptom duration, wk, median (IQR)	43 (19-104)	51 (26-110)	21 (11-46)	52 (28-156)
Presence of hand symptoms, n (%)	987 (84)	327 (80)	582 (86)	78 (83)
Morning stiffness ≥60 min, n (%)	521 (26)	180 (22)	234 (34)	107 (21)
TJC44, median (IQR)	2 (0-6)	2 (0-6)	4 (2-9)	0 (0-3)
Increased CRP, n (%)	372 (17)	113 (14)	163 (22)	96 (14)
RF positivity, n (%)	949 (41)	357 (44)	150 (20)	431 (61)
Low-positive	455 (20)	155 (19)	59 (8)	241 (34)
High-positive	483 (21)	202 (25)	91 (12)	190 (27)
ACPA positivity, n (%)	1,103 (48)	531 (65)	102 (14)	470 (66)
Low-positive	391 (17)	188 (23)	17 (2)	186 (26)
High-positive	712 (31)	343 (42)	85 (12)	284 (40)
Primary identification method for RA risk, n (%)	1,242 (54)	597 (71)	0	645 (89)
Autoantibody positivity with arthralgia/MSK symptoms				
CSA	1,051 (46)	238 (29)	730 (100)	83 (11)
Primary endpoint				
Clinically apparent inflammatory arthritis within 1 y, n (%)	389 (17)	153 (18)	102 (14)	134 (18)

* ACPA, anti-citrullinated peptide antibody; CRP, C-reactive protein; CSA, clinically suspect arthralgia; IQR, interquartile range; MRI, magnetic resonance imaging; MSK, musculoskeletal; RA, rheumatoid arthritis; RF, rheumatoid factor; TJC, tender joint count; US, ultrasound.

Table S6). Combining the aggregated 12 MRI variables with the fixed clinical and serologic risk factors revealed that 6 MRI variables were incrementally associated with the endpoint (Table 2). The AUC of this model was 0.87.

Development of extended risk stratification. Based on these results, an algorithm was derived using the 6 clinical and serologic variables, the cohort variable, and with the 5 US or 6 MRI variables (Figure 1). The possible scores ranged between 0 and 28 (clinical and serologic data), 0 and 40 (with US data), or 0 to 45 (with MRI data). The predicted risks per score are presented in Figure 1. Additionally, the test characteristics and predictive values for various cutoffs are presented.

Performance. The main statistics for performance were the AUC values, which were 0.80, 0.80, and 0.87 for the clinical and serologic data only, with additional US data, and with additional MRI data, respectively (Table 2). Sensitivity and specificity ~80% was only possible when including MRI data and with a risk score ≥12 (sensitivity 79%, specificity 78%) (Figure 1).

Validation. Internal validation showed little variability in the AUC (Table 3). Analyses in two-thirds and one-third of the population also showed only minor changes in AUC. Analysing the 5 groups of cohorts separately showed AUCs ranging from 0.75 to 0.84 (Supplemental Table S7).

Secondary endpoints. Of the individuals, 12% (209 of 1708) developed RA at 1 year according to the 2010 EULAR/ACR classification criteria. The AUC for this secondary endpoint was 0.85 for the clinical and serologic data, 0.84 when US data were added, and 0.93 when MRI data were added (Table 3). Of the individuals, 21% (441 of 2055) developed clinically apparent IA within 2 years. The AUC values were roughly similar to those with this endpoint at 1 year follow-up (Table 3).

Simplified risk stratification. For ease of use, simplifications were applied for the algorithm with the primary endpoint. The cohort variable was omitted because it is an 'artificial variable' that was used to adjust risk estimates for cohort heterogeneity. US-detected inflammation data were omitted as their contribution was negligible in this analysis. For MRI, only tenosynovitis was included because the contribution of the other combined MRI variables was negligible.

Therefore, the clinical and serologic variables in the final criteria were morning stiffness, presence of patient-reported joint swelling, difficulty making a fist, and increased CRP, RF, and ACPA status (Table 4). For imaging, the presence of tenosynovitis of the flexors of the wrist, the extensors of the wrist, the extensors of the metacarpophalangeal joints and the extensors of the MTP joints were included (Table 4). The clinical and serologic score ranged from 0 to 26; when MRI data was added, the score ranged from 0 to 42 (Table 4). The predicted risks were

Table 2. Results from Lasso regression including clinical and serologic variables plus additional ultrasound- or MRI-detected subclinical joint inflammation*

Variable	Clinical and serologic (n = 2,293)	Clinical and serologic + ultrasound (n = 835)	Clinical and serologic + MRI (n = 730)
Morning stiffness			
30-60 min	1.5	1.5	1.5
≥60 min	2.2	2.2	2.2
Patient-reported joint swelling	2.5	2.5	2.5
Difficulty making a fist	3.7	3.7	3.7
Increased CRP	1.3	1.3	1.3
RF			
Low-positive	1.5	1.5	1.5
High-positive	2.4	2.4	2.4
ACPA			
Low-positive	2.9	2.9	2.9
High-positive	8.2	8.2	8.2
2-category cohort variable[a]	1.7	1.7	1.7
Ultrasound			
PD synovitis PIPs	–	2.8	–
PD tenosynovitis	–	1.7	–
PD synovitis MTPs	–	1.4	–
GS synovitis wrist	–	1.2	–
GS tenosynovitis	–	1.2	–
MRI			
Tenosynovitis flexors wrist	–	–	3.3
Tenosynovitis extensors MCPs	–	–	2.7
Tenosynovitis extensors MTPs	–	–	1.9
Tenosynovitis extensors wrist	–	–	1.3
Osteitis wrist	–	–	1.1
Synovitis MCPs	–	–	1.1
AUC (95% CI)	0.80 (0.77-0.83)	0.80 (0.75-0.84)	0.87 (0.82-0.90)

* Values are the OR derived from the regression coefficient. The regression coefficient was used to derive the weights per variable in Figure 1. As Lasso shrinks the coefficient to zero, measures of variations are not informative. The AUC of the clinical and serologic Lasso model in the n = 835 and n = 730 groups were 0.78 (95% CI 0.72–0.83) and 0.84 (95% CI 0.79–0.88), respectively. ACPA, anti-citrullinated peptide antibody; AUC, area under the curve; CI, confidence interval; CRP, C-reactive protein; CSA, clinically suspect arthralgia; GS, grey scale; Lasso, least absolute shrinkage and selection operator; MCP, metacarpophalangeal joint; MRI, magnetic resonance imaging; MTP, metatarsophalangeal joint; OR, odds ratio; PD; power Doppler; PIP, proximal interphalangeal joint; RA, rheumatoid arthritis; RF, rheumatoid factor.
[a] Identification of RA risk based on CSA or autoantibody-positive arthralgia/musculoskeletal symptoms.

plotted per risk for RA score, both for the criteria without and with imaging (Figure 2). Test characteristics were presented for several cutoffs for the risk for RA score (Figure 2; Supplemental Tables S8 and S9).

The simplified clinical and serologic model had an AUC of 0.80; when adding MRI data, the AUC was 0.86. Calibration graphs are presented in Supplemental Figure S3. Sensitivity and specificity ≥75% was only possible with the criteria with MRI data; a risk score of ≥10 points corresponded to a sensitivity of 75% and specificity of 79%.

The percentages of participants classified as low, intermediate, and high risk (<25%, 25%–75%, and ≥75%, respectively) based on clinical and serologic data are shown for those who did and did not reach the endpoint (Figure 3). Patients who did not develop clinically apparent IA were almost exclusively in the low-risk group. Patients who developed the endpoint were in all 3 groups. Exploration revealed that the convertors classified as low risk were ACPA-positive with few symptoms or ACPA-negative with little subclinical joint inflammation (Supplemental Table S10). Importantly, the intermediate risk group, which should ideally be as small as possible, contained 18% of all studied at-risk individuals. This intermediate risk group was the smallest, 11%, when MRI data were included (Figure 3).

Consensus. During voting in the final meeting, 96% of task force members approved the stratification criteria. It was agreed to present both the extended and simplified versions, and agreement was reached on the target population defined in the eligibility criteria (Table 4).

DISCUSSION

We present EULAR/ACR risk stratification criteria for the development of clinically apparent IA and RA, representing the culmination of an international collaborative effort that

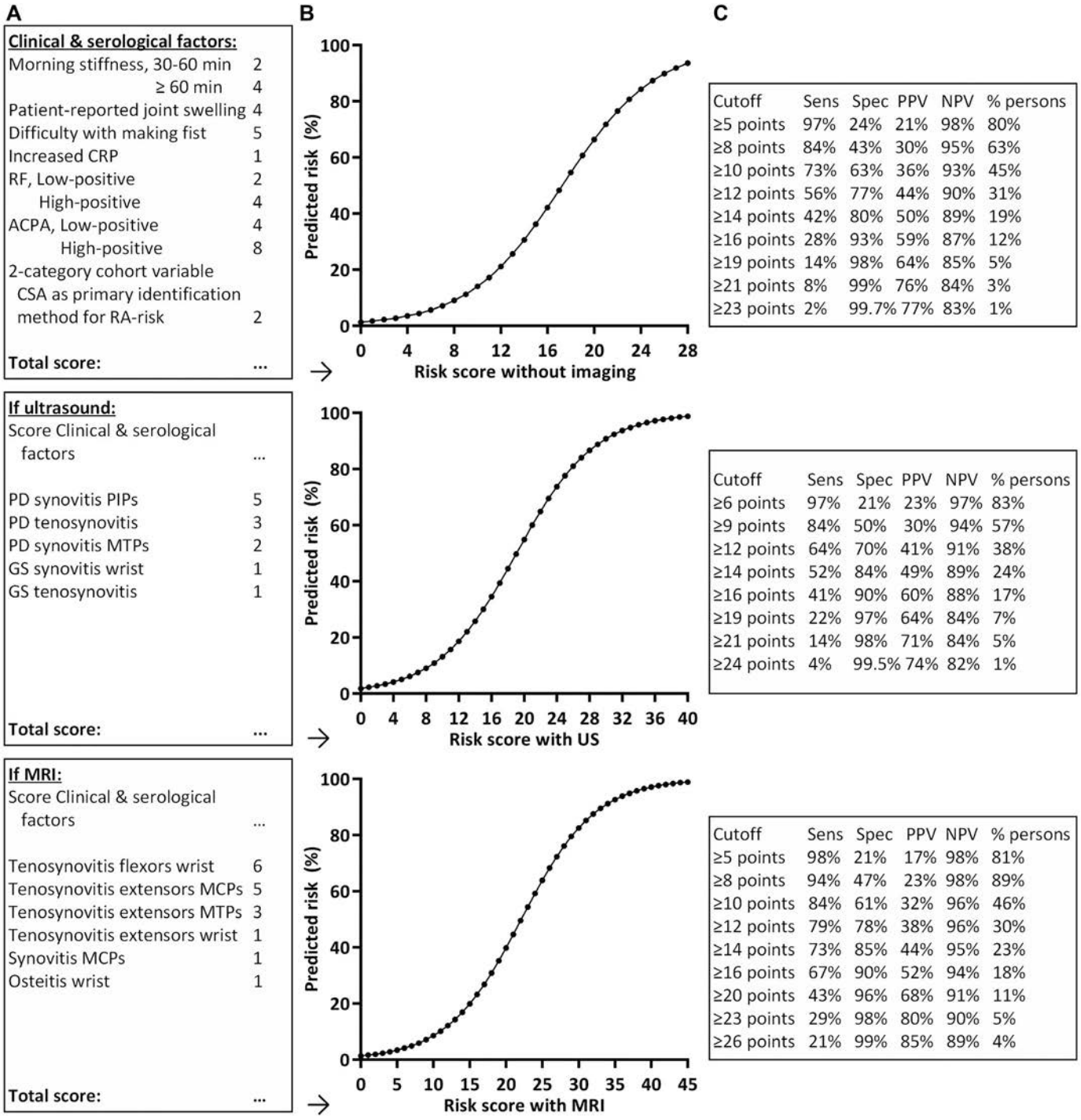


Figure 1. Risk stratification for development of inflammatory arthritis within 1 year using clinical and serologic data and with additional US- or MRI-detected subclinical joint inflammation in arthralgia at risk for RA (A), with predicted risks per risk scores (B), and test characteristics and predictive values for different risk score cutoffs (C). The risk score without imaging ranges from 0 to 28, the risk score with US from 0 to 40, and the risk score with MRI from 0 to 45. The cohort variable was included to adjust for heterogeneity between the cohorts. This was classified in a 2-category variable based on the primary identification method for RA risk, namely identification of being at risk based on autoantibody-positive arthralgia/musculoskeletal symptoms or CSA. ACPA, anti-citrullinated peptide antibody; CRP, C-reactive protein; CSA, clinically suspect arthralgia; GS, grey scale; MCP, metacarpophalangeal joint; MRI, magnetic resonance imaging; MTP, metatarsophalangeal joint; NPV, negative predictive value; PD, power Doppler; PIP, proximal interphalangeal joint; PPV, positive predictive value; RA, rheumatoid arthritis; RF, rheumatoid factor; Sens, sensitivity; Spec, specificity; US, ultrasound.

Table 3. Performance of validation using bootstrapping and data split and for secondary endpoints*

Included variables	Internal validation			Secondary endpoints	
	Bootstrapping AUC (95% CI)	Data $^{2}/_{3}$ AUC (95% CI)	split $^{1}/_{3}$ AUC (95% CI)	Clinical arthritis within 2 y AUC (95% CI)	RA (2010 criteria) within 1 y AUC (95% CI)
Clinical and serologic	0.80 (0.77-0.83)	0.80 (0.77-0.83)	0.80 (0.77-0.83)	0.79 (0.77-0.82)	0.85 (0.83-0.88)
Clinical and serologic + US	0.79 (0.74-0.84)	0.80 (0.76-0.83)	0.79 (0.74-0.82)	0.78 (0.75-0.81)	0.84 (0.78-0.90)
Clinical and serologic + MRI	0.86 (0.81-0.89)	0.84 (0.81-0.86)	0.83 (0.80-0.85)	0.85 (0.81-0.89)	0.93 (0.90-0.97)

* A total of 2,055 persons had 2-year follow-up data; of these, 441 (21.5%) progressed to clinical arthritis within 2 years. A total of 1,708 persons had data on development of clinical arthritis and fulfilled the 2010 EULAR/ACR classification criteria for RA; of these, 209 (12.2%) developed RA within 1 year. Derived models were evaluated after a two-third to one-third split that was performed at the cohort and outcome levels. ACR, American College of Rheumatology; AUC, area under the curve; CI, confidence interval; MRI, magnetic resonance imaging; RA, rheumatoid arthritis; US, ultrasound.

Table 4. EULAR/ACR risk stratification criteria (after simplification) for individuals with arthralgia at risk for RA, to be used in presence or absence of imaging*

Target population (who should be tested?): persons with arthralgia in secondary care without clinical arthritis with the arthralgia not better explained by another disease/condition		
Clinical and serologic characteristics		Score
Morning stiffness[a]	0-30 min	0
	30-60 min	2
	30-60 min	4
Patient-reported joint swelling[b]	No	0
	Yes	4
Difficulty making a fist[c]	No	0
	Yes	5
Increased CRP[d]	No	0
	Yes	1
RF[e]	Negative	0
	Low-positive	2
	High-positive	4
ACPA[e]	Negative	0
	Low-positive	4
	High-positive	8
Clinical and serologic sum score		*See* Figure 2 *for risk*
Imaging characteristics if also MRI		
Tenosynovitis flexors wrist[f]	Absent	0
	Present	6
Tenosynovitis extensors wrist[g]	Absent	0
	Present	2
Tenosynovitis extensors MCPs[h]	Absent	0
	Present	5
Tenosynovitis extensors MTPs[i]	Absent	0
	Present	3
Subscore MRI		
Clinical, serologic, and imaging sum score		*See* Figure 2 *for risk*

* The sum risk score without imaging ranges from 0 to 26 and with imaging from 0 to 42. Predicted risks and test characteristics for several cut-offs of risk scores are presented in Figure 2, with a more extensive list in Supplemental Table S8. The AUC of the criteria without imaging is 0.80 and with MRI data is 0.85. A sensitivity and specificity of ≥75% was only possible with the risk stratification criteria with imaging data and a score of ≥10 points. ACPA, anti-citrullinated peptide antibody; CRP, C-reactive protein; IU, international unit; MCP, metacarpophalangeal joint; MRI, magnetic resonance imaging; MTP, metatarsophalangeal joint; RAMRIS, Rheumatoid Arthritis Magnetic Resonance Imaging Score; RF, rheumatoid factor; ULN, upper limit of normal.

[a] Morning stiffness refers to patient-reported duration of joint stiffness in the morning.

[b] Patient-reported joint swelling refers to presence of a swollen joint as reported by the patient.

[c] Difficulty with making a fist is defined as inability of ≥1 fist with incomplete closure (fingertips do not touch the palm at active closing).

[d] Increased CRP defined as above the local laboratory reference.

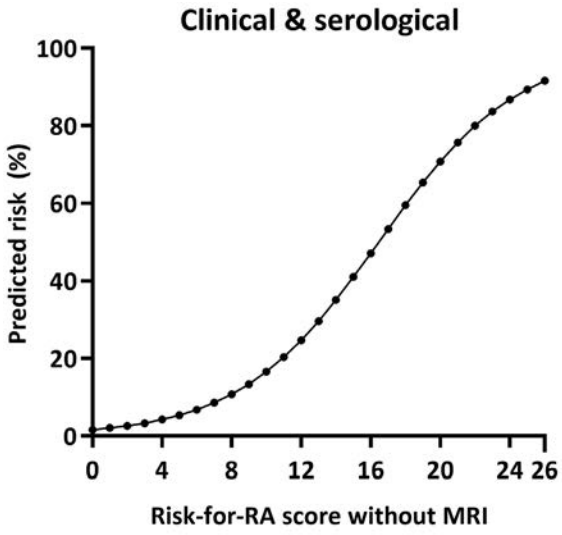
[e] For RF and ACPA, negative refers to IU values less than or equal to the ULN for the laboratory and assay; low-positive refers to IU values that are greater than the ULN but ≤3 times the ULN for the laboratory and assay; high-positive refers to IU values that are >3 times the ULN for the laboratory and assay.

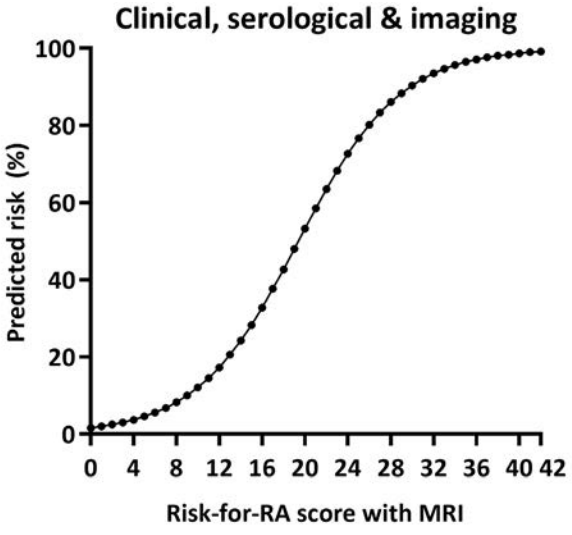
[f] Tenosynovitis of the wrist flexors: presence (RAMRIS ≥1) Of any tenosynovitis of the following: flexor carpi ulnaris, ulnar bursa including flexor digitorum profundus and superficialis tendon quartets, flexor pollicis longus (tendon) in radial bursa, or flexor carpi radialis.

[g] Tenosynovitis of wrist extensors: presence (RAMRIS ≥1) Of any tenosynovitis of the 6 extensor compartments: (I) extensor pollicis brevis, abductor pollicis longus; (II) extensor carpi radialis brevis, extensor carpi radialis longus; (III) extensor pollicis longus; (IV) extensor digitorum communis, extensor indicus proprius; (V) extensor digiti quinti proprius; (VI) extensor carpi ulnaris.

[h] Tenosynovitis of the extensors of the MCPs: presence (RAMRIS ≥1) Of any tenosynovitis of the extensors of MCP2-5.

[i] Tenosynovitis of the extensors of the MTPs: presence (RAMRIS ≥1) Of any tenosynovitis of the extensors of MTP1-5.





Cutoff	Sens	Spec	PPV	NPV
≥4 points	95%	14%	23%	97%
≥6 points	90%	38%	26%	96%
≥8 points	78%	50%	33%	94%
≥10 points	66%	71%	38%	92%
≥12 points	51%	80%	46%	90%
≥14 points	37%	91%	53%	88%
≥16 points	24%	94%	60%	86%
≥18 points	13%	98%	64%	85%
≥20 points	7%	99%	71%	84%

Cutoff	Sens	Spec	PPV	NPV
≥5 points	94%	41%	22%	98%
≥8 points	83%	61%	34%	95%
≥10 points	75%	79%	41%	95%
≥12 points	70%	85%	47%	95%
≥15 points	55%	93%	61%	93%
≥18 points	47%	96%	69%	92%
≥20 points	34%	98%	74%	90%
≥22 points	28%	99%	82%	89%
≥24 points	20%	99%	84%	88%

Figure 2. Predicted risks (graph) and test characteristics and predictive values for several score cutoffs (table) for the risk stratification criteria with- out or with imaging. Predicted risks for the sum scores obtained in Table 4. Test characteristics and predictive values for several cutoffs of risk scores are shown; an extensive list of cutoffs with test characteristics and data on the percentage of persons with these cutoffs is presented in Supplemental Table S8. The area under the curve of the criteria without imaging is 0.80 and with MRI data is 0.85. MRI, magnetic resonance imaging; NPV, negative predictive value; PPV, positive predictive value; RA, rheumatoid arthritis; Sens, sensitivity; Spec, specificity.

included researchers curating arthralgia cohorts, expert rheumatologists, health care professionals, and patient research partners across Europe and North America. The risk stratification criteria consist of symptoms (morning stiffness, patient-reported joint swelling), a sign (difficulty making a fist), serologic markers (CRP, RF, and ACPA), and MRI-detected tenosynovitis. This product has been developed with the key aim of supporting the inclusion of homogeneous risk groups in future prevention trials with individuals with arthralgia in secondary care in whom imminent RA is considered more likely than other causes of arthralgia. Importantly, risk of a disease is not the same as having a disease and therefore, the product is not described as classification criteria.

We aimed to derive an algorithm that is broadly applicable both in settings in which imaging modalities are available to detect subclinical joint inflammation and where such modalities are lacking. The statistical methodology was adapted accordingly, aiming to take maximum advantage of the information from clinical and serologic variables. Most importantly, the information from clinical and serologic data was fixed and prioritized. In the presence of both imaging-detected subclinical inflammation and clinical data, whereby the latter was related to inflammation (eg, morning stiffness is known to be related to subclinical synovitis[21] and difficulty making a fist to subclinical tenosynovitis[22]), multivariable analyses will generally select the imaging variable as the best predictor for RA development because of the most direct relationship. However, for practical application, clinical variables were given priority (set as fixed in the analyses). Consequently, the effect sizes of the clinical and serologic variables were larger than if such an approach had not been applied. In contrast, the effect sizes of imaging variables were lower than would have been the case without this prioritization, and with this approach only revealed the incremental value of adding imaging to what was already known through the clinical and serologic variables. The value of imaging is therefore also lower than previously reported in studies in symptomatic at-risk individuals. This method led to an algorithm that is applicable also in the absence of imaging modalities. The criteria without imaging perform well because they use the underlying correlation of clinical with imaging variables in this

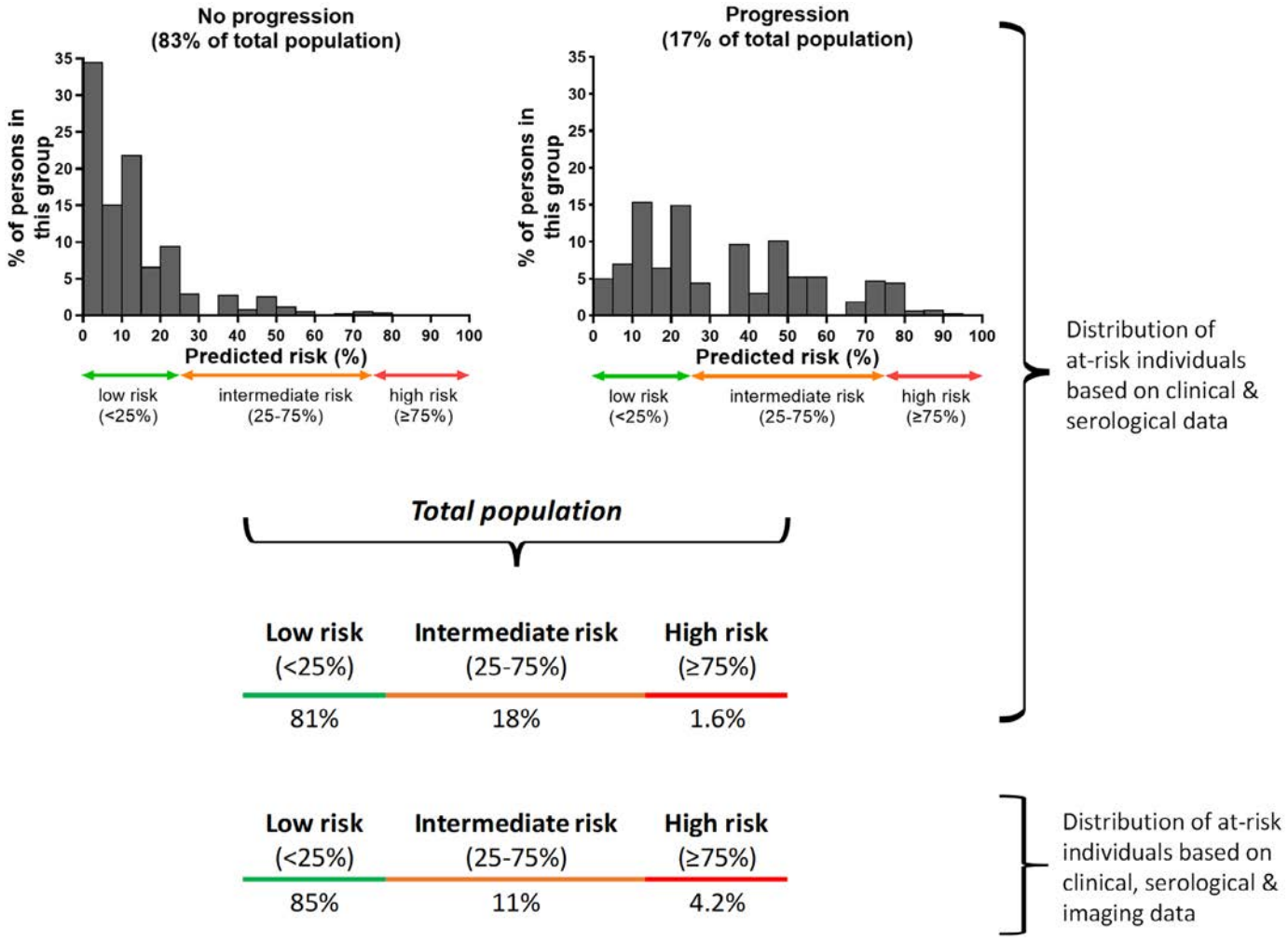


Figure 3. Distribution of subjects who did not and did reach the endpoint based on the predicted risks using the risk stratification criteria with clinical and serologic data in Table 4. Subsequently, risk categories were defined as low (<25% risk), intermediate (25%-75%) and high (≥75%) in the total population based on clinical and serologic data (n = 2,293). The same was done when additionally using MRI data (n = 703). Characteristics of per- sons who developed the endpoint and were categorized as low risk (8% of the total population) are presented in Supplemental Table S10, both for the setting without and with imaging data. MRI, magnetic resonance imaging.

population. Because this is a derivative, the criteria with imaging may be preferred for optimal trial design.

Importantly, the risk stratification criteria are derived from data collected in secondary care and as such are based on the correlation structure between the variables that exist in arthralgia patients in secondary care. For example, patient-reported swelling is related to subclinical joint inflammation in arthralgia at risk for RA.[23] In other situations, eg, the general population in which hand osteoarthritis is more prevalent than imminent RA, this relationship may be different. Consequently, this method cannot be extrapolated for trials in other populations (eg, primary care-based populations) or for triaging referrals from primary to secondary care. The developed criteria are expected to be widely applicable, but only in persons seen with arthralgia in secondary care, where impending RA is more likely than other causes of arthralgia.

The criteria were derived with the idea that the accuracy of data-driven risk stratification performs better than risk stratification by expert opinion of individual rheumatologists. Whether this assumption is true needs to be verified in future studies.

Performance was evaluated using the AUC (combining sensitivity and specificity). In line with the aim of promoting the inclusion of homogeneous risk groups in future studies, this metric evaluates performance for groups of persons. At the design of the project, the intention was to achieve both specificity and sensitivity of ~80%. The algorithm using clinical and serologic data achieved a 'best combination' of sensitivity and specificity of approximately 65% to 70% (Figures 1 and 2). When MRI data were available, a specificity and sensitivity of almost 80% was achieved by a score ≥12 in the expanded version or ≥10 in the simplified version. Therefore, if a sensitivity and specificity of both >75% are desired, the imaging criteria and a cutoff of ≥10 points

are appropriate. However, the expert committee did not choose one cutoff point because, depending on the clinical situation or intervention (with different risks of side effects), either a (very) high sensitivity or specificity may be preferred. With the current presentation, investigators can choose the test characteristics and corresponding cutoff that is felt acceptable by their team of rheumatologic and patient experts.

Importantly, it was not the intention to develop an instrument for individual patients in clinical practice. Absolute risk estimates are more difficult to generalize (than test characteristics) because they depend on prior risks of a disease (context-dependent prevalence). Interestingly, the prior risk of developing clinical arthritis within 1 year was quite similar in all at risk populations studied here (ranging from 14%–18%). Nonetheless, this may be different in other places, and the developed criteria are therefore not suitable for decision-making in individual patients.

Subclinical joint inflammation (particularly tenosynovitis) is a known predictor of clinical arthritis and RA development,[24] and data collected via US and MRI were assessed. MRI-detected subclinical joint inflammation showed some incremental value compared to clinical and serologic data alone. This was demonstrated by an increase in AUC and a higher combination of sensitivity and specificity. This did not occur for subclinical inflammation detected with US. This difference may be explained by previous findings that tenosynovitis is a powerful imaging predictor missed by US in up to 80% of tenosynovitis lesions compared to MRI.[25,26] Additionally, US scanning protocol and gradings that were used in the cohorts were different and could not be standardized in retrospect. The available US data also included less tenosynovial sites than MRI (Supplemental Table S3). US-detected erosions were shared from 1 dataset (Rotterdam SONAR cohort) and were found to be nonpredictive. The taskforce had extensive discussions about the US results, and it was agreed that the data would be presented as they are, with the US in the extended algorithm. Furthermore, the US results presented here apply to the setting of individuals with arthralgia at risk for RA and are not relevant to the value of US in a broader context in the field of RA.

MRI also has disadvantages. Some subtle synovitis and osteitis is present when using contrast-enhanced MRI in the general symptom-free population, especially in older age and in specific joints and bones.[18] These variations need to be considered when defining an abnormal MRI result to prevent false-positive tests.[19] In contrast, tenosynovitis has been found to be rare in the general population[18] and was the only variable that remained in the simplified risk stratification criteria. Tenosynovitis is also relatively easy to detect reliably on MRI. Although MRI is already used for the classification of other rheumatic diseases,[27] MRI may be considered impractical due to the need for intravenous contrast agents, long scan times, and resulting high costs. Recent data have emerged suggesting that a short modified Dixon MRI sequence correlates well with conventional contrast-enhanced MRI for scanning of hands/forefeet.[28] With a scan time of 5 minutes without the need for intravenous contrast, this could make MRI more feasible. The value of this short MRI sequence and its cost-effectiveness are subjects for further research. If positive, this could have a significant impact on the assessment of the burden and costs of MRI compared to the incremental value in risk stratification.

The method for identifying persons at risk of RA varies between centers and countries. In some, the presence of a positive autoantibody test is the starting point, and the presence of any MSK symptom increases the risk. Alternatively, individuals can be identified by a combination of symptoms and signs that resemble RA but where clinical IA is absent, ie, by having CSA. Both identification methods alone have proven to be insufficient. Risk stratification, as shown previously[29–31] and also here, requires a combination of clinical and serologic features combined with imaging of joints. Although all participants in the CSA cohorts were reported to have CSA by their rheumatologists, for participants in the autoantibody-positive cohorts, it was unknown whether they had CSA (according to their rheumatologist), and data from some items of the EULAR definition of CSA[7] were absent/incomplete. Therefore, having CSA at an individual level could not be included in the analyses, but several clinical items that are part of the EULAR definition of CSA have entered the risk stratification criteria (morning stiffness and difficulty making a fist).

The primary endpoint, development of clinically apparent IA, was ascertained by consensus and was a pragmatic choice because data on the 2010 classification criteria for RA were missing in 25%. RA was most likely the final diagnosis in the event of the development of clinical arthritis because at-risk individuals were selected based on RA-related autoantibodies or CSA. Previous studies of cohorts that are included here have demonstrated that diagnoses other than RA or undifferentiated arthritis are indeed rare.[11,29,30] Moreover, the performance for the development of RA was comparable to that for the primary endpoint. Thus, risk stratifications were made for developing IA, an acceptable proxy for RA in this setting.

Heterogeneity between cohorts (eg, individual characteristics, outcomes) poses a challenge when combining datasets.[32,33] Sex, age, and endpoint frequency were similar between the datasets, but there were also differences in clinical and serologic characteristics arising from the method of identifying individuals at risk (eg, the cohorts that included individuals with CSA had higher frequencies of tender joints and the cohorts that selected for autoantibodies showed a higher frequency of ACPA positivity). Unmeasured factors related to geography or other factors may also differ between the cohorts. Cohort heterogeneity was taken into account by including a cohort variable. This had an independent association with the endpoint, suggesting an influence. Although this may suggest some impact on generalizability, exclusion of the cohort variable in the simplification step did not affect performance.

Validation was performed with bootstrapping and a data split. The similarity in AUCs demonstrated robustness. Data from European populations were shared in the taskforce. We did not find cohorts of arthralgia patients from the US, Canada, or Asia who were willing to share data. External validation in non-European populations should be conducted.

A limitation is that the selection of risk variables was based on consensus and availability. Some risk factors previously shown to be predictive in some cohorts (eg, human leukocyte antigen-shared epitope alleles,[29,30,34] functional disability,[35] and T cell subsets[30]), were not assessed in many cohorts and thus not included here. The additional value of these variables can be evaluated in future studies. Including additional variables would be valuable if, for example, the ACPA-negative progressing individuals that are currently classified as low risk would be recognized as high risk. This will not be easy because included ACPA-negative individuals with arthralgia had a prior risk of 9%, and posttest risks depend on prior risks. The high AUC also suggests that there is little room for improvement.

Patient research partners were involved during all phases of the project. Based on their advice, further research into the patients' perspective is needed. Understanding how risk information is perceived and relates to willingness to participate in trials is essential to inform future trials. Perceived risk status has been shown to affect tolerance for side effects of preventive treatments.[16,36] Furthermore, the proportion of at-risk individuals with 'intermediate' risk of RA was relatively small (11%), but their views of risk classification would need further investigation.

In conclusion, risk stratification criteria for arthralgia at risk for RA have been developed based on international cohort data and expert consensus. The developed risk stratification criteria are validated and intended to support the inclusion of homogeneous risk groups in future prevention trials.

ACKNOWLEDGMENTS

The authors are grateful to all the cohorts that shared their data. Daniel Aletaha, Andrew Filer, Georg Schett, and Paul Emery are acknowledged for allowing access to at-risk cohorts. Daniel Aletaha, Duncan Porter, and Iain McInnes are acknowledged for advice. Professor Saskia le Cessie is acknowledged for statistical advice. The authors also wish to thank EULAR and ACR for their financial support, Simona Lupatin from EULAR, and Amy Turner from ACR for administrative support, and EMEUNET for providing a representative (Ilfita Sahbudin).

AUTHOR CONTRIBUTIONS

All authors have contributed to this work, drafted the work or revised it critically and approved the final version. Drs van Steenbergen and van der Helm-van Mil confirm that all authors have provided the final approval of the version to be published and take responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

DISCLOSURES

The authors reported the following competing interests, outside the submitted work. SA–grants or contracts: Novartis, AbbVie; payment for lectures: AbbVie, Eli Lilly, Sandoz. CC –payment or honoraria for lectures: AbbVie, Amgen, Boehringer Ingelheim, Ewopharma, Lilly, Novartis, Pfizer, Sobi; leadership or fiduciary role in other board: President-elect of Romanian Society of Rheumatology, President of Romanian Registry of Rheumatic Diseases. SBC–grants or contracts: investigator and consultant Amgen, AbbVie, BMS, Pfizer; consulting fees: Amgen, AbbVie, BMS, Pfizer; payment or honoraria for lectures: Pfizer; participation on a data safety monitoring board (DSMB): Gilead. APC–grants or contracts: BMS, UCB; consulting fees: BMS, Galvani, Janssen, UCB, Lilly; payment or honoraria for lectures: BMS, Galapagos; leadership or fiduciary role in other board: chair EULAR Research Centre Clinical Research subcommittee. KD–payment or honoraria for lectures: Werfen, ThermoFisher; receipt of equipment, materials: Werfen. BE–grants or contracts: Boehringer Ingelheim; consulting fees: Boehringer Ingelheim. AK–consulting fees: Novartis Deutschland GmbH, Eli Lilly Deutschland GmbH, Gilead Sciences, Medac GmbH. KM–grants or contracts: Gilead, Lilly, Serac Healthcare, Astra Zeneca; consulting fees: Galapagos, Lilly, Serac Healthcare, Deepcure, Zura Bio; payment or honoraria for lectures: AbbVie, Galapagos; sup- port for attending meetings and/or travel: Lilly; participation on a DSMB: Galapagos. LJON–consulting fees: Fresenius Kabi, UCB; payment or honoraria for lectures: AbbVie. AS–grants or contracts: FOREUM; consulting fees: AbbVie, UCB, Novartis; payment of honoraria for lectures: Lilly, AbbVie; support for attending meetings and/or travel: Janssen, Lilly; leadership of fiduciary role in other board: editor in chief ARP Rheumatology, associate editor RMD Open. JAPdS–consulting fees: Amgen, Fresenius Kabi; payment or honoraria for lectures: UCB, Amgen; participation on a DSMB: Fresenius Kabi. JAS–grants or con- tracts: National Institute of Arthritis and Musculoskeletal and Skin Diseases, Rheumatology Research Foundation, R. Bruce and Joan M. Mickey Research Scholar Fund, Gordon and Llura Gund Foundation, BMS, Boehringer Ingelheim, Sonoma Biotherapeutics; consulting fees: AbbVie, Amgen, Boehringer Ingelheim, BMS, Gilead, Inova Diagnostics, Johnson & Johnson, MustangBio, Optum, Pfizer, ReCor, Sana, Sobi, UCB. PS–grants or contracts: FOREUM, Horizon/HADEA, AAL, EUREKA, Jubil€aumsfond der Stadt Wien; consulting fees: AbbVie; payment or honoraria for lectures: AstraZeneca; support for attending meetings and/or travel: Janssen, Galapagos. RL–payment or honoraria for lectures: AbbVie, UCB, Pfizer, Galapagos, Janssen Pharma, Novartis, Lilly; participation on a DSMB: UCB SLE trial; leadership or fiduciary role in other board: ASA board; other financial or nonfinancial interests: director of Rheumatology Consultancy BV, director of Joint Imaging BV. KR –grants or contracts: BMS.

REFERENCES

1. Aletaha D, Neogi T, Silman AJ, et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Ann Rheum Dis 2010;69(9):1580–8.
2. van der Linden MPM, le Cessie S, Raza K, et al. Long-term impact of delay in assessment of patients with early arthritis. Arthritis Rheum 2010;62(12):3537–46.
3. Nielen MMJ, van Schaardenburg D, Reesink HW, et al. Specific autoantibodies precede the symptoms of rheumatoid arthritis: a study of serial measurements in blood donors. Arthritis Rheum 2004;50(2): 380–6.
4. Rantapää-Dahlqvist S, de Jong BAW, Berglin E, et al. Antibodies against cyclic citrullinated peptide and IgA rheumatoid factor predict the development of rheumatoid arthritis. Arthritis Rheum 2003; 48(10):2741–9.
5. van Steenbergen HW, Huizinga TW, van der Helm-van Mil AH. The preclinical phase of rheumatoid arthritis: what is acknowledged and what needs to be assessed? Arthritis Rheum 2013;65(9):2219–32.
6. Gerlag DM, Raza K, van Baarsen LGM, et al. EULAR recommendations for terminology and research in individuals at risk of rheumatoid arthritis: Report from the Study Group for Risk Factors for Rheumatoid Arthritis. Ann Rheum Dis 2012;71(5):638–41.
7. van Steenbergen HW, Aletaha D, Beaart-van de Voorde LJJ, et al. EULAR definition of arthralgia suspicious for progression to rheumatoid arthritis. Ann Rheum Dis 2017;76(3):491–6.
8. Burgers LE, Siljehult F, Ten Brinck RM, et al. Validation of the EULAR definition of arthralgia suspicious for progression to rheumatoid arthritis. Rheumatology (Oxford) 2017;56(12):2123–8.
9. Mankia K, Siddle H, Di Matteo A, et al. A core set of risk factors in individuals at risk of rheumatoid arthritis: a systematic literature review informing the EULAR points to consider for conducting clinical trials and observational studies in individuals at risk of rheumatoid arthritis. RMD Open 2021;7(3):e001768.
10. Mankia K, Siddle HJ, Kerschbaumer A, et al. EULAR points to consider for conducting clinical trials and observational studies in individuals at risk of rheumatoid arthritis. Ann Rheum Dis 2021;80(10): 1286–98.
11. Krijbolder DI, Verstappen M, van Dijk BT, et al. Intervention with methotrexate in patients with arthralgia at risk of rheumatoid arthritis to reduce the development of persistent arthritis and its disease burden (TREAT EARLIER): a randomised, double-blind, placebo-con- trolled, proof-of-concept trial. Lancet 2022;400(10348):283–94.
12. Cope AP, Jasenecova M, Vasconcelos JC, et al. Abatacept in individuals at high risk of rheumatoid arthritis (APIPPRA): a randomised, double-blind, multicentre, parallel, placebo-controlled, phase 2b clinical trial. Lancet 2024;403(10429):838–49.
13. Rech J, Tascilar K, Hagen M, et al. Abatacept inhibits inflammation and onset of rheumatoid arthritis in individuals at high risk (ARIAA): a randomised, international, multicentre, double-blind, placebo-controlled trial. Lancet 2024;403(10429):850–9.
14. Mosor E, Stoffer-Marx M, Steiner G, et al. I would never take preventive medication! Perspectives and information needs of people who underwent predictive tests for rheumatoid arthritis. Arthritis Care Res (Hoboken) 2020;72(3):360–8.
15. Newsum EC, van der Helm-van Mil AHM, Kaptein AA. Views on clinically suspect arthralgia: a focus group study. Clin Rheumatol 2016; 35(5):1347–52.
16. Simons G, Veldwijk J, DiSantostefano RL, et al. Preferences for preventive treatments for rheumatoid arthritis: discrete choice survey in the UK, Germany and Romania. Rheumatology (Oxford) 2023;62(2): 596–605.
17. Padovano I, Costantino F, Breban M, et al. Prevalence of ultrasound synovial inflammatory findings in healthy subjects. Ann Rheum Dis 2016;75(10):1819–23.
18. Mangnus L, van Steenbergen HW, Reijnierse M, et al. Magnetic resonance imaging-detected features of inflammation and erosions in symptom-free persons from the general population. Arthritis Rheumatol 2016;68(11):2593–602.
19. Boer AC, Burgers LE, Mangnus L, et al. Using a reference when defining an abnormal MRI reduces false-positive MRI results-a longitudinal study in two cohorts at risk for rheumatoid arthritis. Rheumatology (Oxford) 2017;56(10):1700–6.
20. van der Heijde D, Aletaha D, Carmona L, et al. 2014 Update of the EULAR standardise operating procedures for EULAR-endorsed recommendations. Ann Rheum Dis 2015;74(1):8–13.
21. Krijbolder DI, Wouters F, van Mulligen E, et al. Morning stiffness precedes the development of rheumatoid arthritis and associates with systemic and subclinical joint inflammation in arthralgia patients. Rheumatology (Oxford) 2022;61(5):2113–8.
22. Wouters F, van der Giesen FJ, Matthijssen XME, et al. Difficulties making a fist in clinically suspect arthralgia: an easy applicable phenomenon predictive for RA that is related to flexor tenosynovitis. Ann Rheum Dis 2019;78(10):1438–9.
23. Boeren AMP, Khidir SJH, de Jong PHP, et al. Patient-reported swelling in arthralgia patients at risk for rheumatoid arthritis: is it of value? Rheumatology (Oxford) 2023;63(12):3415–21.
24. Rogier C, Hayer S, van der Helm-van Mil A. Not only synovitis but also tenosynovitis needs to be considered: why it is time to update textbook images of rheumatoid arthritis. Ann Rheum Dis 2020;79(4):546–7.
25. Ohrndorf S, Boer AC, Boeters DM, et al. Do musculoskeletal ultrasound and magnetic resonance imaging identify synovitis and tenosynovitis at the same joints and tendons? A comparative study in early inflammatory arthritis and clinically suspect arthralgia. Arthritis Res Ther 2019;21(1):59.
26. Wakefield RJ, O'Connor PJ, Conaghan PG, et al. Finger tendon disease in untreated early rheumatoid arthritis: a comparison of ultrasound and magnetic resonance imaging. Arthritis Rheum 2007; 57(7):1158–64.
27. Rudwaleit M, van der Heijde D, Landewé R, al. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. Ann Rheum Dis 2009;68(6):777–83.
28. Boeren AMP, Niemantsverdriet E, Verstappen M, et al. Towards a simplified fluid-sensitive MRI protocol in small joints of the hand in early arthritis patients: reliability between modified Dixon and regular gadolinium enhanced TSE fat saturated MRI-sequences. Skeletal Radiol 2023;52(6):1193–202.
29. van de Stadt LA, Witte BI, Bos WH, van Schaardenburg D. A prediction rule for the development of arthritis in seropositive arthralgia patients. Ann Rheum Dis 2013;72(12):1920–6.
30. Duquenne L, Hensor EM, Wilson M, et al. Predicting inflammatory arthritis in at-risk persons: development of scores for risk stratification. Ann Intern Med 2023;176(8):1027–36.
31. Matthijssen X, Van Dijk B, Wouters F, et al. FRI0542 Obtaining high positive predictive values for the development of clinically apparent arthritis in patients presenting with clinically suspect arthralgia; is it feasible? [Abstract]. Ann Rheum Dis 2020;79:872. https://doi.org/10.1136/annrheumdis-2020-eular.1697.
32. Steyerberg EW, Nieboer D, Debray TPA, et al. Assessment of heterogeneity in an individual participant data meta-analysis of prediction models: an overview and illustration. Stat Med 2019;38(22):4290–309.
33. Studenic P, Hensvold A, Kleyer A, et al. Prospective studies on the risk of rheumatoid arthritis: the European Risk RA Registry. Front Med (Lausanne) 2022;9:824501.

34. Wouters F, Maurits MP, van Boheemen L, et al. Determining in which pre-arthritis stage HLA-shared epitope alleles and smoking exert their effect on the development of rheumatoid arthritis. Ann Rheum Dis 2022;81(1):48–55.
35. Khidir SJH, van Dijk BT, Krijbolder DI, et al. Joint involvement in RA starts predominantly in the hands: functional, clinical and imaging studies in clinically suspect arthralgia and during progression to RA. RMD Open 2023;9(2):e003107.
36. Simons G, Janssen EM, Veldwijk J, et al. Acceptable risks of treatments to prevent rheumatoid arthritis among first-degree relatives: demographic and psychological predictors of risk tolerance. RMD Open 2022;8(2):e002593.

Arthritis & Rheumatology
Vol. 78, No. 3, March 2026, pp 537–547
DOI 10.1002/art.43137

American College of Rheumatology
Empowering Rheumatology Professionals

EULAR/American College of Rheumatology Classification Criteria for Pediatric Chronic Nonbacterial Osteomyelitis

Yongdong Zhao,[1] Melissa S. Oliver,[2] Anja Schnabel,[3] Eveline Y. Wu,[4] Zhaoyi Wang,[1] Achille Marino,[5] Cassyanne L. Aguiar,[6] Jonathan D. Akikusa,[7] Ummusen Kaya Akca,[8] Beverley Almeida,[9] Simone Appenzeller,[10] Erin Balay-Dustrude,[1] Ozge Basaran,[11] Matthew L. Basiaga,[12] Yelda Bilginer,[8] David A. Cabral,[13] Martina Capponi,[14] Nathan Donaldson,[15] Bugra Han Egeli,[16] Emily J. Fox,[17] Antonella Insalaco,[18] Ramesh S. Iyer,[19] Annette F. Jansson,[20] Inna Kostik,[21] Mikhail Kostik,[22] Leonard K. Kovalick,[4] Katia Tomie Kozu,[23] Sivia K. Lapidus,[24] Tzielan C. Lee,[25] Aleksander Lenert,[26] Kamran Mahmood,[9] Edoardo Marrani,[27] Doaa Mosad Mosa,[28] Ian Muse,[1] Alexander Mushkin,[29] Katherine D. Nowicki,[15] Farzana Nuruzzaman,[30] Karen Onel,[31] Manuela Pardeo,[18] Trang Sophia Pham,[1] Lauren Potts,[32] Athimalaipet V. Ramanan,[33] Angelo Ravelli,[34] Nathan D. Rogers,[15] Andrew W. Grim,[35] Micol Romano,[36] Natalie Rosenwasser,[1] Takashi Shawn Sato,[37] Gabriele Simonini,[27] Jennifer B. Soep,[15] Sara M. Stern,[38] Timmy Strauss,[3] Angela Taneja Kohli,[39] Alexander C. Theos,[40] Lori B. Tucker,[13] Leslie F. Vogel,[41] Shima Yasin,[42] Stephen C. Wong,[1] Katerina Bouchalova,[43] Alison M. Hendry,[44] Kevin C. Cain,[45] Hermann J. Girschick,[46] Fatma Dedeoglu,[16] Christian M. Hedrich,[47] Ronald M. Laxer,[48] Polly J. Ferguson,[42] Raymond Naden,[†] and Seza Ozen[8]

This criteria set has been approved by the American College of Rheumatology (ACR) Board of Directors and the EULAR Executive Committee. This signifies that the criteria set has been quantitatively validated using patient data, and it has undergone validation based on an independent data set. All ACR/EULAR-approved criteria sets are expected to undergo intermittent updates.

Classification criteria are essential in clinical and basic science research because they allow investigators to study relatively homogeneous populations of patients recruited from a single or multiple research sites. In clinical settings, diagnoses are made by health care professionals evaluating an individual patient's symptoms, signs, and results of laboratory and imaging studies in order to guide therapeutic recommendations. Patients diagnosed with a particular disease may or may not fulfill classification criteria for that disease. Classification criteria, in the hands of an experienced clinician with expertise in rheumatology, may inform a diagnostic evaluation, but improperly applied classification criteria may lead to misdiagnosis.

The ACR is an independent, professional, medical, and scientific society that does not guarantee, warrant, or endorse any commercial product or service.

Objective. To develop and validate classification criteria for pediatric chronic nonbacterial osteomyelitis (CNO) jointly supported by EULAR and the American College of Rheumatology (ACR).

Methods. This international initiative had 4 phases: (1) candidate items were proposed in a survey of pediatric rheumatologists, (2) criteria definition and reduction by Delphi and nominal group technique exercises, (3) criteria weighting using multicriteria decision analysis, and (4) refinement of weights and threshold score in a development cohort of 441 patients and validation in another cohort of 514 patients.

Results. The new EULAR/ACR classification criteria for CNO require typical radiographic or magnetic resonance imaging findings and bone pain as an obligatory entry criterion and exclusion criteria of malignancy, infection, vitamin C deficiency, and hypophosphatasia, followed by additive weighted criteria in 5 clinical (site of bone lesions, pattern of bone lesions, age at onset, coexisting conditions, fever) and 4 pathology/laboratory domains (bone biopsy findings if done, anemia, C-reactive protein level, and erythrocyte sedimentation rate). A total score ≥55 is required for classification as CNO. The new criteria had a sensitivity of 82% and specificity of 98% in the validation cohort.

Conclusion. These new classification criteria for pediatric CNO developed with international input reflect current views about CNO, have high specificity and good sensitivity, and provide a key foundation for future CNO research.

INTRODUCTION

Chronic nonbacterial osteomyelitis (CNO) is a rare autoinflammatory bone disease of unknown cause with an annual incidence of 0.4 to 2.3 per 100,000 children.[1–3] CNO may significantly impact patients' quality of life and result in permanent bone damage, long-term disability, and deformity if left untreated.[4–7] The term chronic recurrent multifocal osteomyelitis is used when multiple bones are affected over time. We use

This article is also published in *Annals of Rheumatic Diseases*.

Supported by a EULAR/American College of Rheumatology classification criteria award to Yongdong Zhao and Seza Ozen.

[1]Yongdong Zhao, Zhaoyi Wang, Erin Balay-Dustrude, Ian Muse, Trang Sophia Pham, Natalie Rosenwasser, Stephen C. Wong: Pediatric Rheumatology, Department of Pediatrics, University of Washington School of Medicine, Seattle; [2]Melissa S. Oliver: Division of Pediatric Rheumatology, Riley Hospital for Children, Indiana University School of Medicine, Indianapolis; [3]Anja Schnabel, Timmy Strauss: Department of Pediatrics, Faculty of Medicine and University Hospital Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany; [4]Eveline Y. Wu, Leonard K. Kovalick: Division of Pediatric Rheumatology, Department of Pediatrics, The University of North Carolina at Chapel Hill, Chapel Hill; [5]Achille Marino: Pediatric Rheumatology, ASST Gaetano Pini-CTO, Milan, Italy; [6]Cassyanne L. Aguiar; Division of Pediatric Rheumatology, Department of Pediatrics, Children's Hospital of The King's Daughters, Eastern Virginia Medical School, Norfolk; [7]Jonathan D. Akikusa: Rheumatology service, Department of General Medicine, Royal Children's Hospital, Melbourne, Victoria, Australia; [8]Ummusen Kaya Akca, Yelda Bilginer, Seza Ozen: Hacettepe University Faculty of Medicine, Department of Pediatric Rheumatology, Ankara, Turkey; [9]Beverley Almeida, Kamran Mahmood: Department of Paediatric Rheumatology, Alder Hey Children's NHS Foundation Trust, Eaton Road, Liverpool, UK; [10]Simone Appenzeller: Department of Orthopedics, Rheumatology and Traumatology, School of Medical Science, University of Campinas, Brazil; [11]Ozge Basaran: Hacettepe University Faculty of Medicine, Department of Pediatric Rheumatology, Ankara, Turkey, and Division of Pediatric Rheumatology, Hospital for Special Surgery, New York, New York; [12]Matthew L. Basiaga: Division of Pediatric Rheumatology, Department of Pediatric and Adolescent Medicine, Mayo Clinic, Rochester, Minnesota; [13]David A. Cabral, Lori B. Tucker: BC Children's Hospital and University of British Columbia, Vancouver, British Columbia, Canada; [14]Martina Capponi: Department of Maternal Sciences and Urological Sciences, Sapienza University of Rome, Rome, Italy; [15]Nathan Donaldson, Katherine D. Nowicki, Nathan D. Rogers, Jennifer B. Soep: University of Colorado School of Medicine, Aurora; [16]Bugra Han Egeli, Fatma Dedeoglu: Division of Immunology, Rheumatology Program, Boston Children's Hospital, Department of Pediatrics, Harvard Medical School, Boston, Massachusetts; [17]Emily J. Fox: Division of Pediatric Rheumatology, Department of Pediatrics, Children's Mercy Hospital, Kansas City, Missouri; [18]Antonella Insalaco, Manuela Pardeo: Division of Rheumatology, ERN RITA Center, IRCCS Ospedale Pediatrico Bambino Gesù, Rome, Italy; [19]Ramesh S. Iyer: Pediatric Radiology, Department of Radiology, University of Washington School of Medicine, Seattle; [20]Annette F. Jansson: Division of Pediatric Rheumatology and Immunology, Department of Pediatrics, Dr. von Hauner Children's Hospital, University Hospital, Ludwig-Maximilians-Universität München, Munich, Germany; [21]Inna Kostik: Sanatorium for Children 'Detskie Duny', Saint-Petersburg, Russia; [22]Mikhail Kostik: Saint-Petersburg State Pediatric Medical University, Saint-Petersburg, Russia; [23]Katia Tomie Kozu: Pediatric Rheumatology Unit, Instituto da Criança e do Adolescente, Hospital das Clinicas HCFMUSP, Faculdade de Medicina, Universidade de Sao Paulo, Sao Paulo, Brazil; [24]Sivia K. Lapidus: Pediatric Rheumatology Division, Joseph M. Sanzari Children's Hospital, Hackensack University Medical Center and Hackensack Meridian School of Medicine, Hackensack, New Jersey; [25]Tzielan C. Lee: Division of Pediatric Rheumatology, Stanford University School of Medicine, Stanford, California; [26]Aleksander Lenert: Division of Immunology, Department of Internal Medicine, Carver College of Medicine, University of Iowa Hospitals & Clinics, Iowa City; [27]Edoardo Marrani, Gabriele Simonini: Rheumatology Unit, ERN ReCONNET Center, Meyer Children's Hospital IRCCS, Florence, Italy; [28]Doaa Mosad Mosa: Rheumatology Department, Mansoura University Hospitals, Mansoura University Faculty of Medicine, Mansoura, Egypt; [29]Alexander Mushkin: Science-Research Institute of Phthisiopulmonology, Saint-Petersburg, Russia; [30]Farzana Nuruzzaman: Renaissance School of Medicine at Stony Brook University, Stony Brook, New York; [31]Karen Onel: Division of Pediatric Rheumatology, Hospital for Special Surgery, New York, New York; [32]Lauren Potts: Patient Research Partner, California; [33]Athimalaipet V. Ramanan: Bristol Royal Hospital for Children & Translational Health Sciences, University of Bristol, Bristol, UK; [34]Angelo Ravelli: IRCCS Istituto Giannina Gaslini and Università degli Studi di Genova, Genoa, Italy; [35]Andrew W. Grim: Division of Pediatric Rheumatology, Department of Pediatrics, Michigan Medicine, Ann Arbor; [36]Micol Romano: Pediatric Rheumatology, ASST Gaetano Pini-CTO, Milan, Italy, and Department of Pediatrics, Division of Pediatric Rheumatology, Behcet and Autoinflammatory Disease Center, Western University, London, Ontario, Canada; [37]Takashi Shawn Sato: Stead Family Children's Hospital, University of Iowa Hospitals and Clinics, Iowa City; [38]Sara M. Stern: University of Utah, Salt Lake City; [39]Angela Taneja Kohli: Emory University School of Medicine/Children's Healthcare of Atlanta, Georgia; [40]Alexander C. Theos: Patient/Parent Partner, Georgetown University, Washington, DC; [41]Leslie F. Vogel: Department of Rehabilitation Seattle Children's Hospital, Seattle, Washington; [42]Shima Yasin, Polly J. Ferguson: Stead Family Department of Pediatrics, Carver College of Medicine, University of Iowa, Iowa City; [43]Katerina Bouchalova: Department of Pediatrics, Faculty of Medicine and Dentistry, Palacky University Olomouc and University Hospital, Olomouc, Czech Republic; [44]Alison M. Hendry: General Medicine and Rheumatology Service, Division of Medicine, Middlemore Hospital Counties Manukau District Health, Auckland, New Zealand; [45]Kevin C. Cain: Department of Biostatistics, University of Washington, Seattle; [46]Hermann J. Girschick: Department of Pediatrics, Vivantes Clinic Friedrichshain, Berlin, Germany; [47]Christian M. Hedrich: Department of Paediatric Rheumatology, Alder Hey Children's NHS Foundation Trust, Eaton Road, and Department of Women's and Children's Health, Institute of Life Course and Medical Sciences, University of Liverpool, Liverpool, UK; [48]Ronald M. Laxer: The Hospital for Sick Children, St. Michael's Hospital and the University of Toronto, Ontario, Canada.

[†]Dr. Naden is deceased.

Address correspondence via email to Dr. Yongdong Zhao, at yongdong.zhao@seattlechildrens.org.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/art.43137).

Submitted for publication September 11, 2024; accepted in revised form November 14, 2024.

CNO throughout this article for consistency. CNO is difficult to characterize and diagnose because of its insidious onset, variable bone location and symptoms at presentation, and lack of high-quality clinical research describing the disease spectrum. These factors, together with limited awareness among physicians about CNO, often result in misdiagnosis[1] and misclassification for research.

There are no previous classification criteria for CNO. Classification criteria are essential to identify homogeneous groups of patients to better characterize the epidemiology and natural history of CNO and to enroll in clinical trials. In the absence of classification criteria, diagnostic criteria proposed for clinical use have been used as inclusion criteria for research in CNO. CNO diagnostic criteria have been proposed by 3 groups: King et al[8] and Manson et al[9], Jansson et al[10], and Roderick et al.[11] These diagnostic criteria were based on clinical experience from single center cohorts with fewer than 100 patients each and have not been validated in independent cohorts. Thus, with the lack of evidence on the validity of these diagnostic criteria, they should not be used as entry criteria for clinical trials or observational studies.

The roles of advanced imaging methods and bone biopsy need to be carefully evaluated in classification criteria for CNO. Previous CNO diagnostic criteria required presence of lytic and/or sclerotic lesions in radiographs or hyperintensity within bone marrow in fluid-sensitive magnetic resonance imaging (MRI) sequences. Symmetric bone lesions and multifocal lesions occur much more frequently than unifocal lesions.[5,10] Whole-body MRI (WBMRI) is a relatively new modality that better defines patterns of lesion distribution and often identifies clinically 'silent' lesions that may affect the vertebral spine, which can impact treatment choice.[2,12–14]

Bone biopsy may be needed to exclude differential diagnoses of CNO, especially infections and malignancy.[10,11] However, there are no pathognomonic biopsy features for CNO. Common histologic findings include acute and chronic inflammation, marrow fibrosis, or osteonecrosis, and sometimes biopsy reveals only normal bone, likely due to sampling error.[10,15–18] Fibrosis in bone biopsy is present in children with prolonged duration of disease.[19] Microbiological investigations including special stains, cultures, or polymerase chain reaction tests are negative for bacteria, fungus, and mycobacteria, except for occasional contaminant organisms.[20] We aimed to develop the first classification criteria for pediatric CNO combining international expert consensus and data-driven methods and including the patient's perspective.[21,22]

METHODS

Study overview. As in prior EULAR/American College of Rheumatology (ACR) classification criteria[23–27], the study had 4 phases: (1) generation of candidate criteria items, (2) item reduction and definition of items, (3) criteria weighting and identification of threshold for CNO classification, and (4) refinement and validation. The data that support the findings of this study are available from the corresponding author (YZ) upon reasonable request.

Investigators. The study was led by a core group (YZ, MSO, AS, EYW, AMH, KCC, HJG, FD, CMH, PJF, RN, and SO) and overseen by a steering committee of 20 members (core group plus AFJ, LP, AVR, AR, SMS, ACT, and LFV) from North America, Europe (including Turkey), and New Zealand (Supplement S1). Steering committee members were nominated based on CNO expertise or methodological experience. In addition to academic pediatric rheumatologists (7 from Europe and 7 from North America), the steering committee included 1 patient with CNO, 1 parent of a patient with CNO, a physical therapist, methodologists, and a biostatistician. The CNO workgroup of the Childhood Arthritis and Rheumatology Research Alliance (CARRA) assisted with item reduction, and additional international collaborators participated by submitting cases of CNO and mimicking conditions used in the development and validation cohorts for the study.

Phase I: Item generation. Candidate items were identified using a free-text survey sent to members of CARRA, the German Society for Pediatric Rheumatology, the Eurofever registry, and the Dr Peter Dent pediatric rheumatology listserv between September 2017 and May 2018. Responders were asked to list features relevant to CNO and individually rank them in order of importance for classification.

Phase II: Item reduction and definition. Candidate items generated from phase I were reduced and defined in a CARRA CNO work group meeting held in Denver, Colorado, in April 2018 followed by a postmeeting survey. Potential items were retained or removed if 80% agreement of all participants at the meeting was achieved using a nominal group technique. In the postmeeting survey, members of the CARRA CNO work group were asked to score each item as follows: +3 (increases the likelihood of CNO the most), +2 (increases the likelihood of CNO moderately), +1 (increases the likelihood of CNO slightly), 0 (no difference), −1 (decreases the likelihood of CNO slightly), −2 (decreases the likelihood of CNO moderately), −3 (decreases the likelihood of CNO the most). Survey results were discussed by the core group, and no threshold was set to include/exclude these items.

The items remaining, and the items in previously proposed diagnostic criteria[10,11], were included in a case report form (CRF; Supplement S2) used to collect data on CNO cases and mimickers to assemble a development and a validation cohort.[10,11] Pediatric rheumatologists in North America, Europe, Australia, Africa, and South America used the CRF to submit cases and mimickers to an online central database using REDCap software.[28] Seattle Children's Hospital Institution Review Board approved the study as the coordination center (#1115),

and each participating site obtained approval from their institutions. Data received between December 2018 and August 2020 formed the development cohort and data received between August 2020 and May 2021 formed the validation cohort. Submitted CNO cases and mimickers were required to have bone pain, ≥1 imaging study (ie, x-ray and/or MRI), and ≥12 months of follow-up (except for mimickers with a confirmed laboratory or pathologic diagnosis) because of the concern of inadequate observation for mistakenly diagnosed CNO or mimickers. Institutional Review Board or ethics committee approval was obtained at each site. Collaborators submitting CRFs had no prior knowledge of the items likely to advance to final consideration as classification criteria and provided an assessment of their degree of confidence in the cases being CNO or a mimicker, rated from +3 to –3.

Once the development cohort was assembled in 2020, candidate items were further reduced using the prevalence ratio (PR) of each item as a predictor of 'CNO.' PRs were calculated using only cases with moderate to high confidence (ie, +2 or +3 or –2 or –3) according to the submitting physicians. PRs >2 or <0.5 suggested distinguishing features for or against CNO, and a P value <0.05 was considered significant.

The core group met every other week for 24 weeks to consider PRs, discuss example cases or mimickers, further define criteria, and group them in domains and levels within a domain. Thirty ambiguous cases from the development cohort with low confidence levels by the submitting physician (ie, –1 to +1 scores) representing as many clinical features as possible were used by the core group to help establish hierarchical levels within each domain in preparation for multicriteria decision analysis (MCDA).

The CNO classification draft domains prepared by the core group were sent to the steering committee. Members were asked to either agree or disagree with each item and level based on the statistical analysis, clinical feasibility, and summary of discussions from the core group. Responses were reviewed at the first steering committee meeting.

Phase III: Criteria definition and weighting. Ten virtual consensus meetings of the steering committee were held between March 2021 and October 2023 to refine the draft prepared by the core group and assign preliminary weights using MCDA facilitated by 1000minds software (1000minds decision-making and conjoint analysis software, https://www.1000minds.com). The steering committee was presented with hypothetical pairs of cases contrasting 2 domains at a time (Supplement S3) and asked to decide individually and anonymously which case of the pair they were more likely to classify as CNO. Importantly voters were to assume the 2 cases were equal in all other features.

The participants then reviewed the tally of votes together. One hundred percent agreement was considered consensus without further discussion. Incomplete agreement prompted further discussion, moderated by a methodologist (AMH), to achieve consensus as to which case was more likely to be classified as CNO. Discussion was encouraged and facilitated until all members reached consensus. Complex choices in which consensus was not readily achieved were deferred for additional voting rounds. As a result of these forced choices between the alternative pairs, candidate items were weighted and ranked using 1000minds decision-making software.

Threshold identification. The draft weights derived from MCDA were applied to each case of the development cohort. A different set of 30 CNO cases and mimickers that covered the entire range of confidence from very likely to very unlikely CNO, based on the opinion of the submitting clinician, were selected by the statistician, and the case/mimicker status was confirmed by a 3-expert group (SO, AS, and YZ). The total scores for these 30 cases and mimickers were used to examine the performance of the draft weights and to determine a threshold for classification.

The submitting physician's diagnosis was considered the gold standard unless there was disagreement from the 3-expert group. This contrary opinion and its rationale were shared with the submitting physician and they were asked to review their diagnosis. The submitting physician's subsequent diagnosis after review of the expert commentary became the final diagnosis.

An initial threshold score for CNO was identified by examining the additive scores of all cases ranked from high to low. A receiver operating characteristic (ROC) curve analysis was conducted using the additive scores for all cases in the development cohort. This process identified an optimal threshold to distinguish CNO cases from mimickers.

Phase IV: Refinement and validation. The final criteria domains were examined by the steering committee based on the misclassified cases to ensure the clarity of the definitions for the differentially weighted classification system. The updated definitions with minimal change aimed to enhance the ease of application, sensitivity, and specificity of the classification system.

The final criteria were then applied to every CNO case or mimicker confirmed by adjudicators in the validation cohort to obtain their final score and their classification as CNO or not, relative to the agreed threshold. Adjudicators consisted of 5 rheumatologists who did not participate in the steering committee or submit any cases. They adjudicated 600 to 660 cases to ensure triple coverage of each case to verify the diagnosis of cases within the development cohort and exclude disagreed cases from the validation cohort. Sensitivity and specificity of the criteria were calculated and contrasted with the sensitivity and specificity of previously proposed diagnostic criteria.[10,11] Because the original criteria by Jansson et al[10] did not include MRI findings, we tested a version of the Jansson et al criteria modified by adding 'typical MRI findings' defined in the same way as for the new EULAR/ACR classification criteria.

Patient and public involvement. Caregivers of children with CNO and adults who were pediatric patients of CNO were involved in setting the research question and they were intimately involved in design and consensus meetings.

RESULTS

Phase I: Item generation. Overall, 259 of 865 (30%) of surveyed international pediatric rheumatologists responded and provided candidate items for CNO classification. Among responders, 132 (51%) practiced in North America, 77 (30%) in Europe, and 50 (19%) in other continents (South America 9, Asia 8, Australia/New Zealand 9, and other 14). A total of 138 responders (53%) had >10 years of practicing experience, and 108 (42%) had diagnosed and treated >10 CNO patients. The 33 items proposed in free-text survey responses were categorized into 6 domains (clinical presentation, physical examination, laboratory findings, imaging findings, bone biopsy, and treatment response). This is a smaller number of items than in recent criteria for other rheumatic disorders.[24,27] The following 8 items were ranked most important by survey respondents: (1) exclusion of malignancy and infection by bone biopsy, (2) multifocal bone lesions, (3) presence of bone pain or swelling, (4) signs of fibrosis or inflammation in bone biopsy, (5) bone lesions in typical locations (clavicle, metaphysis of long bones, mandible, vertebra), (6) presence of CNO-related comorbidities, (7) normal or mildly elevated C-reactive protein (CRP) or erythrocyte sedimentation rate (ESR), and (8) MRI findings 'typical' of CNO.

Phase II: Item reduction and definition. At the 2018 CARRA CNO workgroup meeting, 39 pediatric rheumatologists (35 from North America, 4 from Europe) and 7 parents of patients with CNO participated in item reduction. Twelve of the 33 items were thoroughly discussed, and the rest deferred to a postmeeting survey. Consensus was reached to exclude 8 items (antinuclear antibody, human lymphocyte antigen B27, lactate dehydrogenase (LDH), uric acid, and alkaline phosphatase (ALP) tests, fever, arthritis, and good clinical response to nonsteroidal anti-inflammatory drugs). LDH, low ALP, and fever were reconsidered in phase III. Physical examination findings were split into 3 items, and ESR and CRP values were expanded from 2 to 6 items to capture the range more accurately. Thus, the total number of items was only modestly reduced from 33 to 31.

In the postmeeting survey of all members of the CARRA CNO work group, 77 of 82 members (94%) scored the 31 items from +3 (increases the likelihood of CNO the most) to –3 (decreases the likelihood of CNO the most). Respondents included 62 pediatric rheumatologists, 9 pediatric rheumatology trainees, 2 adult rheumatologists, and 2 patient/parent representatives. Seventy percent of practitioners had >6 years of experience, and 34% managed >10 CNO patients every year. Supplement S4 shows the average scores for the 31 items. The items with the highest scores were 'multifocal lesions on imaging,' 'ruling out malignancy and infection by biopsy,' and 'typical location on imaging.' The item with the lowest scores (decreases likelihood of CNO the most) was 'CRP and ESR >3 times upper limit of normal.' Results were presented to the steering committee at their first meeting.

CNO cases and mimickers. Demographic information for cases included in the development and validation cohorts is summarized in Table 1. The diagnoses of mimicker cases are reported in Table 2. Among 441 CRFs included in the development cohort, 264 of 285 (93%) CNO cases and 145 of 156 (93%) mimickers had moderate or greater confidence in the diagnosis reported by the submitting physician and were used for calculation of PRs (Supplement S5). Adjudicators disagreed with the diagnosis of 2 CNO cases and 2 mimicker cases that were included in PR calculations. Data missing rates were also assessed. CNO cases more frequently had intermittent pain, clavicular swelling, symmetric bone lesions, and involvement of the thoracic spine, clavicle, sternum/manubrium, pelvic bones, bilateral femur, bilateral tibia, unilateral fibula, and foot bones. Mimickers more frequently had continuous pain, fever, active arthritis (examination findings or imaging findings of arthritis), or imaging findings suggesting infection or malignancy. Finally, complete and sustained response to antibiotic treatment was more frequent in mimickers.

Based on the calculated PRs and discussions of ambiguous cases by the core group (those cases with less than moderate confidence in the diagnosis as reported by the submitting physician, n = 30), candidate criteria were further reduced from 31 to 14, and the levels or options within each domain were defined. Findings from physical exams, disease durations, and pain patterns were among the 17 items removed after reviewing PRs, which were between 0.5 and 2. LDH, fever, and axial arthritis as part of coexisting conditions were added back to the criteria because of their PRs and incidence of >5% in the total population.

Phase III: Criteria definition and weighting. *Phase IIIa: Criteria definition.* Based on discussions by the steering committee, further refinements were made to the definitions for bone lesion pattern and distribution with consensus. The most specific bone sites were determined as the mandible and clavicle, and the least specific sites for CNO were the skull and hands. LDH levels were removed from the classification domains and 'pathological LDH concerning for malignancy' was added as exclusion criteria instead. Although an LDH >700 international units/L had a significant PR of <0.5 to distinguish CNO and mimickers, there was variability across laboratory reporting systems and a lack of expert agreement to establish a cutoff value. Platelet levels were also removed from the classification domains and 'platelets <100,000/mm3' was added as exclusion criteria instead due to concerns for malignancy. Family history of associated conditions, defined as reported axial arthritis, inflammatory bowel disease,

Table 1. Demographic, clinical, and laboratory characteristics of subjects included in development and validation cohorts*

Variables	Development cohort		Validation cohort	
	CNO (n = 285)	Mimickers (n = 156)	CNO (n = 382)	Mimickers (n = 132)
Age at onset (y)	9.5 (3.3)	8.6 (5.1)	9.2 (3.3)	7.3 (5.5)
Time to diagnosis (mo)	13.0 (21.1)	4.6 (8.0)	13.6 (20.5)	6.1 (17.3)
Male sex at birth	115 (40.4%)	96 (61.5%)	159 (42.1%)	84 (63.6%)
Geographic locations				
North America	141 (49.5%)	94 (60.3%)	216 (56.5%)	74 (56.1%)
South America	0 (0%)	0 (0%)	13 (3.4%)	0 (0%)
Europe	117 (41.1%)	59 (37.8%)	147 (38.5 %)	55 (41.7%)
Asia	18 (6.3%)	1 (0.6%)	1 (0.3%)	1 (0.8%)
Africa	0 (0%)	0 (0%)	3 (0.8%)	1 (0.8%)
Oceania	9 (3.2%)	2 (1.3%)	0 (0%)	1 (0.8%)
Other	0 (0%)	0 (0%)	2 (0.5%)	0 (0%)
Race/ethnicity				
Asian	12 (4.2%)	8 (5.1%)	21 (5.5%)	6 (4.5%)
Black	5 (1.8%)	8 (5.1%)	9 (2.4%)	6 (4.5%)
Hispanic/Latino	6 (2.1%)	8 (5.1%)	9 (2.4%)	4 (3.0%)
Middle Eastern/North African	5 (1.8%)	4 (2.6%)	6 (1.6%)	2 (1.5%)
Native American	0 (0%)	4 (2.6%)	3 (0.8%)	1 (0.8%)
Native Hawaiian/Pacific Islander	0 (0%)	0 (0%)	0 (0%)	2 (1.5%)
White	246 (86.3%)	105 (67.3%)	310 (81.2%)	100 (75.8%)
Multiracial	5 (1.8%)	3 (1.9%)	10 (2.6%)	3 (2.3%)
Unknown	6 (2.1%)	16 (10.3%)	14 (3.7%)	8 (6.1%)
Bone biopsy results available	187 (65.6%)	93 (59.6%)	241 (63.1%)	114 (86.4%)
Hemoglobin (g/dL)	12.5 (11.8-13.2)	11.6 (10.3-13.0)	12.3 (11.4-13.2)	11.6 (10.0-12.9)
Platelets (10^9/L)	345 (293-414)	333 (255-427)	330 (275-395)	300 (253-395)
ESR (mm/h)	21 (11-44)	33 (15-73)	24 (10-43)	25 (10-47)
C-reactive protein (mg/L)	8 (3-16)	12 (6-44)	8 (5-15)	12 (3-44)

* Variables are presented as mean (SD), median (median interquartile range), or n (%). The development cohort included 3 CNO cases and 3 mimicker cases that adjudicators disagreed with submitting physicians. The validation cohort excluded 5 CNO cases that adjudicators disagreed with submitting physicians. CNO, chronic nonbacterial osteomyelitis; ESR, erythrocyte sedimentation rate.

psoriasis in a first-degree relative, and an ESR/platelet ratio ≥0.2 (relative thrombocytopenia during inflammation) were later removed due to insignificant weight relative to the other criteria. The remaining criteria were hierarchically organized, and definitions for the domains and their levels were further refined by the steering committee. The final 9 domains include sites of bone lesions and distribution patterns of these lesions based on imaging, bone biopsy pathological findings, age at the onset of symptoms, coexisting conditions, hemoglobin level, associated fever history, ESR, and CRP (Table 3).

The steering committee reached consensus on entry and exclusion criteria to be applied before scoring the 9 domains (5 clinical and 4 pathology/laboratory domains) (Table 3). Entry criteria are as follows: (1) aged <18 years at symptom onset, (2) presence of bone pain and/or functional musculoskeletal limitation for ≥6 weeks, and (3) abnormal radiographic or MRI imaging findings at nonarthritic bone sites. Nonarthritic bone sites refer to bones whose endings, if forming a joint, do not demonstrate signs of arthritis on physical examination or imaging (ie, no synovial thickening, contrast enhancement, or moderate to large joint effusions). Exclusion criteria are as follows: (1) confirmatory evidence of mutually exclusive mimicker diseases such as malignancy, infection, scurvy, or hypophosphatasia; (2) pathological LDH concerning for malignancy; (3) platelets <100,000/mm3; and (4) complete and sustained clinical and laboratory response to antimicrobial treatment alone.

Phase IIIb: Criteria weighting. During 8 virtual consensus meetings of the steering committee, 59 MCDA pairwise trade-off decisions were made. Criteria weights were calculated by the 1000minds software based on these decisions. The final list of criteria and their assigned weighted scores are shown in Table 3.

Phase IV: Refinement and validation. During the initial weighting, LDH, response to antibiotics, family history, and ESR/platelet ratio were included. LDH and complete response to antibiotics were moved to exclusion criteria due to the difficulty to define absolute threshold and >50% missing data. The latter 2 domains were removed due to low weight (1%–2%). The optimal threshold for CNO classification, based on Youden index of the ROC curve analysis in the development cohort (Supplement S6), was identified as 55 of a maximum of 100 when inclusion criteria and the composite domain score were applied. Cases that would have met exclusion criteria were kept in the pool for the exercise of threshold setting. There was a mixed cluster of CNO cases and mimicker cases in the development cohort with weighted scores between 52 and 54. These CNO cases did not have a bone biopsy, and it was probable that

Table 2. Detailed distribution of mimicker cases within development and validation cohorts*

Mimicker diagnosis	Development (n = 156)	Validation (n = 32)
Infection	49 (31%)	57[a] (43%)
Malignancy	61 (39%)	61 (46%)
Inflammatory arthritis	21 (13%)	4 (3%)
Other mimickers	25 (16%)	10 (8%)
Vitamin C deficiency	10	2
Benign bone tumor	7	4
Hypophosphatasia	1	2
Amplified pain syndrome	2	0
Fracture	1	0
Hyperostosis hypophosphatemia	1	0
NEMO deficiency syndrome	1	0
Traumatic injury	1	0
Unicameral bone cyst	1	0
Other autoinflammatory disease	0	1
Vascular malformation	0	1

* The development cohort included 1 case of inflammatory arthritis, 1 case of hyperostosis hypophosphatemia, and 1 case of amplified pain syndrome that adjudicators disagreed with submitting physician. NEMO, nuclear factor kappa B essential modulator.
[a] Includes 33 patients with bone tuberculosis.

biopsy findings could have moved their weighted score above the threshold for classification. After discussion, the steering committee agreed that a patient should be classified as having CNO if their score was ≥55. This cutoff value corresponded to a sensitivity of 92% and a specificity of 99% in the development cohort.

The final CNO criteria were then applied to the validation cohort to determine their performance. Given that no previous CNO classification criteria exist, their performance was compared to proposed diagnostic criteria by Jansson et al[10] and Roderick et al.[11] The demographic information of the validation cohort (n = 382 CNO patients) is summarized in Table 1, and the distribution of mimickers (n = 132) is presented in Table 2. The new EULAR/ACR CNO classification criteria had a sensitivity of 82% (95% confidence interval [CI], 78-86) and specificity of 98% (95% CI, 94-100) in the validation cohort. Two false positive cases had benign bone tumors. Sensitivity compared favorably to the Roderick et al diagnostic criteria (63%) and similarly to the modified Jansson et al criteria (85%) while specificity was similar to the Roderick et al criteria and superior to the Jansson et al criteria (Table 4).[10,11] Since all 3 criteria include bone biopsy results, a bone biopsy showing evidence of infection or malignancy is considered an exclusion for all 3 criteria, ie, the case is not classified as CNO if such evidence is found. However, the other exclusions given in Table 3 are only applied to the EULAR/ACR classification criteria because they are not specified as exclusions in the other 2 criteria. In our dataset, confirmatory evidence of mutually exclusive mimicker diseases was based on vitamin C measurement, bone marrow aspiration biopsy, blood culture, and test for tuberculosis or atypical mycobacteria, in addition to bone biopsy, which was the source of such evidence in most mimicker cases. In the validation cohort, 88% of mimicker cases met one of the exclusion criteria in Table 3, including 69% with a bone biopsy showing evidence of infection or malignancy. That said, a bone biopsy is not required to apply the classification criteria, and approximately one-third of all CNO cases had no bone biopsy.

DISCUSSION

The EULAR/ACR endorsed classification criteria for pediatric CNO presented here represent a milestone in CNO care and research. Through a 4-phase, iterative process, an international collaborative group defined weighted criteria to classify CNO, distinguishing it from mimicker conditions. The excellent specificity (98%) and sensitivity (82%) of these criteria will help define homogenous CNO cohorts for research studies. A key strength of the criteria is that patients can be classified accurately as having CNO, even in the absence of bone biopsies. However, it is important to reiterate that classification criteria are not diagnostic criteria and that biopsies may be essential in certain settings to exclude differential diagnoses of CNO.

These EULAR/ACR pediatric CNO classification criteria were built on a robust and widely recognized multiphase methodological approach that defined entry and exclusion criteria. The absolute exclusion criteria cover important mimickers of pediatric CNO such as malignancy, infection, and metabolic bone disease. Scurvy and hypophosphatasia can be diagnosed via laboratory testing and may present with a clinical picture resembling CNO, and it is important for researchers to exclude these 2 conditions prior to classifying a patient as CNO. In prospective use of the criteria to select subjects for future studies, exclusion criteria are to be applied during case review before attempting to add criteria weights. If confirmatory evidence of an alternative diagnosis such as malignancy or vitamin C deficiency is already available, if the patient had received monotherapy with antibiotics that fully resolved the symptoms, or if thrombocytopenia or marked elevation of LDH are documented, the patient should not be classified as CNO. These exclusion criteria do not mean that every subject needs to have bone biopsies or have received full courses of antibiotics before they can be classified as CNO. If none of these is documented, then proceed to add criteria weights.

Some patients with a clinical diagnosis of pediatric CNO will not fulfill classification criteria. We focused classification criteria development efforts on patients with 'typical' and 'common' manifestations with the goal of enrolling homogeneous populations into observational studies and clinical trials. As reported here, some patients with a diagnosis of CNO without bone biopsy did not accrue sufficient points to be classified as CNO. It is possible that if a biopsy was performed, those patients would have had sufficient points for classification. That said, in the hands of experienced clinicians, biopsy may not be necessary in all cases.

Table 3. EULAR/ACR classification criteria for pediatric chronic nonbacterial osteomyelitis*

Step 1, Verify entry criteria (ALL should be present):	
1. Bone pain and/or musculoskeletal functional limitation ≥6 wk	
2. Age of onset <18 y	
3. Abnormal findings from radiograph and/or advanced imaging including MRI, CT, bone scintigraphy at nonarthritic bone sites[a]	
Step 2, Verify exclusion criteria (NONE should be present):	
1. Confirmatory evidence of mutually exclusive mimicker diseases[b]	
2. Platelet <100,000/mm3	
3. Pathological LDH concerning for malignancy[c]	
4. Complete and sustained clinical and laboratory response to antimicrobial treatment alone	
Step 3, Add the scores: Add the highest value from each of the 9 domains below. A score of ≥55 is required to classify a patient as having CNO.	
Criteria domains/levels	
Bone biopsy[d]	Score
• Signs of inflammation[e] AND fibrosis	17
• Signs of fibrosis only	11
• Signs of inflammation only	9
• No signs of inflammation or fibrosis	0
Age at the onset of symptoms	
• ≥3 y	17
• <3 y	0
Sites of bone lesions based on imaging	
• Clavicle and/or mandible	16
• Sites other than clavicle, mandible, skull, or hand	9
• Skull or hand without clavicle or mandible	0
Distribution pattern of bone lesions based on imaging	
• Multifocal lesions (in ≥2 bones) with symmetrical pattern (bilateral involvement of at least one bone)	9
• Multifocal without any symmetrical pattern of bone involvement	7
• Unifocal without whole-body imaging such as WBMRI, PET/CT performed	3
• Unifocal with whole-body imaging performed	0
Coexisting conditions prior to the diagnosis of CNO	
• Inflammatory bowel disease (IBD)	9
• Cutaneous condition[f] without IBD	8
• Axial arthritis[g] without cutaneous condition or IBD	5
• None	0
Hemoglobin (normal range varies by age)	
• ≥10 g/dL	8
• <10 g/dL	0
Fever (oral/temporal temperature above 38°C and not related to common infections)	
• Absence of fever	8
• Presence of fever	0
Erythrocyte sedimentation rate (normal range: 0-20 mm/h)	
• <60 mm/h	8
• ≥60 mm/h	0
C-reactive protein (normal range: 0-10 mg/L)	
• <30 mg/L	8
• ≥30 mg/L	0

* Fever is defined as temperature >38°C without common infections including clinical symptoms of upper airway infection or urinary tract infection or abscess. It does not need to be present throughout 6 wk but any stage between the onset and the diagnosis. ACR, American College of Rheumatology; CNO, chronic nonbacterial osteomyelitis; CT, computed tomography; LDH, lactate dehydrogenase; MRI, magnetic resonance imaging; PET, positron emission tomography; WBMRI, whole-body MRI.
[a] Typical imaging findings include but are not limited to lytic, sclerotic lesions with or without periosteal reaction, hyperostosis on radiograph or CT, bone marrow edema or hyperintensity in a fluid-sensitive sequence of MRI. Nonarthritic bone sites refer to bones whose joint-forming ends do not demonstrate imaging signs of arthritis including synovial thickening, enhancement, and/or moderate to large joint effusions. See Supplement S7.
[b] Primary or metastatic malignancy in bone, leukemia, lymphoma, neuroblastoma, infectious osteomyelitis, metabolic bone disease, vitamin C deficiency, hypophosphatasia, and monogenic bone diseases including Majeed or deficiency of interleukin-1 receptor antagonist (DIRA).
[c] Due to differences in laboratory assays across institutions, no absolute threshold is defined. Levels ≥700 IU or >2 times the upper normal limit may be considered pathological.
[d] Bone biopsy is not required before applying these criteria.
[e] Defined as the presence of immune cells including neutrophils, monocytes, lymphocytes, and/or plasma cells.
[f] Cutaneous conditions include psoriasis, palmoplantar pustulosis, pyoderma gangrenosum, acne fulminans, and hidradenitis suppurativa.
[g] Axial arthritis is defined as imaging confirmation of inflammation within the sacroiliac joint or intervertebral joint.

Differential weighting of criteria represents their relative contribution to an individual child's classification as having CNO. For CNO, having a lesion at a typical site, such as the mandible and clavicle, in the absence of lesions at rarely affected sites, such as the skull and hands, carries the most weight. While the clavicle was proposed as a pathognomonic site by Roderick et al[11], based on

Table 4. Sensitivity and specificity of EULAR/ACR classification criteria compared to previously published diagnostic criteria using validation cohort data*

Criteria	Sensitivity	Specificity
EULAR/ACR classification criteria	82% (78%-86%)	98% (94%-100%)
Modified Jansson et al diagnostic criteria[a]	85% (81%-89%)	80% (72%-87%)
Roderick et al diagnostic criteria[b]	63% (58%-68%)	95% (89%-98%)

* ACR, American College of Rheumatology.
[a] Jansson et al diagnostic criteria require 2 of 4 major criteria or 1 of 4 major criteria and 3 of 6 minor criteria to diagnose chronic nonbacterial osteomyelitis.[10]
[b] Roderick et al diagnostic criteria require the presence of typical clinical findings and typical radiological findings and 1 of 2 criteria depending on bone sites and the number of lesions as well as the level of C-reactive protein.[11]

the personal experience of steering committee members, mimicker diseases, such as Langerhans cell histiocytosis and primary bone malignancy, may also present with clavicular lesions.[29] For this reason, clavicular lesions were not chosen as a sufficient criterion to classify as CNO. Mandibular CNO lesions have been reported extensively by oral maxillofacial surgeons[30–37], and it is important to recognize their usefulness for classifying an individual as having CNO.

The pattern of bone lesions (unifocal, multifocal, symmetrical) is also very important for CNO classification. Unifocal lesions confirmed by whole-body imaging are not characteristic of CNO. If whole-body imaging was not performed, multifocal or symmetrical lesions may be missed. Bone biopsy findings are not sufficient by themselves for CNO classification. This may suggest consideration of obtaining WBMRI before deciding on a bone biopsy.

Several items that may be useful for recognizing CNO in clinical practice, such as family history of associated diseases and ESR/platelet ratio, were not included in the criteria because they carried very low weight in the 1000minds exercise. Coexisting conditions such as inflammatory arthritis may pose a challenge in classifying CNO. Some children with inflammatory arthritis may subsequently develop CNO, and future longitudinal studies may help determine the prevalence of such occurrence by applying our CNO criteria.

Our study has several strengths. First, a large cohort of nearly 1000 patients with CNO or a mimicking condition was assembled by an international group of investigators. These criteria can be generalized to patients worldwide regardless of racial composition. Second, the experts involved in the consensus exercise and decision analysis (MCDA) also had international representation.

While it represents a significant step forward in planning randomized controlled trials and improving the quality of clinical and laboratory studies, this study also has limitations. First, the laboratory, imaging, and pathological findings were not measured or confirmed centrally, and detailed imaging findings were not available. Second, not all the cases had received all diagnostic tests considered within the inclusion criteria, including ESR, CRP, and hemoglobin. Therefore, the authors had to score 'not available' as '0.' Finally, ascorbic acid (vitamin C) deficiency, benign bone tumors, and hypophosphatasia are difficult to distinguish from CNO without applying exclusion criteria, which highlights the importance of obtaining appropriate history and laboratory results prior to applying classification criteria. The sensitivity and specificity results should be interpreted cautiously for 3 reasons. First, there is no true gold standard; the sensitivity and specificity presented reflect how well each of the classification rules agrees with the diagnosis given by the physician who submitted the case and confirmed by independent adjudicators. Second, specificity may be different if different mimicking conditions are included; the results presented here apply to the mix of mimicker conditions in our data and may be different with a different mix. Third, patients who had incomplete data may have been misclassified.

These EULAR/ACR endorsed classification criteria for pediatric CNO are the first criteria set developed through international collaboration using large development and validation cohorts and rigorous methodology. The criteria had excellent specificity (98%) and good sensitivity (82%). These criteria are easy to apply in a pediatric population because clinical, laboratory, and imaging findings are incorporated into the classification without the absolute requirement for a bone biopsy. The criteria will help researchers to easily identify a homogeneous study population for future clinical and laboratory studies.

DISCLOSURES

YZ: ACR, EULAR, CARRA (Grants to Institution, Finance Committee), Bristol-Myer Squibbs (Grants to Institution), Rheumatology Research Foundation (Grants to Institution), CTIP (Grants to Institution), UpToDate CRMO/CNO (Royalties/License), Novartis (Consulting Fees to Self), U.S. Provisional Application No. 63/356977 (Patent Issued), Seattle Children's CRMO Warriors Guild (Leadership/Fiduciary Role), American Board of Pediatrics, Pediatric Rheumatology Subboard (Leadership/Fiduciary Role). MSO: CARRA (AF-Small Grants from 2022 to 2024, Total: $48,399.00). AS: EULAR (Fellow with Funding of €15,000.00, Grants or Contracts). JDA: Novartis (Participation on Advisory Board 2023). MLB: Sjogren Foundation (Grants to Institution), Childhood Arthritis and Rheumatology Research Alliance (Grants to Institution). SKL: Time Hearst Foundation Physician Scientist Award (Grants), ACR (Support for Attending Meetings), Hackensack Meridian Health (Support for Attending Meetings), CARRA PFAPA/Autoinflammatory (Work Group Leader), APMC (Pediatrics Committee). TCL: Crohns & Colitis Congress Talk 2024 (Honoraria for Self, Support for Travel), Packard Children's Health Alliance Board of Directors (Board Member, Unpaid), Lupus Foundation of Northern California (Board Member, Unpaid). AL: National Institution of Health Award Number K23AR082966 (Grants to Institution). TSS: Invicro (Consulting Fees to Institution), Hippo Education (Payment to Self). GS: Sobi (Consulting Fees), Novartis (Consulting Fees), AbbVie (Consulting Fees, Payment/Honoraria). LBT: Cassie & Friends

Society (Board Member, Unpaid). LFV: Trinds 2022-2025 (Contract for Trials, Payment to Self), Wolters Kluwer 2021 (Payment for Authorship to Self), Scholar Rock 2024 (Support for Travel). SCW: Arthritis Foundation Great West Region (Medical Advisory Board, Unpaid), Aicardi Goutieres Syndrome Advocacy Association (ECHO Faculty, Unpaid). KB: Czech Ministry of Health (DRO FNOl, 00098892, Support for the Manuscript). AMH: Seattle Children's Research Institute (Support for the Manuscript, Consulting Fees). FD: UpToDate (Royalties or Licences), ISSAID (Support for Attending Meetings in 2023, Education Committee Chair). CMH: Merck (Unrestricted Research Grant), Novartis (Unrestricted Research Grant Ended in 2022, Advisory Board on the Use of IL-1 blockers in CNO). RML: UpToDate (Royalties or Licences), Elsevier (Royalties or Licences), Sanofi (Consulting Fees to Institution, Participation on Board), Eli Lilly Canada (Consulting Fees to Self, Participation on Board), Sobi (Consulting Fees to Self), Novartis (Consulting Fees to Self, Participation on Board), Akros Pharma (Consulting Fees to Self), Children's Hospital of Philadelphia (Participation on Board). PJF: American Board of Pediatrics (Chair, Payment or Honoraria to Self), Pediatric Rheumatology Subcommittee (Chair, Payment or Honoraria to Self), Current Opinion in Rheumatology Section Editor (Payment or Honoraria to Self), CARRA (Support for Travel), CNO Meeting in Liverpool 2022 (Support for Travel), PRES Invited Speaker 2022 (Support for Travel). RN: ACR/EULAR Development for Classification for Anti-Phospholipid Syndrome and IgG4-Related Diseases (Consulting Fees), Genentech Development of Pediatric Glucocorticoid Toxicity Index (Consulting Fees, Support for Travel). SO: Novartis (Consulting Fees), Sobi (Consulting Fees). EYW, ZW, AM, CLA, UKA, BA, SA, EB-D, OB, YB, DAC, MC, ND, BHE, EJF, AI, RSI, AFJ, IK, MK, LKK, KTK, KM, EM, DMM, IM, AM, KDN, FN, KO, MP, TSP, LP, AVR, AR, NDR, AWG, MR, NR, JBS, SMS, TS, ATK, ACT, SY, KCC, and HJG have no conflicts of interest to disclose.

ACKNOWLEDGEMENT

We would like to thank Dr Sindhu Johnson for generously sharing her experience and assisting us during the planning phase of this project. American College of Rheumatology and EULAR provided funding and guidance through the completion of the project via Drs Amy Turner, Jaime Guzman, Zahi Touma, and Simona Lupatin. We appreciated the dedicated assistance in data entry from each center. Authors would like to thank all members of the CNO/Chronic Recurrent Multifocal Osteomyelitis (CRMO) workgroup who have participated in the discussion of this project and the surveys. Funding from Seattle Children's Research Institute and CRMO Warriors Guild made initial planning of the project possible.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr Zhao confirms that all authors have provided the final approval of the version to be published, and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. Jansson AF, Grote V, Study Group ESPED. Nonbacterial osteitis in children: data of a German Incidence Surveillance Study. Acta Paediatr 2011;100(8):1150–7.
2. Aden S, Wong S, Yang C, et al. Increasing cases of chronic nonbacterial osteomyelitis in children: a series of 215 cases from a single tertiary referral center. J Rheumatol 2022;49(8):929–34.
3. Schnabel A, Range U, Hahn G, et al. Unexpectedly high incidences of chronic non-bacterial as compared to bacterial osteomyelitis in children. Rheumatol Int 2016;36(12):1737–45.
4. Hofmann SR, Kapplusch F, Girschick HJ, et al. Chronic recurrent multifocal osteomyelitis (CRMO): presentation, pathogenesis, and treatment. Curr Osteoporos Rep 2017;15(6):542–54.
5. Wipff J, Costantino F, Lemelle I, et al. A large national cohort of French patients with chronic recurrent multifocal osteitis. Arthritis Rheumatol 2015;67(4):1128–37.
6. Catalano-Pons C, Comte A, Wipff J, et al. Clinical outcome in children with chronic recurrent multifocal osteomyelitis. Rheumatology (Oxford) 2008;47(9):1397–9.
7. Huber AM, Lam PY, Duffy CM, et al. Chronic recurrent multifocal osteomyelitis: clinical outcomes after more than five years of follow-up. J Pediatr 2002;141(2):198–203.
8. King SM, Laxer RM, Manson D, et al. Chronic recurrent multifocal osteomyelitis: a noninfectious inflammatory process. Pediatr Infect Dis J 1987;6(10):907–11.
9. Manson D, Wilmot DM, King S, et al. Physeal involvement in chronic recurrent multifocal osteomyelitis. Pediatr Radiol 1989;20(1-2):76–9.
10. Jansson A, Renner ED, Ramser J, et al. Classification of non-bacterial osteitis: retrospective study of clinical, immunological and genetic aspects in 89 patients. Rheumatology (Oxford) 2007;46(1):154–60.
11. Roderick MR, Shah R, Rogers V, et al. Chronic recurrent multifocal osteomyelitis (CRMO) – advancing the diagnosis. Pediatr Rheumatol Online J 2016;14(1):47.
12. Andronikou S, Mendes da Costa T, Hussien M, et al. Radiological diagnosis of chronic recurrent multifocal osteomyelitis using whole-body MRI-based lesion distribution patterns. Clin Radiol 2019;74(9) 737.e3–15.
13. Zhao Y, Sato TS, Nielsen SM, et al. Development of CROMRIS (ChRonic nonbacterial Osteomyelitis MRI Scoring) tool and evaluation of its interrater reliability. J Rheumatol 2019;47(5):739–47.
14. Capponi M, Pires Marafon D, Rivosecchi F, et al. Assessment of disease activity using a whole-body MRI derived radiological activity index in chronic nonbacterial osteomyelitis. Pediatr Rheumatol Online J 2021;19(1):123.
15. Borzutzky A, Stern S, Reiff A, et al. Pediatric chronic nonbacterial osteomyelitis. Pediatrics 2012;130(5):e1190–7.
16. Zhao Y, Dedeoglu F, Ferguson P, et al. Physicians' perspectives on the diagnosis and treatment of chronic nonbacterial osteomyelitis. Int J Rheumatol 2017;2017:7694942.
17. Girschick HJ, Krauspe R, Tschammler A, et al. Chronic recurrent osteomyelitis with clavicular involvement in children: diagnostic value of different imaging techniques and therapy with non-steroidal anti-inflammatory drugs. Eur J Pediatr 1998;157(1):28–33.
18. Girschick HJ, Raab P, Surbaum S, et al. Chronic non-bacterial osteomyelitis in children. Ann Rheum Dis 2005;64(2):279–85.
19. Morbach H, Hedrich CM, Beer M, et al. Autoinflammatory bone disorders. Clin Immunol 2013;147(3):185–96.

20. Girschick HJ, Huppertz HI, Harmsen D, et al. Chronic recurrent multifocal osteomyelitis in children: diagnostic value of histopathology and microbial testing. Hum Pathol 1999;30(1):59–65.
21. Singh JA, Solomon DH, Dougados M, et al. Development of classification and response criteria for rheumatic diseases. Arthritis Rheum 2006;55(3):348–52.
22. Dougados M, Gossec L. Classification criteria for rheumatic diseases: why and how? Arthritis Rheum 2007;57(7):1112–5.
23. Aringer M, Costenbader K, Daikh D, et al. 2019 European League Against Rheumatism/American College of Rheumatology classification criteria for systemic lupus erythematosus. Arthritis Rheumatol 2019;71(9):1400–12.
24. Aringer M, Costenbader K, Daikh D, et al. 2019 European League Against Rheumatism/American College of Rheumatology classification criteria for systemic lupus erythematosus. Ann Rheum Dis 2019; 78(9):1151–9.
25. van den Hoogen F, Khanna D, Fransen J, et al. 2013 classification criteria for systemic sclerosis: an American College of Rheumatology/-European League Against Rheumatism collaborative initiative. Ann Rheum Dis 2013;72(11):1747–55.
26. Shiboski CH, Shiboski SC, Seror R, et al. 2016 American College of Rheumatology/European League Against Rheumatism classification criteria for primary Sjogren's syndrome: a consensus and data-driven methodology involving three international patient cohorts. Arthritis Rheumatol 2017;69(1):35–45.
27. Wallace ZS, Naden RP, Chari S, et al. The 2019 American College of Rheumatology/European League Against Rheumatism classification criteria for IgG4-related disease. Ann Rheum Dis 2020;79(1):77–87.
28. Harris PA, Taylor R, Thielke R, et al. Research Electronic Data Capture (REDCap)–a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42(2):377–81.
29. Priemel MH, Stiel N, Zustin J, et al. Bone tumours of the clavicle: histopathological, anatomical and epidemiological analysis of 113 cases. J Bone Oncol 2019;16:100229.
30. Yamazaki Y, Satoh C, Ishikawa M, et al. Remarkable response of juvenile diffuse sclerosing osteomyelitis of mandible to pamidronate. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2007;104(1):67–71.
31. Padwa BL, Dentino K, Robson CD, et al. Pediatric chronic nonbacterial osteomyelitis of the jaw: clinical, radiographic, and histopathologic features. J Oral Maxillofac Surg 2016;74(12):2393–402.
32. Baur DA, Altay MA, Flores-Hidalgo A, et al. Chronic osteomyelitis of the mandible: diagnosis and management—an institution's experience over 7 years. J Oral Maxillofac Surg 2015;73(4): 655–65.
33. van Merkesteyn JPR, Groot RH, Bras J, et al. Diffuse sclerosing osteomyelitis of the mandible: A new concept of its etiology. Oral Surg Oral Med Oral Pathol 1990;70(4):414–9.
34. Soubrier M, Dubost JJ, Ristori JM, et al. Pamidronate in the treatment of diffuse sclerosing osteomyelitis of the mandible. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2001;92(6):637–40.
35. Chun CSY. Chronic recurrent multifocal osteomyelitis of the spine and mandible: case report and review of the literature. Pediatrics 2004;113(4):e380–4.
36. Suei Y, Taguchi A, Tanimoto K. Diffuse sclerosing osteomyelitis of the mandible: its characteristics and possible relationship to synovitis, acne, pustulosis, hyperostosis, osteitis (SAPHO) syndrome. J Oral Maxillofac Surg 1996;54(10):1194–9 discussion 1199–200.
37. Suei Y, Tanimoto K, Taguchi A, et al. Chronic recurrent multifocal osteomyelitis involving the mandible. Oral Surg Oral Med Oral Pathol 1994;78(2):156–62.

Arthritis & Rheumatology
Vol. 78, No. 3, March 2026, pp 548–557
DOI 10.1002/art.43427

Biomarkers of Lupus Nephritis Histopathology: Where Do We Stand?

Valentina Querin,[1,2] Leonardo Palazzo,[3] Agnes B. Fogo,[4] Mark Haas,[5] David R. W. Jayne,[6] Liz Lightstone,[7] Brad H. Rovin,[8] Surya V. Seshan,[9] Augusto Vaglio,[2,10] Ioannis Parodis,[3,11] and Ingeborg M. Bajema,[1] for the Lupus Nephritis Classification Steering Committee

Objective. Lupus nephritis (LN) is characterized by a variable disease course, necessitating continuous monitoring. There is an urgent need to identify noninvasive biomarkers. By reviewing and critically assessing the quality of existing studies on LN biomarkers correlating with histopathology, we here explore the challenges in promoting their use in clinical practice and identify promising candidates for future validation.

Methods. A systematic literature search was conducted, including studies on adult patients with biopsy-proven LN published between 2012 and 2024. The search focused on studies demonstrating correlations between biomarkers and histologic findings, particularly the National Institutes of Health activity and chronicity indices, the International Society of Nephrology/Renal Pathology Society histologic classes, and specific active and chronic lesions. The quality of selected articles was assessed by a multidisciplinary panel of experts using a standardized scoring system.

Results. Ninety-three articles investigating the potential utility of >100 biomarkers were selected. In quality assessment, 68% of the articles were weak or very weak (score <5, scale 2–9). From the remaining articles (adequate or robust) we identified five biomarkers with a potential for implementation in clinical practice: transforming growth factor β1, pentraxin-3, (soluble) CD163, CD11b, and interleukin-16.

Conclusion. The methodologic heterogeneity across studies and the overall moderate score of the quality of the articles represent obstacles to the implementation of new biomarkers into clinical practice. The LN working group advocates further research on selected biomarkers that demonstrated good capacity to reflect specific histopathologic features outperforming traditional tests. At present, the choice of biomarkers that perform to these standards is very limited.

INTRODUCTION

Given the complexity and variable course of systemic lupus erythematosus (SLE), in combination with potential adverse effects of therapy, continuous clinical and laboratory assessments are required to predict and evaluate organ involvement, response to treatment, and disease progression.[1] In current clinical practice, kidney biopsy, in combination with the assessment of well-known serum and urinary parameters reflecting the severity of kidney involvement in SLE, are used to monitor patients affected by lupus nephritis (LN). However, it is known that these routine clinical biomarkers do not adequately mirror the evolving lesions that we know occur in due course in the kidneys of patients with LN.[2] Many new biomarkers have been suggested to hold promise for additional clinical value above those of traditional tests.

Searching MEDLINE for "biomarkers" and "lupus nephritis" through PubMed yields nearly 2,000 results over a period of 48 years, with the most hits in recent years.[3] New biomarkers

[1]Valentina Querin, MD, Ingeborg M. Bajema, MD, PhD: Department of Pathology and Medical Biology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands; [2]Valentina Querin, MD, Augusto Vaglio, MD, PhD: Department of Biomedical, Experimental and Clinical Sciences "Mario Serio," University of Florence, Florence, Italy; [3]Leonardo Palazzo, MD, Ioannis Parodis, MD, PhD: Division of Rheumatology, Department of Medicine Solna, Karolinska Institutet and Karolinska University Hospital, Stockholm, Sweden; [4]Agnes B. Fogo, MD: Department of Pathology, Microbiology and Immunology, Vanderbilt University Medical Center, Nashville, Tennessee; [5]Mark Haas, MD, PhD: Department of Pathology and Laboratory Medicine, Cedars-Sinai Medical Center, Los Angeles, California; [6]David R. W. Jayne, MD: Department of Medicine, University of Cambridge, Cambridge, United Kingdom; [7]Liz Lightstone, MBBS, PhD: Department of Immunology and Inflammation, Faculty of Medicine, Imperial College London, London, United Kingdom; [8]Brad H. Rovin, MD: Ohio State University, Columbus; [9]Surya V. Seshan, MD: Department of Pathology and Laboratory Medicine, Weill Cornell Medicine, New York, New York; [10]Augusto Vaglio, MD, PhD: Nephrology and Dialysis Unit, Meyer Children's Hospital IRCCS, Florence, Italy; [11]Ioannis Parodis, MD, PhD: Department of Rheumatology, Faculty of Medicine and Health, Örebro University, Örebro, Sweden.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/art.43427).

Author disclosures and graphical abstract are available at https://onlinelibrary.wiley.com/doi/10.1002/art.43427.

Address correspondence via email to Valentina Querin, MD, at v.querin@umcg.nl.

Submitted for publication July 10, 2025; accepted in revised form October 6, 2025.

have been proposed to be easily measured and valuable for disease monitoring. These biomarkers aim to serve a variety of purposes, such as aiding in diagnosis, indicating disease severity and prognosis, guiding treatment decisions, and helping to identify flares. No single biomarker is likely to address all these aspects; most probably only a panel of multiple biomarkers, with complementary predictive value, could help to improve current clinical practice, and the challenge lies in determining the perfect combination of biomarkers that could serve to do so.

Certain biomarkers correlate with histologic findings, offering the potential to supplement kidney biopsy in monitoring disease activity. Although the kidney biopsy provides a wealth of information regarding the pattern of injury at a specific point in time, its use for longitudinal assessment, with repeat biopsies, is cumbersome. In contrast, a "liquid biopsy," based on biomarkers derived primarily from blood or urine, may offer a noninvasive, rapid, highly specific, and sensitive assessment of LN activity and damage.[4] Ideally, such biomarkers should faithfully recapitulate kidney biopsy findings, and we here set out a search to identify them.

Despite several LN biomarkers identified using omics approaches,[5,6] none has been integrated into clinical practice yet. Biomarkers are often tested in cross-sectional studies and the results are rarely validated in independent cohorts. Most importantly, only a few studies examine how biomarkers and histopathology change over time although it seems evident that this is what they are primarily needed for. The longitudinal correlation of a biomarker with histopathology must be established to qualify a biomarker for clinical use. Furthermore, a qualified biomarker should outperform or offer significantly added value compared with existing clinical markers used for LN management.

We conducted a systematic analysis of the recent literature on potential noninvasive biomarkers that specifically associate with histologic findings detected in kidney biopsies to gain insights into which biomarkers could eventually be helpful in representing tissue findings in LN. Furthermore, we provide an assessment of the quality of the studies and address why qualification of biomarkers is cumbersome. Lastly, we identify which biomarkers appear to be of most interest for future validation.

MATERIALS AND METHODS

Search strategy and selection criteria. We performed a systematic literature search on biomarkers in LN, following up on that by Palazzo et al of 2022,[3] but restricting our search to studies with a histologic focus. A Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram illustrating the selection process is shown in Figure 1. We selected studies in English published between January 1, 2012, and August 22, 2024, using terms including but not limited to "lupus nephritis," "systemic lupus erythematosus," "biomarker," "urine," "serum," "kidney," "tissue," "antibody," "antigen," "cytokine," "protein," "RNA," "peptide," and "receptor" (detailed in Supplementary Table S1). We included studies conducted on adult patients (aged ≥18 years) with biopsy-proven LN that highlighted a correlation between one or more biomarkers and histologic findings, with a focus on the National Institutes of Health (NIH) activity index (AI) and chronicity index (CI), the International Society of Nephrology/Renal Pathology Society (ISN/RPS) histologic classes (I–VI), and specific active or chronic lesions in the glomerular, tubulointerstitial, and vascular compartments. We excluded case reports, case series, reviews, animal studies, and studies on pediatric patients. We also subsequently conducted a manual search covering the same timeframe and inclusion criteria; this involved analyzing the references of the previously selected articles and searching the MEDLINE through PubMed for individual biomarkers already identified.

Data analysis. Considering the mixed design of many studies included in this review, the risk of bias (RoB) was assessed using the Joanna Briggs Institute (JBI) Critical Appraisal Checklist for Analytical Cross-Sectional Studies or the JBI Critical Appraisal Checklist for Cohort Studies (Supplementary Tables S2 and S3), based on the main design of the study and the methodology used to explore biomarker-histopathologic correlations. We assembled a multidisciplinary group of experts and researchers in the field of LN—one nephrologist (DRWJ), one rheumatologist (IP), one pathologist (IMB), and two junior researchers (VQ and LP)—and asked them to provide their opinion about the quality of each article, based on the grading system of Guyatt et al. This system classifies the quality of evidence in four levels—high, moderate, low and very low—according to the likelihood that future research will change the confidence in the estimated effect, namely the performance of the biomarker.[7] A score of 2 was for "very low quality," 5 was for "low quality," 7 was for "moderate quality," and 9 was for "high quality." We then calculated the average score for each article. Finally, we calculated the score for each biomarker based on the average of the quality evaluations from the various articles reviewed (one or multiple, depending on the biomarker).

RESULTS

Different categories of biomarkers reflect a broad spectrum of histopathologic lesions. From the systematic search, we extracted 74 publications and through a manual search we ultimately selected 19 additional articles. We investigated the potential utility of >100 biomarkers, dividing them into categories and describing their associations with histopathologic features (Supplementary Table S4). The biomarkers' categories included autoantibodies (n = 8); complement factors (n = 5); cytokines, chemokines, and growth factors (n = 26); cell adhesion molecules and cell surface molecules (n = 7); immune cells (n = 8); markers of kidney damage (n = 17); micro-RNA (n ≈ 12); and others (n ≈ 29). The histopathologic parameters included solitary or combined lesions from the cell level to combined

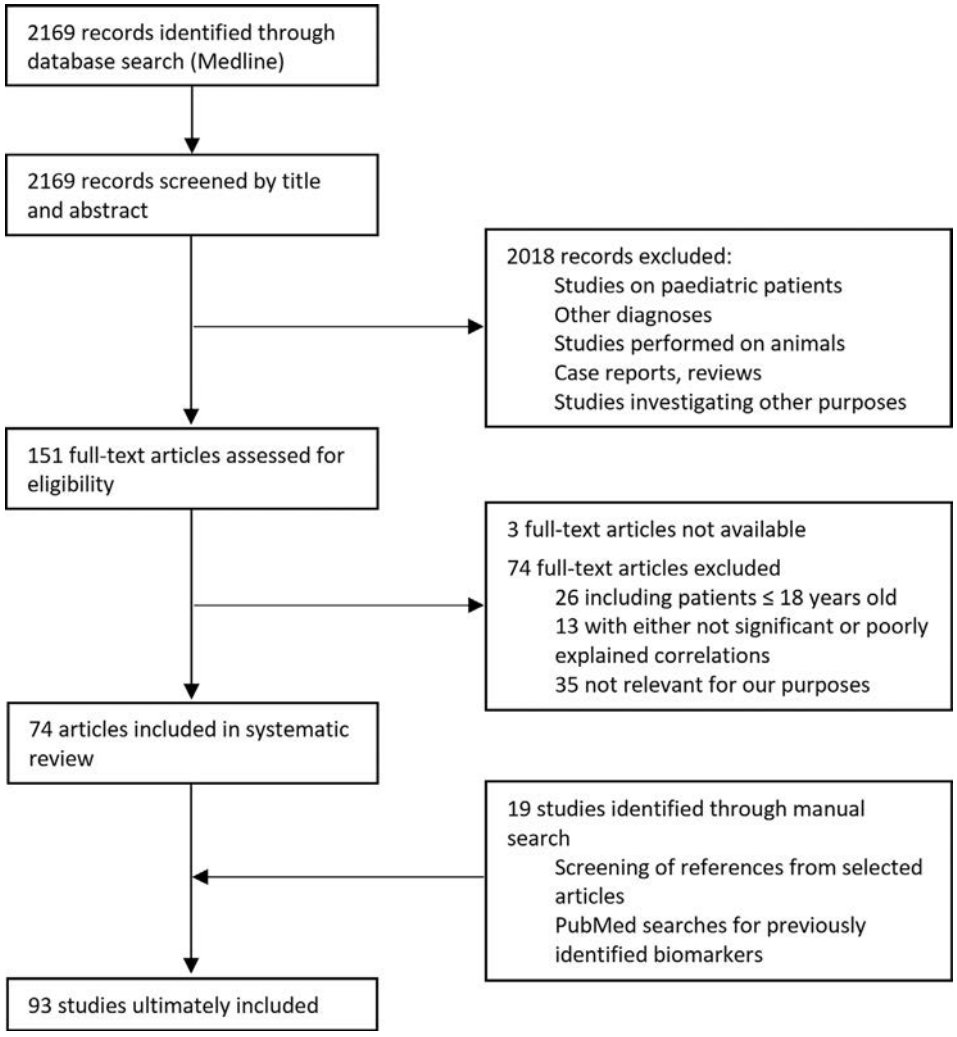


Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram of the included studies.

measurements such as the NIH indices and ISN/RPS classes. Figure 2 shows biomarkers in relation to the histopathologic lesions, as described in the 93 publications.

A low overall quality of the articles affects the reliability of analyzed biomarkers. More than 50% of the articles received low overall scores in critical appraisal, with 18% scoring <5. One article was scored above 8, whereas the remaining 42% displayed moderate scores (Supplementary Table S5). Importantly, the RoB did not necessarily align with the overall quality of the articles, for which other aspects such as the clinical relevance, clarity of results and methodology, and applicability of findings are also important.

Approximately 10% of the biomarkers were scored within the highest score range, 10% scored within the lowest range, and several biomarkers clustered around a score of 6, the threshold above which quality is considered acceptable (Figure 3). In the next section we will primarily focus on biomarkers that obtained a good overall score (≥6).

Autoantibodies mostly reflect active kidney injury, whereas immune cells have not proven useful as biomarkers. Numerous antibodies, including anti–double-stranded DNA (anti-dsDNA) and antinucleosome antibodies, have been investigated for their usefulness in assessing activity in LN.[8–12] The study designs differed considerably and correlated antibody levels or positivity with various histologic parameters. For instance, in a study of 107 patients the positivity for anti-dsDNA and antinucleosome antibodies was associated with NIH AI scores.[10] Other studies looked more generally at the ability of various autoantibodies to distinguish "proliferative" classes III/IV from "nonproliferative" and/or membranous class V LN.[13,14] In a different approach, the absence of anti-Sjogren's-syndrome-related antigen B (anti-SSB) and anti-Smith (anti-Sm) antibodies was protective against the development of severe LN.[15] Anti-neutrophil cytoplasmic autoantibodies (ANCAs) do not have a defined role in LN with contradictory results regarding their correlation with NIH AI and CI scores or with a variety of lesions.[11,16,17] Some investigators suggested that LN that was more vasculitic in nature would also be ANCA positive,[18] but that was never verified.

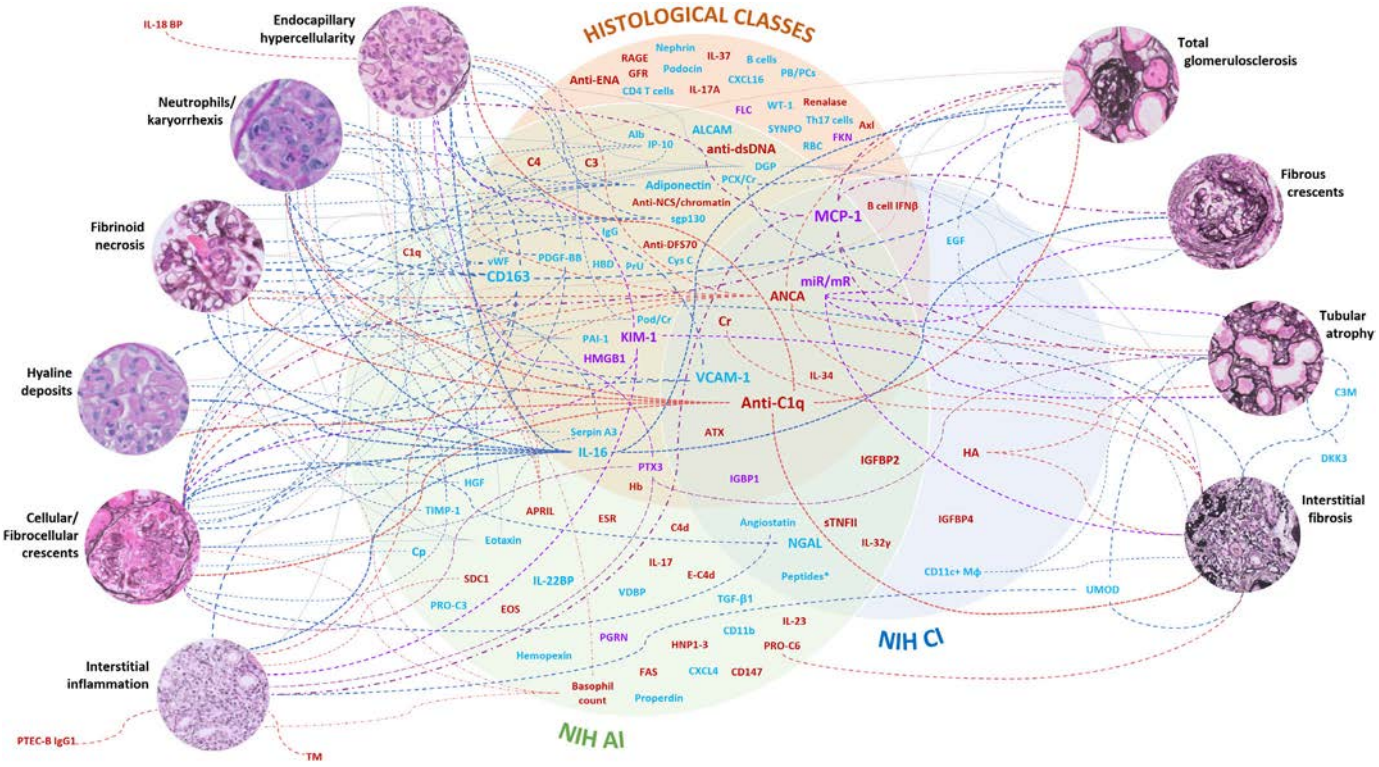


Figure 2. All biomarkers measured in serum (in red), urine (in blue), or both (in purple) that have shown a significant correlation with one or more histologic features identifiable in renal biopsy, particularly the NIH AI (green circle), the NIH CI (blue circle), the histologic classes (orange circle), and the active and chronic lesions that determine the indices (smaller outer circles with histologic pictures). The arrows indicate the correlations with the active and chronic lesions. Only biomarkers extensively described in the articles are included. The biomarkers and arrows are displayed in varying font sizes and thicknesses, respectively, based on the number of published articles on each of them. *A total of 137 peptides correlate with AI and 7 with CI. AI, activity index; Alb, albumin; ANCA, antineutrophil cytoplasmic antibody; anti-C1q, anticomplement component 1q antibody; anti-dsDNA, anti–double-stranded DNA; anti-ENA, antiextractable nuclear antigen antibody; APRIL, A proliferation-inducing ligand; BP, binding protein; C3, complement component 3; C4d, complement component 4d; CI, chronicity index; Cp, ceruloplasmin; Cr, creatinine; CXCL4, chemokine ligand 4; Cys C, cystatin C; DGP, granulocytes degranulation product; DKK3, Dickkopf 3; EOS, eosinophil count; ESR, erythrocyte sedimentation rate; FAS, fas cell surface death receptor; FKN, fractalkine; GFR, glomerular filtration rate; HA, hyaluronan; Hb, hemoglobin; HMGB1, high mobility group box 1; HNP1-3, human neutrophil peptides 1-3; IGFBP4, insulin-like growth factor binding protein 4; IL, interleukin; IP-10, interferon-γ inducible protein 10; KIM-1, kidney injury molecule 1; MCP-1, monocyte chemoattractant protein-1; miR, micro-RNA; mR, messenger RNA; Mϕ, macrophage; NIH, National Institutes of Health; PAI-1, plasminogen activator inhibitor type 1; PGRN, progranulin; Pod, podocyte; PTEC, proximal kidney tubular epithelial cell; PTX3, pentraxin 3; RAGE, receptor for advanced glycation end-products; RBC, red blood cell; SDC1, syndecan-1; sTNFRII, soluble tumor necrosis factor receptor II; TGF-β1, transforming growth factor β1; TIMP-1, tissue inhibitor of metalloproteinase-1; TM, thrombomodulin; VDBP, vitamin D binding protein; vWF, Von Willebrand factor.

Most studies investigating biomarkers based on the presence of inflammatory cells in urine and serum did not obtain a good score in the quality assessment. Of note is a study that identified that a urinary count exceeding 800 CD4 T cells/mL exhibited a remarkable sensitivity and specificity (area under the curve = 0.9969) for identifying proliferative forms of LN.[19]

Urinary interleukin-16, (soluble)CD163, transforming growth factor β1, and CD11b are promising biomarkers of LN activity. Cytokines are more commonly tested in the urine from patients with LN because of their presumed greater specificity in reflecting kidney injury or damage. Several studies investigated their correlation with histopathologic findings and study designs usually aimed to distinguish between proliferative and nonproliferative LN[20–22] or to correlate the biomarker levels with the NIH AI and CI scores.[23–28] Only a few high-quality studies also described correlations with specific histopathologic lesions.[23,29–32]

In a discovery study, urinary levels of interleukin (IL)-16, macrophage surface molecule CD163, and transforming growth factor β1 (TGF-β1) were postulated as biomarkers correlating best with the NIH AI scores.[6] Interestingly, there was abundant glomerular and interstitial IL-16 expression in proliferative forms of LN and an association between urinary levels of IL-16 and the number of IL-16$^+$ cells in the kidney. This suggests that urinary IL-16 levels are likely derived from the kidney and, in agreement with previous studies,[21,29] could potentially serve as a good biomarker of active forms of LN.[6] In a subsequent study involving

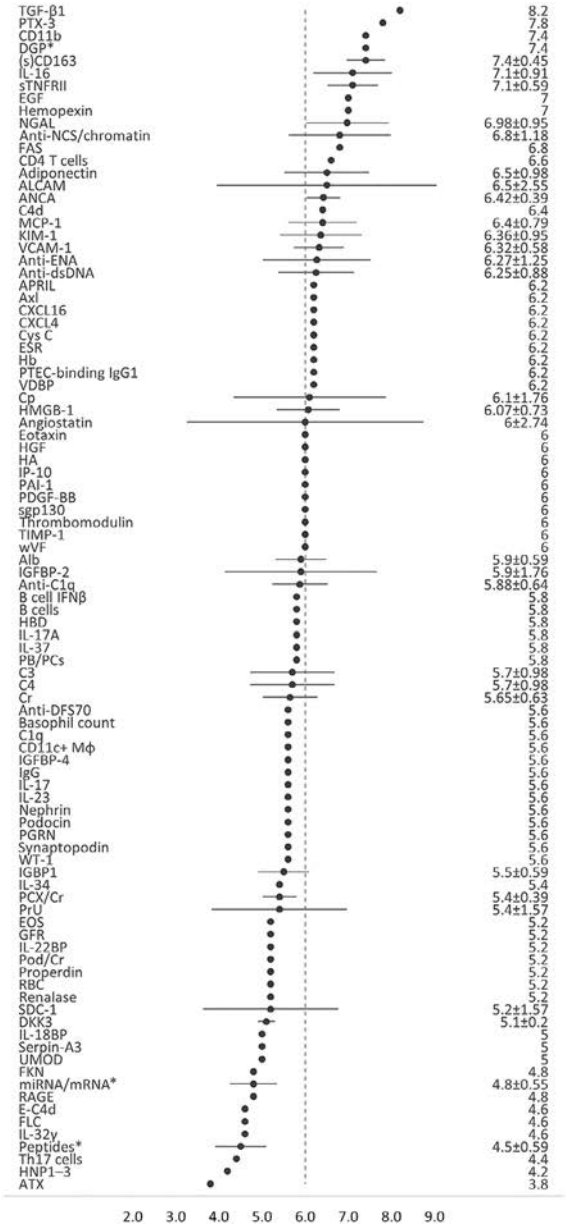


Figure 3. Forest plot showing the mean score and the 95% confidence interval for each biomarker according to the quality evaluation of the articles by the expert panel. Only biomarkers extensively described in the articles are included. *These categories encompass numerous biomarkers; peptides include both collagen and noncollagen types (details for each biomarker can be found in the corresponding articles listed in Supplementary Table S4). Alb, albumin; ALCAM, activated leukocyte cell adhesion molecule; ANCA, antineutrophil cytoplasmic antibody; anti-C1q, anticomplement component 1q antibody; anti-DFS70, antidense fine speckled 70-kDa MW antibody; anti-dsDNA, anti–double-stranded DNA; anti-ENA, antiextractable nuclear antigen antibody; anti-NCS, antinucleosome antibody; APRIL, A proliferation-inducing ligand; ATX, autotaxin; Axl, axl receptor tyrosine kinase; C3, complement component 3; C4, complement component 4; C4d, complement component 4d; CI, chronicity index; Cp, ceruloplasmin; Cr, creatinine; CXCL4, chemokine ligand 4; CXCL16, chemokine ligand 16; Cys C, cystatin C; DGP, granulocytes degranulation product; DKK3, Dickkopf 3; E-C4d, endothelial complement component 4d; EGF, epidermal growth factor; EOS, eosinophil count; ESR, erythrocyte sedimentation rate; FAS, fas cell surface death receptor; FKN, fractalkine; FLC, free light chains; GFR, glomerular filtration rate; HA, hyaluronan; Hb, hemoglobin; HBD, hemoglobin subunit delta; HGF, hepatocyte growth factor; HMGB1, high mobility group box 1; HNP1-3, human neutrophil peptides 1-3; IFNβ, interferon β; IGBP1, Ig binding protein 1; IGFBP2, insulin-like growth factor binding protein 2; IGFBP4, insulin-like growth factor binding protein 4; IL, interleukin; IP-10, interferon-γ inducible protein 10; KIM-1, kidney injury molecule 1; MCP-1, monocyte chemoattractant protein-1; miRNA, micro-RNA; mRNA, messenger RNA; NGAL, neutrophil gelatinase-associated lipocalin; PAI-1, plasminogen activator inhibitor type 1; PB, plasmablast; PC, plasma cell; PCX, podocalyxin; PDGF-BB, platelet-derived growth factor subunit B; PGRN, progranulin; Pod, podocyte; PrU, proteinuria; PTEC, proximal kidney tubular epithelial cell; PTX3, pentraxin 3; RAGE, receptor for advanced glycation end-products; RBC, red blood cell; SDC-1, syndecan-1; sgp130, soluble gp130; sTNFRII, soluble tumor necrosis factor receptor II; TGF-β1, transforming growth factor β1; TIMP-1, tissue inhibitor of metalloproteinase-1; UMOD, uromodulin; VCAM-1, vascular cell adhesion molecule 1; VDBP, vitamin D binding protein; vWF, Von Willebrand factor; WT-1, Wilms tumor 1.

225 patients, the proliferative LN signature was predominantly characterized by elevated urinary levels of CD163 and IL-16, in addition to many granulocyte degranulation products.[32] Previous high-quality studies had already reported the ability of CD163 and its soluble form (sCD163) to predict activity in LN, correlating both with the NIH AI scores and proliferative classes.[24,33,34] Specifically, an analysis of a limited cohort of repeat biopsies showed that levels of urinary CD163 could help distinguish even minimal degrees of activity (AI > 1) from quiescent biopsies (AI = 0).[33] Importantly, CD163 outperformed proteinuria in reflecting histologic activity.[32,33] In another study, urinary levels of the leukocyte surface molecule CD11b were significantly increased and related to histologic activity, and this marker was superior to CD163 in predicting proliferative forms of LN.[35] Consistently, urinary levels of CD11b were also correlated with the number of CD11b+ cells in glomeruli and, to a lesser extent, in the interstitium.[35] These studies suggest that IL-16, CD163, TGF-β1, and CD11b are some of the most promising biomarkers for inclusion in a hypothetical liquid biopsy panel.

Other potential biomarkers need further investigation to define their specific role in LN. Some cytokines, such as the soluble form of tumor necrosis factor receptor II (sTNFRII) or the monocyte chemoattractant protein-1 (MCP-1), exhibit dual correlations with both the NIH AI and CI scores.[24–27,31,36] This makes them difficult to use as biomarkers, unless they are combined with other potential candidates, such as urinary neutrophil gelatinase-associated lipocalin (NGAL), a well-known marker of more acute kidney injury that demonstrated a correlation with the NIH AI scores and, to a lesser extent, the NIH CI scores.[26,37,38] The combination of urinary MCP-1 with NGAL and kidney injury molecule 1 (KIM-1), enhanced the ability to predict active and chronic tubulointerstitial lesions in patients with LN.[30] Serum levels of sTNFRII were particularly effective in detecting chronic rather than active damage, associating with long-term estimated glomerular filtration rate deterioration.[25] Similarly, urinary levels of a molecule normally expressed by kidney tubules, epidermal growth factor, were negatively correlated with the NIH CI score and most of its components and appeared to reflect the onset of chronic damage before creatinine levels rose.[39]

Numerous studies have demonstrated the utility of adhesion molecules as biomarkers to detect activity in LN.[28,40–43] In a cohort of 96 patients, urinary levels of activated leukocyte cell adhesion molecule correlated positively with NIH AI scores and performed well in discriminating proliferative LN from membranous LN exceeding the ability of traditional markers such as C3, anti-dsDNA, and 24-h urine protein to indicate disease activity.[43] Urinary levels of vascular cell adhesion molecule 1 (VCAM-1), which plays a role in inflammatory cell trafficking, were positively correlated with NIH AI scores and active lesions.[28,40–42] In a large study involving 154 patients, the combination of urinary VCAM-1 with cystatin C and KIM-1 appeared to be more effective than VCAM-1 alone in differentiating proliferative from membranous LN, reaching a sensitivity of 94.7% and improving overall diagnostic performance.[42] In the same cohort, urinary VCAM-1 levels also demonstrated a positive correlation with the NIH CI scores[42]; however, this aspect has not been demonstrated in previous studies. Therefore, further investigation is required to confirm the specificity of VCAM-1 as a biomarker of histologic activity.

New emerging biomarkers with a precise pathophysiologic localization within the kidney parenchyma demonstrated greater capacity to delineate specific aspects of damage occurring in the context of LN. For instance, KIM-1, a transmembrane glycoprotein expressed by proximal tubular cells, was not only associated with the NIH AI scores and proliferative forms of LN[26,30,42] but also showed an interesting correlation with acute and chronic tubulointerstitial damage.[30,44,45] Another insight for the possibility of detecting whether a tubulointerstitial component of damage is present and more significant comes from the study by Yap et al; patients with moderate-to-severe tubulointerstitial inflammation exhibited a higher incidence of proximal kidney tubular epithelial cell–binding IgG1 seropositivity compared with those with no or mild inflammation, suggesting also a pathogenic role of this subclass of IgG in LN tubular injury.[46] Finally, increased levels of vitamin D binding protein in urine, which is normally undetectable because of active tubular reabsorption, demonstrated a significant correlation with the NIH AI scores, suggesting a potential role in detecting even a subclinical tubular damage.[42]

Pentraxin 3 correlates with specific active lesions in LN. Elevated serum levels of a systemic inflammation indicator, pentraxin 3 (PTX3), were associated not only with the NIH AI scores but particularly with class IV LN, endocapillary hypercellularity, and leukocyte infiltration. Its urinary levels also exhibited a significant correlation with cellular crescents and interstitial inflammation, potentially making PTX3 a reliable marker of histologic activity in LN.[47] Further studies with multiple marker panels may allow integrated assessments of which combined markers allow the best understanding of disease states.

DISCUSSION

LN is a severe manifestation of SLE in the kidney, leading to considerable patient morbidity and mortality.[48] More and more studies show that in addition to the valuable role of the histologic evaluation in establishing the diagnosis of LN, repeat kidney biopsies are useful in making decisions about the management of LN patients during long-term follow-up.[2,49] However, taking a kidney biopsy always carries a certain, although often minor, risk of complications.[50] Therefore, there is a continuous call for urinary and serum biomarkers as surrogates for the kidney biopsy. To obtain insight into the current status of biomarkers related to the histopathology of LN, we critically reviewed studies published in the last

12 years that presented one or more biomarkers with a relation to the spectrum of lesions in LN. We investigated how these biomarkers could be used in making the first steps of what a liquid biopsy could look like. Our aim was to identify biomarkers of LN histopathology with strong evidence for their applicability in clinical practice to select those that should be prioritized in future feasibility studies to facilitate their advancement toward regulatory approval and clinical usage.

Our results reveal an abundance of studies describing numerous biomarkers that correlate with many different histopathologic parameters and were conducted through a variety of methodologic approaches. Only a small number of studies were graded as being of good quality according to a vigorous quality control assessment. Low- or moderate-quality studies were usually characterized by very small cohorts from different populations, a lack of robust control groups and/or specificity about the timeframe between taking the biopsy and the sample for biomarker assessment, and an oversimplification of histopathology. From the studies graded to be of good quality, we identified a role for the urinary biomarkers (s)CD163 and IL-16 as surrogate markers for active lesions in LN,[24,29,32–34] for PTX3 to be related to endocapillary hypercellularity and cellular crescents,[47] and for urinary CD11b to be correlated to infiltrating cells in the glomerular and tubulointerstitial compartments.[35] TGF-β1 was shown to have a correlation with the total score of the NIH AI.[6] Notably, the combination of these biomarkers seems to capture a broad scope of the spectrum of active lesions in LN, and it would be tempting to investigate in a future study how much of the variation within the LN biopsy profile could be predicted by the combination of these biomarkers in a new clinical test.

Despite these promising findings, it was disappointing to realize that the heterogeneity of methodologic approaches, in combination with the overall low level of quality of most of the studies, considerably limited the interpretation of >50% of the identified articles. Importantly, almost 40% of the analyzed articles did not accurately report the timing of sample collection in relation to the kidney biopsy, affecting the reliability of the results. This might constitute room for improvement in future studies. Moreover, quality was often negatively influenced by oversimplification of the complexity of histopathologic lesions. Many studies made gross generalizations by lumping together a great number of subtle lesions to demarcate, for example, "active" versus "inactive" LN or "proliferative" versus "nonproliferative" LN. Furthermore, comparisons between studies were hampered by methodologic issues, for example with regard to different cutoffs for the NIH AI used in studies to investigate how biomarkers were able to detect various levels of disease activity. It is important to underline that many studies stopped at a point at which a distinction would have been informative in the diagnostic phase, whereas the clinical interest of new biomarkers lies more in their capacity to monitor the evolution of disease over time and guide therapy accordingly, sometimes obviating the need for a kidney biopsy. Considering the difficulty in establishing standardized cutoffs for the NIH AI and CI to evaluate biomarkers and their potential usefulness, it would be more appropriate to develop panels with biomarkers that each reflect specific histologic lesions rather than the overall NIH AI or CI score.

Of note, most of the biomarker studies focus on indicators of active inflammatory lesions; almost none investigated how to predict for the membranous component of LN. One of the challenges in the clinical management of patients with LN is to determine whether ongoing proteinuria is related to class V lesions or, for instance, to chronic lesions such as segmental glomerulosclerosis. For biomarkers being used as a "liquid biopsy," this is an important area of interest that is still understudied. Extending this argument to the development of chronic lesions in general, the focus on purely active lesions alone risks neglecting the subtle, long-term damage caused by low-grade and chronic inflammation, which can progressively impair renal function.

That many of the publications included in our review had only a mediocre quality, as assessed by our scoring system, is one of the limitations of our study together with the fact that we had to rely on the published results of studies with various methodologies. Almost all the studies that obtained a higher-quality rank established associations with the NIH AI and/or CI or focused on correlations between a biomarker and single histopathologic lesions. However, practically none of them investigated whether biomarkers driven by histopathology could be used for decisions with regard to new therapies of LN. Considering the large number of drugs in development for LN that aim to target specific pathogenetic pathways such as the complement system, the interferon system, or specific inflammatory cells, it would be obvious to look for signs in the kidney tissue that capture the presence and level of activity of these drug targets, followed by the development of a biomarker that could monitor the level of activity continuously as a tool for therapy response and for renal outcome. It seems that the premises for such studies are available, and hopefully they will be performed in the very near future.

In LN trials, multiple biomarker panels could prove useful for better stratifying patients and effectively evaluating therapy response, overcoming the limitation of traditional markers, such as proteinuria. Next to that, an increasing body of research is focusing on the use of transcriptomics to detect tissue biomarkers that could serve in the diagnostic phase to tailor treatments and, thus, improve response rates. The combined use of tissue biomarkers with liquid biopsy panels could fundamentally change LN management and prognosis.

From our experience with the articles we have reviewed and analyzed here, we recommend for future studies that they adhere to several quality control parameters, which we have listed in Table 1. Standardizing validation cohorts by unifying treatment protocols and adjusting for patients' characteristics and

Table 1. Recommendations for future research on LN biomarkers that aim to reflect kidney histopathology*

Essential components of high-quality biomarker studies	Main characteristics of essential components	Other aspects to consider and further suggestions
Discovery cohort	• Representative of the spectrum of affected patients with LN • Adequate size • Kidney biopsy data available	–
Control groups	• Highest possible similarity to the cases in terms of age, sex, and ethnicity • Selected based on the research question	–
Serum/urinary sample	• Rigorous protocols for collection, processing, storage, and analysis • Collected within 2–3 weeks of the kidney biopsy (at diagnosis or flare), preferably before immunosuppressive therapy	Clinical trials should store samples for biomarker studies
Biomarker measurement	• According to manufacturer instructions • Reference to standardized cutoff values	• Evaluate confounding factors • Consider costs and ease of detection
Outcomes	• Associations with histopathologic findings, especially with specific features of activity and/or chronicity • Prediction of clinical or histopathologic outcome over time	• Centralized biopsy evaluation would mitigate interobserver variability and could thereby facilitate more structured procedures for biomarker interpretation • Biomarkers should prove superior to traditional markers in reflecting outcomes and exhibit high sensitivity and specificity
Validation process	• Standardized cohorts spanning the heterogeneity of LN • Independent prospective longitudinal studies that overcome the limitations of cross-sectional design	Centers enrolled in biomarker studies should align treatment protocols to improve interpretation of results

* LN, lupus nephritis.

comorbidities affecting outcomes should be prioritized. Furthermore, future validation processes require prospective longitudinal studies that can overcome the limitations of cross-sectional evaluations and investigate biomarker changes over time in relation to kidney histopathology.

In conclusion, there is an urgent need for closer monitoring of LN through the liquid biopsy that could prevent or predict overt reactivation of disease or clinical flares. Only by performing high-quality studies, including biomarker evaluations in clinical trials, and further investigating the clinical utility of multiple biomarkers that would reflect the complex and alternating histopathology of LN in time will we be able to add an extra dimension to the care of patients with LN.

ACKNOWLEDGMENTS

We thank Julius Lindblom for his work on the original search at Karolinska Institute and all librarians involved. Gemini was used to improve the readability and language of the manuscript. After using this tool, the authors reviewed and edited the text as needed and take full responsibility for the content of the publication.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr Querin confirms that all authors have provided the final approval of the version to be published and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. Dörner T, Furie R. Novel paradigms in systemic lupus erythematosus. Lancet 2019;393(10188):2344–2358.
2. Malvar A, Pirruccio P, Alberton V, et al. Histologic versus clinical remission in proliferative lupus nephritis. Nephrol Dial Transplant 2017; 32(8):1338–1344.
3. Palazzo L, Lindblom J, Mohan C, et al. Current insights on biomarkers in lupus nephritis: a systematic review of the literature. J Clin Med 2022;11(19):5759.
4. Świecka M, Maślińska M. Liquid biopsy and its emerging role in rheumatology. Crit Rev Immunol 2022;42(1):13–19.
5. Stanley S, Vanarsa K, Soliman S, et al. Comprehensive aptamer-based screening identifies a spectrum of urinary biomarkers of lupus nephritis across ethnicities. Nat Commun 2020;11(1):2197.
6. Fava A, Rao DA, Mohan C, et al; Accelerating Medicines Partnership in Rheumatoid Arthritis and Systemic Lupus Erythematosus Network. Urine proteomics and renal single-cell transcriptomics implicate interleukin-16 in lupus nephritis. Arthritis Rheumatol 2022;74(5): 829–839.
7. Guyatt GH, Oxman AD, Vist GE, et al; GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008;336(7650):924–926.
8. Chen Z, Wang GS, Wang GH, et al. Anti-C1q antibody is a valuable biological marker for prediction of renal pathological characteristics in lupus nephritis. Clin Rheumatol 2012;31(9):1323–1329.

9. Tan Y, Song D, Wu LH, et al. Serum levels and renal deposition of C1q complement component and its antibodies reflect disease activity of lupus nephritis. BMC Nephrol 2013;14(1):63.
10. Moroni G, Quaglini S, Radice A, et al. The value of a panel of autoantibodies for predicting the activity of lupus nephritis at time of renal biopsy. J Immunol Res 2015;2015:106904.
11. Li C, Zhou ML, Liang DD, et al. Treatment and clinicopathological characteristics of lupus nephritis with anti-neutrophil cytoplasmic antibody positivity: a case-control study. BMJ Open 2017;7(7):e015668.
12. Chen D, Zhao L, Dai Y, et al. Anti-DFS70 antibodies are associated with proliferative lupus nephritis and renal pathological activity. Front Immunol 2022;13:810639.
13. Kwon OC, Park JH, Park HC, et al. Non-histologic factors discriminating proliferative lupus nephritis from membranous lupus nephritis. Arthritis Res Ther 2020;22(1):138.
14. Renaudineau Y, Chauveau D, Faguer S, et al. Urinary soluble CD163 is useful as "liquid biopsy" marker in lupus nephritis at both diagnosis and follow-up to predict impending flares. J Transl Autoimmun 2024; 9:100244.
15. Kianmehr N, Khoshmirsafa M, Shekarabi M, et al. High frequency of concurrent anti-C1q and anti-dsDNA but not anti-C3b antibodies in patients with Lupus Nephritis. J Immunoassay Immunochem 2021; 42(4):406–423.
16. Wang Y, Huang X, Cai J, et al. Clinicopathologic characteristics and outcomes of lupus nephritis with antineutrophil cytoplasmic antibody: a retrospective study. Medicine (Baltimore) 2016;95(4):e2580.
17. Wang S, Shang J, Xiao J, et al. Clinicopathologic characteristics and outcomes of lupus nephritis with positive antineutrophil cytoplasmic antibody. Ren Fail 2020;42(1):244–254.
18. Turner-Stokes T, Wilson HR, Morreale M, et al. Positive antineutrophil cytoplasmic antibody serology in patients with lupus nephritis is associated with distinct histopathologic features on renal biopsy. Kidney Int 2017;92(5):1223–1231.
19. Enghard P, Rieder C, Kopetschke K, et al. Urinary CD4 T cells identify SLE patients with proliferative lupus nephritis and can be used to monitor treatment response. Ann Rheum Dis 2014;73(1):277–283.
20. Singh RG, Usha, Rathore SS, et al. Urinary MCP-1 as diagnostic and prognostic marker in patients with lupus nephritis flare. Lupus 2012; 21(11):1214–1218.
21. Häyry A, Faustini F, Zickert A, et al. Interleukin (IL) 16: a candidate urinary biomarker for proliferative lupus nephritis. Lupus Sci Med 2022; 9(1):e000744.
22. Rahmé Z, Franco C, Cruciani C, et al. Characterization of serum cytokine profiles of patients with active lupus nephritis. Int J Mol Sci 2023; 24(19):14883.
23. Treamtrakanpon W, Tantivitayakul P, Benjachat T, et al. APRIL, a proliferation-inducing ligand, as a potential marker of lupus nephritis. Arthritis Res Ther 2012;14(6):R252.
24. Endo N, Tsuboi N, Furuhashi K, et al. Urinary soluble CD163 level reflects glomerular inflammation in human lupus nephritis. Nephrol Dial Transplant 2016;31(12):2023–2033.
25. Parodis I, Ding H, Zickert A, et al. Serum soluble tumour necrosis factor receptor-2 (sTNFR2) as a biomarker of kidney tissue damage and long-term renal outcome in lupus nephritis. Scand J Rheumatol 2017; 46(4):263–272.
26. Gulati G, Bennett MR, Abulaban K, et al. Prospective validation of a novel renal activity index of lupus nephritis. Lupus 2017;26(9): 927–936.
27. Taha HA, Abdallah NH, Salem MN, et al. Urinary and tissue monocyte chemoattractant protein1 (MCP1) in lupus nephritis patients. Egypt Rheumatol 2017;39(3):145–150.
28. Mok CC, Soliman S, Ho LY, et al. Urinary angiostatin, CXCL4 and VCAM-1 as biomarkers of lupus nephritis. Arthritis Res Ther 2018; 20(1):6.
29. Landolt-Marticorena C, Prokopec SD, Morrison S, et al. A discrete cluster of urinary biomarkers discriminates between active systemic lupus erythematosus patients with and without glomerulonephritis. Arthritis Res Ther 2016;18(1):218.
30. Ding Y, Nie LM, Pang Y, et al. Composite urinary biomarkers to predict pathological tubulointerstitial lesions in lupus nephritis. Lupus 2018;27(11):1778–1789.
31. Pérez-Arias AA, Méndez-Pérez RA, Cruz C, et al. The first-year course of urine MCP-1 and its association with response to treatment and long-term kidney prognosis in lupus nephritis. Clin Rheumatol 2023; 42(1):83–92.
32. Fava A, Buyon J, Magder L, et al; Accelerating Medicines Partnership in RA/SLE network. Urine proteomic signatures of histological class, activity, chronicity, and treatment response in lupus nephritis. JCI Insight 2024;9(2):e172569.
33. Mejia-Vilet JM, Zhang XL, Cruz C, et al. Urinary soluble CD163: a novel noninvasive biomarker of activity for lupus nephritis. J Am Soc Nephrol 2020;31(6):1335–1347.
34. Zhang T, Li H, Vanarsa K, et al. Association of urine sCD163 with proliferative lupus nephritis, fibrinoid necrosis, cellular crescents and intrarenal M2 macrophages. Front Immunol 2020;11:671.
35. Kitagawa A, Tsuboi N, Yokoe Y, et al. Urinary levels of the leukocyte surface molecule CD11b associate with glomerular inflammation in lupus nephritis. Kidney Int 2019;95(3):680–692.
36. Wu T, Ding H, Han J, et al. Antibody-array-based proteomic screening of serum markers in systemic lupus erythematosus: a discovery study. J Proteome Res 2016;15(7):2102–2114.
37. Torres-Salido MT, Cortés-Hernández J, Vidal X, et al. Neutrophil gelatinase-associated lipocalin as a biomarker for lupus nephritis. Nephrol Dial Transplant 2014;29(9):1740–1749.
38. Susianti H, Iriane VM, Dharmanata S, et al. Analysis of urinary TGF-β1, MCP-1, NGAL, and IL-17 as biomarkers for lupus nephritis. Pathophysiology 2015;22(1):65–71.
39. Mejia-Vilet JM, Shapiro JP, Zhang XL, et al. Association between urinary epidermal growth factor and renal prognosis in lupus nephritis. Arthritis Rheumatol 2021;73(2):244–254.
40. Singh S, Wu T, Xie C, et al. Urine VCAM-1 as a marker of renal pathology activity index in lupus nephritis. Arthritis Res Ther 2012;14(4): R164.
41. Soliman S, Mohamed FA, Ismail FM, et al. Urine angiostatin and VCAM-1 surpass conventional metrics in predicting elevated renal pathology activity indices in lupus nephritis. Int J Rheum Dis 2017; 20(11):1714–1727.
42. Liu L, Wang R, Ding H, et al. The utility of urinary biomarker panel in predicting renal pathology and treatment response in Chinese lupus nephritis patients. PLoS One 2020;15(10):e0240942.
43. Ding H, Lin C, Cai J, et al. Urinary activated leukocyte cell adhesion molecule as a novel biomarker of lupus nephritis histology. Arthritis Res Ther 2020;22(1):122.
44. Nozaki Y, Kinoshita K, Yano T, et al. Estimation of kidney injury molecule-1 (Kim-1) in patients with lupus nephritis. Lupus 2014; 23(8):769–777.
45. Carbayo J, Verdalles Ú, Díaz-Crespo F, et al. Tubular biomarkers and histology in lupus nephritis. Int J Rheum Dis 2024;27(6):e15210.
46. Yap DYH, Yung S, Zhang Q, et al. Serum level of proximal renal tubular epithelial cell-binding immunoglobulin G in patients with lupus nephritis. Lupus 2016;25(1):46–53.
47. Pang Y, Tan Y, Li Y, et al. Pentraxin 3 is closely associated with tubulointerstitial injury in lupus nephritis: a large multicenter cross-sectional study. Medicine (Baltimore) 2016;95(3):e2520.

48. Mok CC, Kwok RCL, Yip PSF. Effect of renal disease on the standardized mortality ratio and life expectancy of patients with systemic lupus erythematosus. Arthritis Rheum 2013;65(8):2154–2160.
49. Pagni F, Galimberti S, Goffredo P, et al. The value of repeat biopsy in the management of lupus nephritis: an international multicentre study in a large cohort of patients. Nephrol Dial Transplant 2013;28(12): 3014–3023.
50. Poggio ED, McClelland RL, Blank KN, et al; Kidney Precision Medicine Project. Systematic review and meta-analysis of native kidney biopsy complications. Clin J Am Soc Nephrol 2020;15(11):1595–1602.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis & Rheumatology
Vol. 78, No. 3, March 2026, pp 558–565
DOI 10.1002/art.43316

REVIEW

Mechanisms of B Cell Tolerance in Health and Autoimmunity

Sivasankaran Munusamy Ponnan[1] and Shaun W. Jackson[2]

Case presentation: A 27-year-old man with end-stage renal disease secondary to X-linked Alport syndrome received a deceased donor kidney transplant after three years on hemodialysis. His posttransplant course was initially unremarkable, with stable creatinine levels around 1.2 mg/dL. However, 10 months post transplantation, he presented with acute onset of hematuria, proteinuria (3.5 g/day), and rapidly deteriorating allograft function with creatinine levels rising to 4.6 mg/dL over just one week. Urinalysis revealed dysmorphic red blood cells and red blood cell casts. Allograft biopsy demonstrated linear IgG deposition along the glomerular basement membrane (GBM) with crescent formation in 60% of glomeruli, together with high titer circulating anti-GBM antibodies, confirming the diagnosis of de novo anti-GBM disease in a kidney allograft. The consulting rheumatology team was puzzled by the immunopathology of the case.

Introduction

To understand the underlying immunology, we will comprehensively review the mechanisms of B cell tolerance. Specifically, we will examine the layered and redundant mechanisms that the immune system employs to generate a diverse B cell repertoire able to respond to distinct pathogens while maintaining immune tolerance. Using this background, we will return to the specifics of this unique case to consider how it informs the understanding of autoimmune pathogenesis.

Fundamentals of B cell development and tolerance

B cells are critical parts of the adaptive immune system, providing for humoral immunity and protection against invasive pathogens. However, the capacity of B cells to generate a vast repertoire of antigen receptors through random variable, diversity, and joining (VDJ) recombination inherently carries the risk of producing B cells that recognize self-antigens. For this reason, B cell development and selection are carefully coordinated processes that ensure the formation of a diverse yet self-tolerant antibody repertoire.[1] When these tolerance mechanisms fail, as seen in diseases such as systemic lupus erythematosus (SLE) and rheumatoid arthritis, autoreactive B cells can escape control. This leads to the production of harmful autoantibodies, which cause inflammation and tissue damage.[2] Therefore, understanding the molecular and cellular pathways that govern B cell tolerance not only sheds light on the origins of autoimmunity but also informs therapeutic strategies aimed at restoring immune balance.

B cells develop from bone marrow (BM) hematopoietic stem cells via the tightly regulated progress through specific developmental stages—from pro–B cells to pre–B cells to immature B cells. Each pro–B cell undergoes rearrangement of the VDJ gene segments of the Ig heavy-chain locus. Successful heavy-chain rearrangement results in expression of the pre–B cell receptor (pre-BCR), which first promotes pro–B cell proliferation, followed by cell cycle exit and Ig light-chain gene recombination.[3] This coordinated developmental program ultimately leads to the expression of a mature BCR as surface IgM.

Practicing rheumatologists regularly prescribe highly targeted monoclonal antibodies and naturally think of these therapeutic agents as exquisitely specific. However, it is less widely recognized that antibodies derived from the preimmune B cell repertoire exhibit remarkable polyreactivity.[4] This intrinsic polyreactivity serves an important evolutionary purpose. Because the universe of potential foreign antigens is near infinite, a repertoire composed exclusively of highly specific monoclonal antibodies would render the frequency of naïve B cells capable of recognizing novel pathogens too low to mount an effective humoral

The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the NIH.

Supported by the National Institutes of Health (grants R01-AR-073938, R01-AR-075813, and R01-DK-136980 to Dr Jackson). Additional funding support provided the Lupus Research Alliance Lupus Mechanisms and Targets Award, Global Team Science Award, and Targeted Research Program on Engineered Cell Therapies for Lupus (to Dr Jackson).

[1]Sivasankaran Munusamy Ponnan, PhD: Seattle Children's Research Institute, Seattle, Washington; [2]Shaun W. Jackson, MBChB, MD: Seattle Children's Research Institute and University of Washington School of Medicine, Seattle.

Author disclosures and graphical abstract are available online at https://onlinelibrary.wiley.com/doi/10.1002/art.43316.

Address correspondence via email to Shaun W. Jackson, MBChB, MD, at shaun.jackson@seattlechildrens.org.

Submitted for publication May 21, 2025; accepted in revised form July 1, 2025.

response.[5] Nevertheless, this intrinsic polyreactivity, coupled with the propensity for random VDJ recombination to generate self-reactive B cells, increases the potential for autoimmunity. To address this challenge, both central and peripheral tolerance mechanisms have evolved to eliminate or functionally silence autoreactive B cells, thereby preventing autoimmunity (Figure 1).

Central tolerance

Central tolerance refers to processes that eliminate or modify autoreactive B cells in the BM before they can enter the periphery and encounter their cognate self-antigens. During B cell lymphopoiesis, immature B cells expressing surface IgM encounter self-antigens presented by BM stromal cells, which express a wide range of intracellular and cell surface proteins. The strength and duration of BCR engagement with these self-antigens determines the cellular fate through several distinct mechanisms, including the following:

1. Clonal deletion describes the process by which immature B cells that strongly recognize self-antigens in the BM undergo programmed cell death (apoptosis), effectively eliminating potentially autoreactive B cell clones before they can mature and enter the circulation. This critical tolerance checkpoint was first identified using transgenic mouse models in which B cells express a single self-reactive BCR in the presence of their cognate self-antigen. In this context, strong BCR cross-linking induces developmental arrest at the BM stage.[6–8] Although initially identified using animal models, subsequent studies revealed that ~55% to 75% of developing BM B cells in humans exhibit reactivity to nuclear or cytoplasmic antigens. These high-affinity self-reactive B cells are efficiently purged from the developing repertoire,[4] underscoring the critical importance of clonal deletion to human B cell tolerance.
2. Receptor editing is a crucial salvage pathway for self-reactive immature B cells that encounter self-antigens but avoid immediate deletion. Rather than undergoing apoptosis, these cells reactivate recombination-activating gene (RAG1/2) expression, resulting in further Vκ–Jκ recombination in an attempt to replace their existing κ light chains through secondary rearrangements at the Igκ locus. If κ locus rearrangements fail to eliminate self-reactivity, the B cell may then proceed to rearrange the λ light-chain locus, resulting in multiple possibilities to

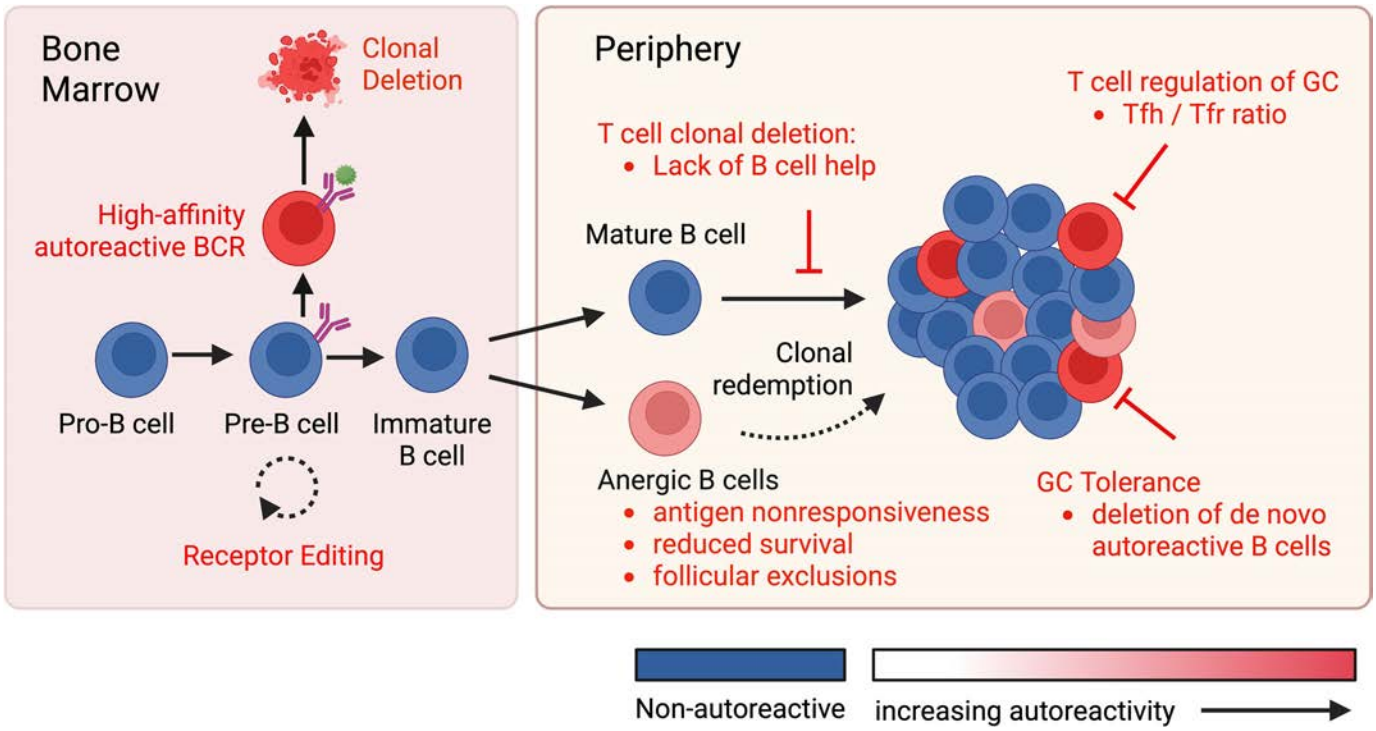


Figure 1. Central and peripheral mechanisms of B cell tolerance. This figure illustrates the layered checkpoints that maintain B cell tolerance. Central tolerance in the bone marrow (left panel) involves clonal deletion of high-affinity autoreactive B cells and receptor editing of self-reactive cells to alter B cell receptor binding. During periphery tolerance (right panel), mature B cells that are self-reactive become anergic, characterized by antigen unresponsiveness, reduced survival, and follicular exclusion. GC tolerance involves the deletion of de novo autoreactive B cells generated by somatic hypermutation. T cells also regulate B cell tolerance via central deletion of potential B cell helpers and by controlling GC responses, influenced by the balance of Tfh and Tfr cells. The concept of clonal redemption highlights how anergic B cells can be rescued to participate in immune responses against foreign antigens. A color gradient from blue (nonautoreactive) to red (increasing autoreactivity) indicates the spectrum of self-reactivity within the B cell repertoire. BCR, B cell receptor; GC, germinal center; Tfh, T follicular helper; Tfr, T follicular regulatory.

rescue the autoreactive B cells.[9] Estimates that 25% to 50% of peripheral B cells have undergone at least one round of receptor editing highlight its critical contribution to shaping a self-tolerant BCR repertoire while preserving immunologic diversity.[10–13]

Peripheral tolerance

Although central tolerance mechanisms eliminate most highly autoreactive B cells, this process is incomplete, and several key mechanisms operate in the periphery to ensure that B cell tolerance is maintained:

1. Anergy: Anergy is defined as a state of functional unresponsiveness in response to encountering cognate antigen. This tolerogenic state is typically induced when immature or transitional B cells encounter soluble self-antigens, a scenario that triggers insufficient BCR signaling to induce apoptosis but enough to initiate inhibitory pathways.[14] Anergic B cells exhibit several hallmark features, including reduced surface expression of IgM (while maintaining normal IgD levels), increased expression of inhibitory receptors such as CD22, and impaired responsiveness to antigen binding. This BCR signaling defect manifests via chronically elevated basal intracellular calcium levels paired with paradoxically reduced calcium mobilization after antigen stimulation.[15] Importantly, the presence or absence of B cell anergy does not function as a binary process but rather exists on a continuum within the mature B cell repertoire based on relative self-reactivity.[16]
2. Follicular exclusion: In addition to functional unresponsiveness, self-reactive B cells exhibit reduced ability to access B cell follicles in secondary lymphoid organs when competing with nonautoreactive B cells for limited microenvironmental niches.[17,18] After exiting the BM, preimmune B cells migrate to the spleen as immature transitional 1 B cells, which subsequently upregulate CXCR5, the receptor for the follicular chemokine CXCL13, thereby facilitating migration into primary follicles.[19] Within a diverse preimmune repertoire, anergic B cells down-regulate CXCR5, resulting in retention of self-reactive B cells at the T cell–B cell border.[20] This spatial segregation strategically positions autoreactive B cells away from survival niches, T cell help, and germinal center (GC) reactions, effectively limiting their potential to contribute to autoimmune responses.
3. Reduced survival of self-reactive B cells: In parallel with follicular exclusion, anergic autoreactive B cells developing in competition with nonautoreactive counterparts have short lifespans of only two to three days.[17,18,21] The cytokine BAFF (also known as B lymphocyte stimulator) is a critical survival factor necessary for B cell maturation beyond the transitional stage.[22,23] Both mouse models and human studies have linked increased circulating BAFF levels with the development of humoral autoimmunity, in particular SLE.[24–28] These data suggest that BAFF might breach B cell tolerance checkpoints by creating a permissive environment for autoreactive B cell survival. However, murine modeling has revealed that BAFF overexpression neither disrupts central B cell tolerance nor enables high-affinity self-reactive B cells to survive and enter follicles in competitive settings. Instead, increased BAFF level subtly modulates transitional B cell selection stringency by enhancing the survival of lower-affinity self-reactive B cells.[29–31]
4. Clonal ignorance: Although not strictly a B cell–intrinsic tolerance mechanism, clonal ignorance refers to the process in which autoreactive B cells never encounter their corresponding self-antigen. This occurs when self-antigens are present at low levels or remain sequestered within cellular structures and physiologically restricted zones.[32,33] Under these circumstances, the development, lifespan, and activation potential of these B cells remains unchanged.

GC B cell tolerance

GCs are transient structures within secondary lymphoid organs where antigen-activated B cells undergo rapid proliferation and somatic hypermutation (SHM). SHM introduces random mutations into the Ig variable region genes, diversifying the BCR repertoire and fueling the iterative selection of higher affinity B cells via a process termed affinity maturation. Ultimately, B cells with high affinity for the specific activating antigen differentiate into memory B cells and plasma cells, crucial components of humoral immunity.[34] However, the random nature of SHM inevitably carries the risk of generating B cells with de novo self-reactivity within the GC. Therefore, multiple checkpoints have evolved to maintain self-tolerance and prevent the emergence of pathogenic autoantibodies from the GC reaction. Autoreactive B cells within GCs can arise from two primary pathways: (1) recruitment of pre-existing self-reactive B cells that evaded central tolerance mechanisms in the BM and (2) generation of de novo autoreactivity via SHM within the GC itself.

Although central tolerance significantly depletes the pool of self-reactive B cells emigrating from the BM, self-reactive B cells persist in the periphery in an anergic (functionally unresponsive) state. These anergic B cells could potentially become activated and enter a GC response if they successfully acquire T cell help. For this reason, a key peripheral tolerance checkpoint limits the recruitment of self-reactive B cells into GCs during initial cognate interactions between T helper cells and B cells. Upon recognizing peptides presented on B cell major histocompatibility complex

class II, T helper cells up-regulate both CD40L, delivering CD40-activating signals to the B cell, and Fas ligand, which engages the death receptor Fas. Because of inherently weak BCR-derived survival signals, anergic B cells are particularly vulnerable to Fas-mediated death upon receiving T cell help.[35–37] This mechanism eliminates escaped autoreactive B cells, preventing their GC participation and subsequent autoantibody production.

Additional layers of tolerance operate directly within the GC to manage de novo self-reactivity from SHM. Uncovering these intra-GC mechanisms has proven technically challenging due to rapid cellular turnover and mutation rates. Nevertheless, seminal studies demonstrated that delivery of high doses of soluble antigen (mimicking ubiquitous self-antigen) triggers widespread apoptosis of established GC B cells,[38–40] a phenomenon analogous to clonal deletion during central tolerance. Moreover, SHM does not only select for increased foreign antigen affinity. Ig receptors also rapidly accrue mutations that decrease affinity for self-ligands.[41] Despite these safeguards, GC tolerance is not infallible. GCs are likely a crucial source of autoantibody-producing cells in autoimmune diseases, as evidenced by extensive SHM in autoreactive B cell clones and the characteristic expansion of spontaneous GCs in both murine lupus models and patients with SLE.[42–44]

Regulation of B cell tolerance by T cells

Although developmental tolerance is primarily governed by BCR signals, T lymphocytes act as critical extrinsic gatekeepers of B cell tolerance. Most self-reactive T cells are eliminated by effective central tolerance in the thymus. This limits activation of autoreactive B cells that have bypassed central tolerance checkpoints, particularly those specific to T-dependent antigens, as the necessary cognate T helper cell support for their activation is largely absent.

The K/BxN mouse arthritis model elegantly demonstrates this T cell dependency. Autoreactive B cells targeting the ubiquitous self-antigen glucose-6-phosphate isomerase (GPI) only trigger disease when cognate $CD4^+$ T cells expressing the KRN T cell receptor (specific for GPI peptides) are present.[45,46] This highlights how availability of T cell help can be the decisive factor in breaching B cell tolerance, even when autoreactive B cells have evaded central tolerance mechanisms. The clinical significance of this linked tolerance in humans is underscored by the characteristic autoantibody profiles in autoimmune polyglandular syndrome type 1 (also known as autoimmune polyendocrinopathy–candidiasis–ectodermal dystrophy), a disorder resulting from *AIRE* gene mutations that impair *AIRE*-dependent deletion of self-reactive T cells in the thymus.[47,48]

Beyond deletion of autoreactive T helper cells, Treg cells actively promote B cell tolerance by both modulating the preimmune repertoire and limiting autoreactive B cell activation. Compelling evidence comes from studies of patients with immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome with mutations in the *FOXP3* gene, which encodes a transcription factor essential for Treg cell function. In addition to developing severe systemic autoimmunity, these patients exhibit a specific defect in peripheral B cell tolerance checkpoints.[49] Moreover, a specialized subset of Treg cells, termed T follicular regulatory (Tfr) cells, is crucial for controlling GC responses and preventing autoreactive B cell activation.[50] Tfr cells exhibit suppressive activity in vitro and limit T follicular helper (Tfh) cell and GC B cell numbers in vivo.[51,52] Recent studies indicate that Tfr cells maintain GC quality by selectively suppressing non–antigen-specific B cells, potentially including autoreactive B cells. This represents an elegant regulatory mechanism whereby the Tfh differentiation pathway is co-opted by Treg cells to control GC responses, thus maintaining B cell tolerance while allowing productive immune responses against foreign antigens.

B cell tolerance on a spectrum: clonal redemption and biologic rationale for anergic B cell maintenance

Each of the previous sections describes the overlapping mechanisms used to maintain B cell tolerance. However, the purpose of this redundancy is not clear if one considers that robust B cell tolerance might be readily achieved via clonal deletion of all autoreactive B cells, not only those with the highest self-affinity. Instead, the immune system employs a nuanced approach to B cell tolerance in which low-affinity autoreactive B cells can exit the BM and persist within the preimmune repertoire under the control of peripheral tolerance mechanisms. This observation raises a biologic paradox—why is B cell tolerance incomplete despite the risk of humoral autoimmunity?

This apparent contradiction is best understood by considering the consequences of overly strict B cell tolerance. Because many pathogens express surface epitopes that mimic self-antigens, a repertoire entirely devoid of self-reactivity would leave "gaps" in immune coverage rendering individuals vulnerable to invasive pathogens. For this reason, despite the potential that autoreactive B cells might escape inhibitory control, the immune system elects to maintain cells with low self-antigen affinity in a state of functional anergy.

Of course, for anergic B cells to be useful in the setting of infection requires the state of anergy to be reversible. Multiple studies in both animal models and human studies have indicated that this does in fact occur. Clonal redemption refers to the process by which anergic self-reactive B cells can be "redeemed" to productively contribute to protective immune responses against pathogens.[53] Using an elegant approach, Burnett and colleagues demonstrated that antibodies derived from anergic B cells acquire mutations that can distinguish between nearly identical foreign and self-antigens.[41] To do this, they used mice

engineered to express hen egg lysozyme (HEL) ubiquitously, which induces anergy in anti-HEL B cells. These mice were immunized with duck egg lysozyme (DEL), a foreign protein structurally similar to HEL but differing by four amino acids at the interaction interface. Despite their initial self-reactivity, HEL-binding B cells entered GCs and underwent SHM characterized by a distinct two-phase selection: first, rapid selection for mutations decreasing affinity for the self-antigen (HEL), followed by slower selection for epistatic mutations that increased affinity for the foreign antigen (DEL). Thus, previously anergic autoreactive B cells can be recruited into protective humoral immune responses by exploiting subtle structural differences to evolve high specificity toward foreign antigens.

Pathways to breaks in B cell tolerance in human autoimmunity

Although the multilayered system of B cell tolerance is remarkably effective, failure can occur, leading to the generation of pathogenic autoantibodies and the development of autoimmune disease. Tolerance can be breached through various mechanisms, often involving a combination of genetic predisposition, environmental triggers, and stochastic events. One mechanism linked to the breakdown of self-tolerance is genetic risk factors that subtly alter B cell selection and activation thresholds. For example, genome-wide association studies have uncovered genetic polymorphisms impacting key B cell pathways, including BCR signaling components (*PTPN22*, *BANK1*, *BLK*, and *LYN*), endosomal toll-like receptors (*TLR7*/*TLR9*) crucial for autoantibody development in SLE, and regulators of circulating BAFF levels. Although these genetic mechanisms have been extensively reviewed elsewhere,[2,54,55] we will focus on a critical, yet underexplored, contributor to autoimmunity: the role of neoantigens (Figure 2).

The case vignette discussed in this article represents a striking co-occurrence of two rare human diseases. Alport syndrome is a genetic disorder characterized by slowly progressive kidney dysfunction and hearing loss, with a population prevalence of approximately[56] 1 in 5,000 to 10,000. In contrast, anti-GBM disease in an autoimmune syndrome resulting in rapidly progressive glomerulonephritis with an annual incidence of less than two per million.[57] The development of two rare, and seemingly unrelated, kidney diseases in one patient raises interesting questions pertinent to the pathogenesis of humoral autoimmunity.

As we have discussed, B cell tolerance is established early during development and maintained by layered peripheral tolerance mechanisms. However, when the immune system encounters antigens that were previously hidden or absent or are modified versions of self-proteins, autoimmunity can result because appropriate tolerance has not been established. One dramatic example relates to the clinical case described previously. Alport syndrome is a genetic disorder caused by pathogenic mutations in genes encoding subunits of the type IV collagen protein family, most commonly *COL4A5* in the X-linked form.[58] These mutations lead to the inability to form the normal alpha-3-4-5 type IV collagen network, a key component of basement membranes in tissues such as the kidney, eye, and cochlea. Because the normal collagen IV network is absent, the immune system remains naïve to specific collagen subunits. Consequently, when a kidney allograft expressing the full complement of type IV collagen is transplanted into a patient with Alport syndrome, the recipient's immune system may recognize these previously unseen collagen subunits as foreign. This can trigger an alloimmune response, leading to the development of de novo anti-GBM disease targeting the transplanted kidney. Interestingly, standard anti-GBM antibody testing in Alport post-transplant nephritis may yield false-negative results because commercial assays are optimized to detect antibodies against the NC1 domain of alpha-3(IV) (Goodpasture antigen) rather than the alpha-5(IV) chain mutated in X-linked Alport disease, highlighting the exquisite specificity of B cell tolerance.

Beyond this unique situation of organ transplantation for Alport syndrome, the concept that neoantigens can initiate breaks in immune tolerance applies to various autoimmune diseases. For example, posttranslational modifications of self-proteins can create novel epitopes recognized as foreign by the immune system. A classic example is the citrullination of proteins in rheumatoid arthritis, leading to the production of anti–citrullinated protein antibodies, which drive inflammation and joint damage.[59] Similarly, aberrant proteolytic processing of proinsulin peptides in islet beta cells may promote the development of type 1 diabetes.[60] Beyond posttranslational modification, tissue damage or cellular stress can expose normally hidden or "cryptic" self-epitopes, such as exposure to cardiac antigens triggering post–myocardial infarction pericarditis (Dressler syndrome).[61] Molecular mimicry refers to the phenomenon in which pathogen-associated foreign antigens share structural or sequence similarities with self-antigens, a process best exemplified by Guillain-Barré syndrome in which antibodies against *Campylobacter jejuni* lipooligosaccharides cross-react with gangliosides on peripheral nerves.[62]

Finally, paraneoplastic syndromes highlight how cancer can instigate breaks in B cell tolerance. In antibody-mediated paraneoplastic disorders, this may result from the ectopic expression of self-antigens normally confined to immune-privileged sites (eg, voltage-gated calcium channels in Lambert–Eaton myasthenic syndrome)[63] or from the expression of mutated proteins generating tumor-specific neoantigens. This latter mechanism is best exemplified by temporal links between cancer and scleroderma in patients with autoantibodies to RNA polymerase III subunit (RPC1).[64] In this context, cancer-specific mutations within the *POLR3A* locus, encoding RPC1, can elicit the production of antitumor antibodies, which subsequently exhibit cross-reactivity with the native RPC1 protein. This molecular mimicry can then

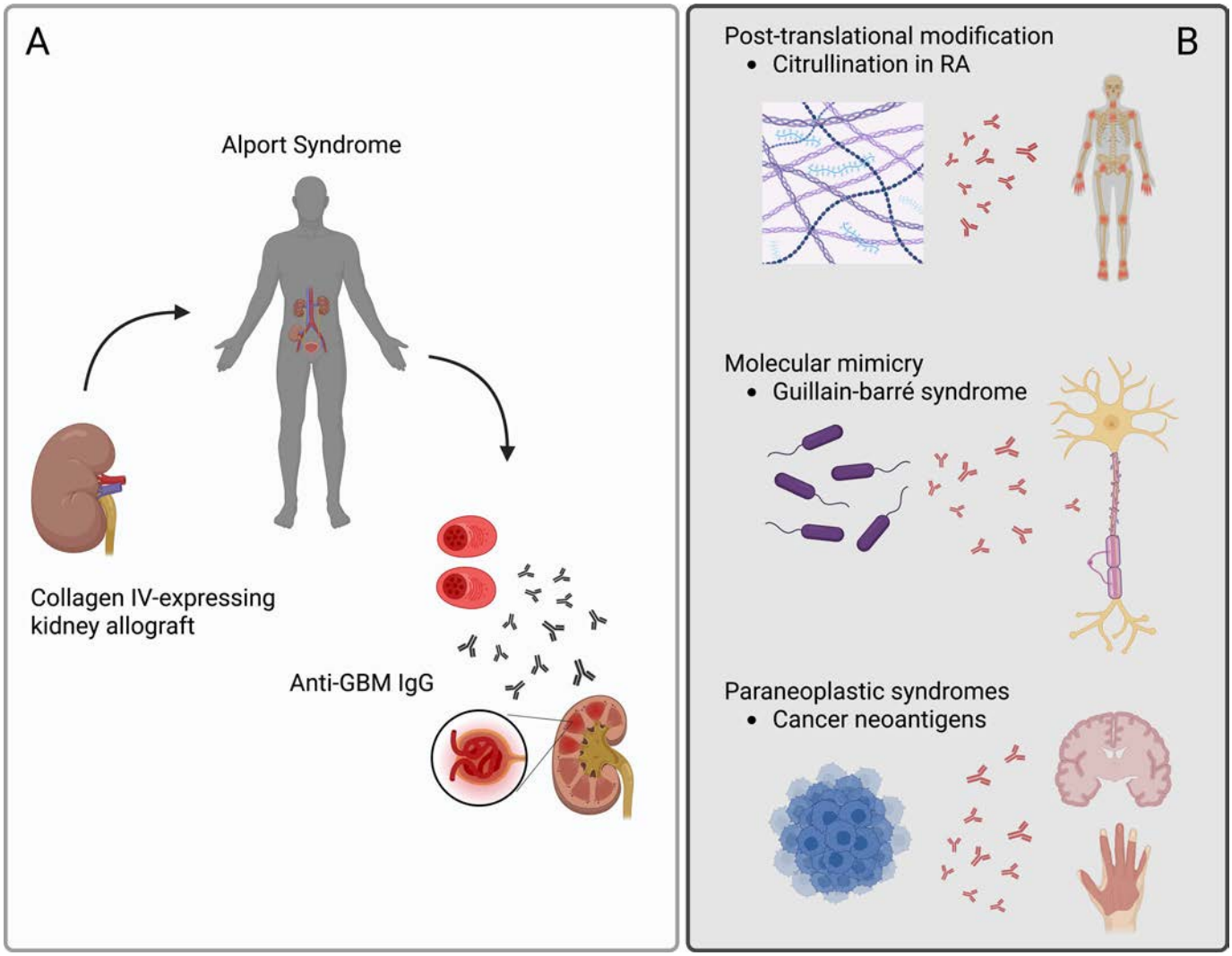


Figure 2. Contribution of neoantigens to breaks in B cell tolerance. This figure illustrates how neoantigens can initiate breaks in B cell tolerance and lead to autoimmunity. (A) The case of de novo anti-GBM disease in a patient with Alport syndrome post kidney transplant, in which introduction of previously absent collagen IV antigens triggers an alloimmune response and the production of anti-GBM antibodies. (B) Additional mechanisms for tolerance breaks through the formation of neoantigens, including via posttranslational modifications (eg, citrullination in RA), molecular mimicry (eg, *Campylobacter jejuni* antigens resembling nerve gangliosides in Guillain-Barré syndrome), and cancer neoantigens (eg, Lambert–Eaton myasthenic syndrome and *POLR3A* mutations in paraneoplastic scleroderma). GBM, glomerular basement membrane; RA, rheumatoid arthritis. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.43316/abstract.

drive systemic inflammation, tissue damage, and fibrosis characteristic of scleroderma.[65]

Conclusions

In conclusion, B cell tolerance functions on a complex and finely tuned spectrum rather than a simple binary state of survival versus deletion. The seemingly paradoxical maintenance of low-affinity autoreactive, anergic B cells is a biologic imperative, preserving a diverse repertoire capable of recognizing foreign antigens, including those mimicking self. Crucially, the process of clonal redemption provides a mechanism for these anergic B cells to be safely recruited into protective immune responses. Nevertheless, maintenance of B cell tolerance is not infallible, with factors such as genetic predisposition toward B cell activation, defects in T cell tolerance, and the emergence of neoantigens each contributing to tolerance breaks. Understanding the mechanisms shaping the B cell repertoire and governing the maintenance of B cell tolerance is paramount for developing effective strategies to prevent and treat humoral autoimmunity.

AUTHOR CONTRIBUTIONS

Both authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding

author, Dr Jackson confirms that all authors have provided the final approval of the version to be published and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. Manjarrez-Orduño N, Quách TD, Sanz I. B cells and immunological tolerance. J Invest Dermatol 2009;129(2):278–288.
2. Rawlings DJ, Metzler G, Wray-Dutra M, et al. Altered B cell signalling in autoimmunity. Nat Rev Immunol 2017;17(7):421–436.
3. Nemazee D. Mechanisms of central tolerance for B cells. Nat Rev Immunol 2017;17(5):281–294.
4. Wardemann H, Yurasov S, Schaefer A, et al. Predominant autoantibody production by early human B cell precursors. Science 2003; 301(5638):1374–1377.
5. Perelson AS, Oster GF. Theoretical studies of clonal selection: minimal antibody repertoire size and reliability of self-non-self discrimination. J Theor Biol 1979;81(4):645–670.
6. Nemazee DA, Burki K. Clonal deletion of B lymphocytes in a transgenic mouse bearing anti-MHC class I antibody genes. Nature 1989; 337(6207):562–566.
7. Hartley SB, Crosbie J, Brink R, et al. Elimination from peripheral lymphoid tissues of self-reactive B lymphocytes recognizing membrane-bound antigens. Nature 1991;353(6346):765–769.
8. Hartley SB, Cooke MP, Fulcher DA, et al. Elimination of self-reactive B lymphocytes proceeds in two stages: arrested development and cell death. Cell 1993;72(3):325–335.
9. Nemazee D. Receptor editing in lymphocyte development and central tolerance. Nat Rev Immunol. 2006;6(10):728–740.
10. Gay D, Saunders T, Camper S, et al. Receptor editing: an approach by autoreactive B cells to escape tolerance. J Exp Med 1993;177(4): 999–1008.
11. Tiegs SL, Russell DM, Nemazee D. Receptor editing in self-reactive bone marrow B cells. J Exp Med 1993;177(4):1009–1020.
12. Duong BH, Ota T, Aoki-Ota M, et al. Negative selection by IgM superantigen defines a B cell central tolerance compartment and reveals mutations allowing escape. J Immunol 2011;187(11):5596–5605.
13. Halverson R, Torres RM, Pelanda R. Receptor editing is the main mechanism of B cell tolerance toward membrane antigens. Nat Immunol 2004;5(6):645–650.
14. Goodnow CC, Crosbie J, Adelstein S, et al. Altered immunoglobulin expression and functional silencing of self-reactive B lymphocytes in transgenic mice. Nature 1988;334(6184):676–682.
15. Merrell KT, Benschop RJ, Gauld SB, et al. Identification of anergic B cells within a wild-type repertoire. Immunity 2006;25(6):953–962.
16. Zikherman J, Parameswaran R, Weiss A. Endogenous antigen tunes the responsiveness of naive B cells but not T cells. Nature 2012; 489(7414):160–164.
17. Cyster JG, Goodnow CC. Antigen-induced exclusion from follicles and anergy are separate and complementary processes that influence peripheral B cell fate. Immunity 1995;3(6):691–701.
18. Cyster JG, Hartley SB, Goodnow CC. Competition for follicular niches excludes self-reactive cells from the recirculating B-cell repertoire. Nature 1994;371(6496):389–395.
19. Förster R, Mattis AE, Kremmer E, et al. A putative chemokine receptor, BLR1, directs B cell migration to defined lymphoid organs and specific anatomic compartments of the spleen. Cell 1996;87(6): 1037–1047.
20. Ekland EH, Forster R, Lipp M, et al. Requirements for follicular exclusion and competitive elimination of autoantigen-binding B cells. J Immunol 2004;172(8):4700–4708.
21. Phan TG, Amesbury M, Gardam S, et al. B cell receptor-independent stimuli trigger immunoglobulin (Ig) class switch recombination and production of IgG autoantibodies by anergic self-reactive B cells. J Exp Med 2003;197(7):845–860.
22. Moore PA, Belvedere O, Orr A, et al. BLyS: member of the tumor necrosis factor family and B lymphocyte stimulator. Science 1999; 285(5425):260–263.
23. Schneider P, MacKay F, Steiner V, et al. BAFF, a novel ligand of the tumor necrosis factor family, stimulates B cell growth. J Exp Med 1999;189(11):1747–1756.
24. Khare SD, Sarosi I, Xia XZ, et al. Severe B cell hyperplasia and autoimmune disease in TALL-1 transgenic mice. Proc Natl Acad Sci U S A. 2000;97(7):3370–3375.
25. Gavin AL, Duong B, Skog P, Ait-Azzouzene D, Greaves DR, Scott ML, et al. deltaBAFF, a splice isoform of BAFF, opposes full-length BAFF activity in vivo in transgenic mouse models. J Immunol 2005;175(1): 319–328.
26. Mackay F, Woodcock SA, Lawton P, et al. Mice transgenic for BAFF develop lymphocytic disorders along with autoimmune manifestations. J Exp Med 1999;190(11):1697–1710.
27. Cheema GS, Roschke V, Hilbert DM, et al. Elevated serum B lymphocyte stimulator levels in patients with systemic immune-based rheumatic diseases. Arthritis Rheum 2001;44(6):1313–1319.
28. Zhang J, Roschke V, Baker KP, et al. Cutting edge: a role for B lymphocyte stimulator in systemic lupus erythematosus. J Immunol 2001;166(1):6–10.
29. Thien M, Phan TG, Gardam S, et al. Excess BAFF rescues self-reactive B cells from peripheral deletion and allows them to enter forbidden follicular and marginal zone niches. Immunity 2004;20(6): 785–798.
30. Lesley R, Xu Y, Kalled SL, et al. Reduced competitiveness of autoantigen-engaged B cells due to increased dependence on BAFF. Immunity 2004;20(4):441–453.
31. Hondowicz BD, Alexander ST, Quinn WJ, et al. The role of BLyS/BLyS receptors in anti-chromatin B cell regulation. Int Immunol 2007;19(4): 465–475.
32. Akkaraju S, Canaan K, Goodnow CC. Self-reactive B cells are not eliminated or inactivated by autoantigen expressed on thyroid epithelial cells. J Exp Med 1997;186(12):2005–2012.
33. Aplin BD, Keech CL, de Kauwe AL, et al. Tolerance through indifference: autoreactive B cells to the nuclear antigen La show no evidence of tolerance in a transgenic model. J Immunol 2003;171(11):5890–5900.
34. Victora GD, Nussenzweig MC. Germinal Centers. Annu Rev Immunol 2022;40:413–442.
35. Rathmell JC, Townsend SE, Xu JC, et al. Expansion or elimination of B cells in vivo: dual roles for CD40- and Fas (CD95)-ligands modulated by the B cell antigen receptor. Cell 1996;87(2):319–329.
36. Garrone P, Neidhardt EM, Garcia E, et al. Fas ligation induces apoptosis of CD40-activated human B lymphocytes. J Exp Med 1995; 182(5):1265–1273.
37. Rathmell JC, Cooke MP, Ho WY, et al. CD95 (Fas)-dependent elimination of self-reactive B cells upon interaction with CD4+ T cells. Nature 1995;376(6536):181–184.
38. Shokat KM, Goodnow CC. Antigen-induced B-cell death and elimination during germinal-centre immune responses. Nature 1995; 375(6529):334–338.
39. Pulendran B, Kannourakis G, Nouri S, et al. Soluble antigen can cause enhanced apoptosis of germinal-centre B cells. Nature 1995; 375(6529):331–334.

40. Han S, Zheng B, Dal Porto J, et al. In situ studies of the primary immune response to (4-hydroxy-3-nitrophenyl)acetyl. IV. Affinity-dependent, antigen-driven B cell apoptosis in germinal centers as a mechanism for maintaining self-tolerance. J Exp Med 1995;182(6): 1635–1644.
41. Burnett DL, Langley DB, Schofield P, et al. Germinal center antibody mutation trajectories are determined by rapid self/foreign discrimination. Science 2018;360(6385):223–226.
42. Vinuesa CG, Sanz I, Cook MC. Dysregulation of germinal centres in autoimmune disease. Nat Rev Immunol 2009;9(12):845–857.
43. Aloisi F, Pujol-Borrell R. Lymphoid neogenesis in chronic inflammatory diseases. Nat Rev Immunol 2006;6(3):205–217.
44. Wellmann U, Letz M, Herrmann M, et al. The evolution of human anti-double-stranded DNA autoantibodies. Proc Natl Acad Sci U S A 2005;102(26):9258–9263.
45. Matsumoto I, Staub A, Benoist C, et al. Arthritis provoked by linked T and B cell recognition of a glycolytic enzyme. Science 1999; 286(5445):1732–1735.
46. Kouskoff V, Korganow AS, Duchatelle V, et al. Organ-specific disease provoked by systemic autoimmunity. Cell 1996;87(5):811–822.
47. Vazquez SE, Ferré EM, Scheel DW, et al. Identification of novel, clinically correlated autoantigens in the monogenic autoimmune syndrome APS1 by proteome-wide PhIP-Seq. Elife. 2020;9:e55053.
48. Wang EY, Dai Y, Rosen CE, et al. High-throughput identification of autoantibodies that target the human exoproteome. Cell Rep Methods 2022;2(2):100172.
49. Kinnunen T, Chamberlain N, Morbach H, et al. Accumulation of peripheral autoreactive B cells in the absence of functional human regulatory T cells. Blood 2013;121(9):1595–1603.
50. Fu W, Liu X, Lin X, et al. Deficiency in T follicular regulatory cells promotes autoimmunity. J Exp Med 2018;215(3):815–825.
51. Linterman MA, Pierson W, Lee SK, et al. Foxp3+ follicular regulatory T cells control the germinal center response. Nat Med 2011;17(8): 975–982.
52. Chung Y, Tanaka S, Chu F, et al. Follicular regulatory T cells expressing Foxp3 and Bcl-6 suppress germinal center reactions. Nat Med 2011;17(8):983–988.
53. Burnett DL, Reed JH, Christ D, et al. Clonal redemption and clonal anergy as mechanisms to balance B cell tolerance and immunity. Immunol Rev 2019;292(1):61–75.
54. Jackson SW, Kolhatkar NS, Rawlings DJ. B cells take the front seat: dysregulated B cell signals orchestrate loss of tolerance and autoantibody production. Curr Opin Immunol 2015;33:70–77.
55. Deng Y, Tsao BP. Updates in Lupus Genetics. Curr Rheumatol Rep 2017;19(11):68.
56. Hertz JM, Thomassen M, Storey H, et al. Clinical utility gene card for: Alport syndrome – update 2014. Eur J Hum Genet 2015;23(9):1269.
57. Kuang H, Jiang N, Jia XY, et al. Epidemiology, clinical features, risk factors, and outcomes in anti-glomerular basement membrane disease: a systematic review and meta-analysis. Autoimmun Rev 2024; 23(4):103531.
58. Kashtan CE, Ding J, Garosi G, et al. Alport syndrome: a unified classification of genetic disorders of collagen IV alpha345: a position paper of the Alport Syndrome Classification Working Group. Kidney Int 2018;93(5):1045–1051.
59. Deane KD, El-Gabalawy H. Pathogenesis and prevention of rheumatic disease: focus on preclinical RA and SLE. Nat Rev Rheumatol 2014; 10(4):212–228.
60. Jin N, Wang Y, Crawford F, et al. N-terminal additions to the WE14 peptide of chromogranin A create strong autoantigen agonists in type 1 diabetes. Proc Natl Acad Sci U S A 2015;112(43):13318–13323.
61. Malik J, Zaidi SMJ, Rana AS, et al. Post-cardiac injury syndrome: an evidence-based approach to diagnosis and treatment. Am Heart J Plus 2021;12:100068.
62. Ang CW, Jacobs BC, Laman JD. The Guillain-Barré syndrome: a true case of molecular mimicry. Trends Immunol 2004;25(2):61–66.
63. Tarr TB, Wipf P, Meriney SD. Synaptic pathophysiology and treatment of Lambert-Eaton myasthenic syndrome. Mol Neurobiol 2015;52(1): 456–463.
64. Shah AA, Rosen A, Hummers L, et al. Close temporal relationship between onset of cancer and scleroderma in patients with RNA polymerase I/III antibodies. Arthritis Rheum 2010;62(9):2787–2795.
65. Joseph CG, Darrah E, Shah AA, et al. Association of the autoimmune disease scleroderma with an immunologic response to cancer. Science 2014;343(6167):152–157.

Arthritis & Rheumatology
Vol. 78, No. 3, March 2026, pp 566–581
DOI 10.1002/art.43326

American College of Rheumatology
Empowering Rheumatology Professionals

REVIEW

Case-Based Immunology: B Cells and Systemic Sclerosis Interstitial Lung Disease

Nina Goldman, Voon Ong, and Christopher P. Denton

Interstitial lung disease (ILD) is an important complication of systemic sclerosis (SSc), with high mortality and morbidity. Recent clinical studies in SSc-ILD have led to US Food and Drug Administration–approved therapies in SSc-ILD. Importantly, evidence from these studies has been extrapolated to guide management of ILDs of other systemic autoimmune rheumatic diseases. Pathogenesis of SSc-ILD involves interplay between fibroblasts and the innate and adaptive immune system. A central role for the B cell compartment is supported by clinical and translational studies. We use a case from our center as a basis to discuss the pathogenesis of SSc-ILD, autoantibodies in SSc-ILD, and the role of B cells in the disease. We go on to consider treatment options for the case, the decision-making algorithm for treatment, and risks associated with treatment.

The clinical case

The patient, a 69-year-old man diagnosed with diffuse cutaneous systemic sclerosis[1] (dcSSc) in 2014, returned for clinic review in May 2022. He reported increasing dyspnea and worsening exercise tolerance over the preceding year without significant cough, chest pain, or ankle edema. He has had recurrent urinary tract infections requiring multiple antibiotic courses but no documented chest infection.

A history review revealed first presentation with Raynaud phenomenon symptoms in September 2013, followed by digital puffiness in January 2014 with progressive skin thickening over the upper limbs and anterior chest over the next six months. He reported reflux, proximal muscle weakness, and weight loss of 6 kg over a similar interval. He was referred to his local rheumatologist and was diagnosed with dcSSc in August 2014. The modified Rodnan skin score (mRSS) peaked at 19/51 in September 2014. Initial investigations at disease presentation in August 2014 showed a hemoglobin level of 127 g/L, a creatine kinase level of 757 IU, a C-reactive protein (CRP) level of <1 mg/L, an erythrocyte sedimentation rate (ESR) of 2 mm/hr, and an N-terminal pro–brain natriuretic peptide (NT-proBNP) level of 26 pmol/L (high >236 pmol/L). Esophageal dysmotility was confirmed with pH manometry and motility studies. Interstitial lung disease (ILD) was confirmed on a high-resolution computed tomography (HRCT) scan at the time of diagnosis.

In August 2014, he was started on mycophenolate mofetil (MMF) 2 g daily for active skin disease and ILD, with prednisolone 20 mg daily for myositis. In late 2014, he developed complete heart block necessitating permanent pacemaker implantation. A follow-up echocardiogram in 2015 showed a left ventricular ejection fraction of 45% and regional wall motion abnormality with slight hypokinesia of the left ventricular apex.

His maintenance immunosuppressive therapies comprise MMF 2 g and prednisolone 7.5 mg daily. At the clinic review, the patient reports ongoing SSc gastrointestinal involvement alongside his respiratory symptoms. Despite high-dose proton pump inhibitors and H_2 receptor antagonists for gastroesophageal reflux disease, prokinetics for esophageal dysmotility, and rotational antibiotics for small intestinal bacterial overgrowth, he remained symptomatic with vomiting, diarrhea, and bloating. His weight was maintained with high-calorie supplements.

On examination, the mRSS was stable at 9/51. Chest auscultation demonstrated fine bilateral inspiratory crackles. Autoimmune serology demonstrated speckled-pattern antinuclear autoantibodies (ANAs) at a titer of 1:1,280 and negative extractable nuclear antigens, with no SSc-specific antibody identified by standard immunoblot. Available laboratory test results

Dr Goldman's work was supported by the Medical Research Council and Scleroderma & Raynaud's UK (grant MR/V030108/1).

Nina Goldman, MB BChir, Voon Ong, PhD, FRCP, Christopher P. Denton, PhD, FRCP, FMedSci: Centre for Rheumatology, Division of Medicine, University College London, London, United Kingdom.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/art.43326).

Author disclosures and graphical abstract are available at https://onlinelibrary.wiley.com/doi/10.1002/art.43326.

Address correspondence via email to Christopher P. Denton, PhD, FRCP, FMedSci, at c.denton@ucl.ac.uk.

Submitted for publication September 26, 2024; accepted in revised form July 10, 2025.

included a hemoglobin level of 154 g/L, a CRP level of 2 mg/L, and an ESR of 5 mm/hr. Cardiac markers performed included the following results: an NT-proBNP level of 230 ng/L and a troponin T level of 19 ng/L. The recent echocardiogram demonstrated a low probability of pulmonary hypertension.

Case progress

Results of a repeat lung function test are shown in Supplemental Table 1 and Supplemental Figure 1. The lung function test showed forced vital capacity (FVC) and diffusing capacity for carbon monoxide (DL_{CO}) decline of 18.2% and 11.8%, respectively, over the preceding year from May 2021 to May 2022.

This case highlights development of SSc-ILD and provides a platform for discussion of recent progress in understanding pathogenesis, risk stratification, and treatment options for this complication, which is one of the leading causes of mortality and morbidity in SSc.[2,3] Informed consent was obtained for use of the clinical information from the patient for publication.

Pathogenesis of SSc-ILD

It is hypothesized that initiation of SSc-ILD involves both endothelial and epithelial cell damage triggered by environmental factors in a susceptible individual. This cellular damage results in the release of locally active mediators such as endothelin 1 and chemokines that may precede activation of the immune system.[4] Immune system activation then results in an inflammatory cell infiltrate including B cells, macrophages, T cells, and natural killer cells and profibrotic pathway activation, particularly transforming growth factor β (TGFβ) pathways. Subsequently, fibroblasts are activated locally; transdifferentiate from other cell lineages, including epithelial cells and vascular components such as pericytes and endothelial cells[5]; and are recruited and differentiate into profibrotic cells producing excessive extracellular matrix, resulting in deposition of the extracellular matrix within the lung parenchyma and further TGFβ release.[6] This perpetuates the cycle of lung damage alongside ongoing inflammation and abnormal healing. Other factors such as infection, aspiration, and increased lung stiffness also play a role in exacerbating lung damage.

Genetic and epigenetics changes have been demonstrated to influence an individual's susceptibility to SSc-ILD, with HLA-related and non–HLA-related genetic polymorphisms implicated. The strongest genetic association with ILD is linked to ANA subgroup. Studies have confirmed that ANA reactivity links closely to major histocompatibility complex (MHC) class II haplotype, with *HLA-DRB1*11* conferring considerable risk in patients who are positive for anti–topoisomerase I antibodies (ATAs).[7]

Many risk factors for development and progression of SSc-ILD have been identified. Patients with dcSSc are at higher risk of clinically significant ILD compared to patients with limited cutaneous SSc (lcSSc).[8] Other risk factors for progressive disease include early disease duration, male sex, African American race, autoantibody subset, reflux, low baseline FVC and DL_{CO}, CT extent above 20%, and FVC decline of at least 10% from baseline.[9] Elevated levels of interleukin-6 (IL-6), the proinflammatory pleiotropic cytokine, have also been shown to be predictive of progression of early ILD, and high CRP levels have also been associated with worse SSc-ILD outcome.[10]

Understanding the autoantibody profile

The presence of ANAs is a hallmark of SSc, found in more than 95% of patients with SSc.[11] Due to their mutual exclusivity, likely a consequence of MHC restriction,[12] ANA specificities have emerged as central in helping to stratify the risk of organ complications for patients, with the most frequent reactivities being anticentromere antibody (ACA), ATA, and anti–RNA polymerase III antibody (anti–RNAP III), found in >50% of patients with SSc.[11,13]

The focus in SSc-ILD remains on ATA-positive patients, with ATA positivity resulting in a much higher risk of clinically significant ILD than low-risk ACA and intermediate-risk anti–RNAP III.[8] However, other SSc-specific autoantibodies, including anti-Th/To, anti–eukaryotic initiation factor 2b (anti-eIF2b), and anti–U11/U12 RNP, and SSc overlap–associated antibodies anti-PM/Scl and anti-Ku are also associated with a high risk of SSc-ILD. Other novel autoantibodies, including anti–nucleolar organizer region antibodies and anti–Bicaudal D protein antibodies, have been variably associated with ILD[14–16] (Table 1). The ANA immunofluorescence pattern can be of assistance, particularly with autoantibodies not routinely tested in clinical practice, with anti-PM/Scl and anti-Th/To resulting in a nucleolar pattern, anti–U11/12 RNP (also known as anti-RNPC3) and anti-Ku resulting in a speckled pattern, and anti-eIF2b resulting in a cytoplasmic pattern.[17] The American College of Rheumatology (ACR) and the American College of Chest Physicians (CHEST) guidelines list ATA positivity and ANA with a nucleolar pattern as risk antibodies for progressive SSc-ILD; however, other rarer high-risk antibodies are being defined and should be considered in SSc-ILD screening.[18]

Anti–U11/U12 RNP has been found in 3% to 8% of patients with SSc.[19–22] U11 and U12 are components of the spliceosome and are involved in splicing of pre–messenger RNA.[23] U11/U12 autoantibodies that react with the 65-kDa protein component in the U11/U12 small nuclear RNP particle complex were first reported in a patient with SSc by Gilliam and Steitz[24] and occurs in both skin subtypes of SSc and has been associated with severe ILD.[19,21] Patients with this antibody are more likely to be male and have moderate to severe gastrointestinal involvement.[25] In one study, anti–U11/U12 RNP has also been associated with malignancy, with a short interval between SSc diagnosis and cancer diagnosis, although this was not replicated in a subsequent study.[22,26]

A repeat chest HRCT scan was performed, and images are shown in Figure 1. The fibrotic changes had significantly

Table 1. Main associations of autoantibody profile in scleroderma-associated ILD*

Antibody	ANA pattern	Intracellular target	Prevalence in SSc cohort	Cutaneous subset association	Lung involvement	Comments
ATA	Homogeneous or nucleolar/speckled	Type I topoisomerase	20%–30%	Diffuse in 60%	80% develop ILD, of which up to 50% may be progressive	Prognostic for lung function decline irrespective of skin subset; predictive of response to tocilizumab
anti–RNAP III	Nucleolar/homogenous	RNA polymerase type 3	4%–20%	Diffuse phenotype with early severe skin disease followed by potential for rapid spontaneous resolution	Propensity for late lung function decline	–
Anti-Th/To	Nucleolar	Nucleolar 7-2/8-2 RNA protein complex	2%–6%	Limited or sine SSc	50% develop, of which 30% progress	Coexisting pulmonary hypertension may occur
Anti-U11/U12	Speckled	U11/U12 RNA polymerase	1%–3%	Limited/diffuse	80% develop often severe and progressive ILD	Coexisting severe gut involvement with increased risk of cancer
Anti-PM-Scl	Nucleolar	Nucleolar PM/Scl macromolecular complex	3%–6% (25% of SSc–myositis overlap)	Limited or sine SSc	35%–87% develop with good outcome	Organizing pneumonia pattern may occur in context of myositis overlap phenotype
Anti-Ku	Speckled	Ku complex (p70/p80 heterodimer)	2%–7% (15% of SSc–myositis overlap)	Limited	Up to 76% develop ILD with good outcome	–
Anti–U1 RNP	Coarse speckled	Small nuclear ribonucleoproteins	5%–35% (100% in MCTD)	Limited	35% develop, of which 20% progress	–
Anti-eIF2B	ANA-negative with cytoplasmic staining	Eukaryotic initiation factor 2b	1%–2%	Diffuse	High incidence and up to 100% develop	Possible malignancy association
Anti-RuvBL1/2	Speckled	RuvBL1/2 double hexamer	1%–2%	Diffuse and myositis overlap	Higher frequency in men and higher age at onset with myositis, including cardiac involvement and gut dysmotility	–
Anti-RNPC3	Speckled	Small nuclear ribonucleoproteins U11 and U12	3%–5%	Diffuse and limited	Coexisting PAH and GI disease and myositis and malignancy	–
Anti-BICD2	Speckled and/or nucleolar?	Bicaudal D protein	20%–35%	Diffuse	Associated myositis, may coexist with ACA	–
Anti-NOR90	Nucleolar punctate		2%–3%	Limited	Good prognosis	Variably associated with increased risk of ILD
Anti-PRMT5	Enzyme belonging to arginine methyltransferases	Unknown	31%	Diffuse	May coexist with ATA	–
Anti-TERF1	Unknown		9%	Limited and diffuse	May occur with anti-Ku and anti–U1 RNP	–

* ACA, anticentromere antibody; ANA, antinuclear autoantibody; anti-eIF2b, anti–eukaryotic initiation factor 2b; anti–RNAP III, anti–RNA polymerase III antibody; ATA, anti–topoisomerase I antibody; GI, gastrointestinal; ILD, interstitial lung disease; MCTD, mixed connective tissue disease; PAH, pulmonary arterial hypertension; SSc, systemic sclerosis.

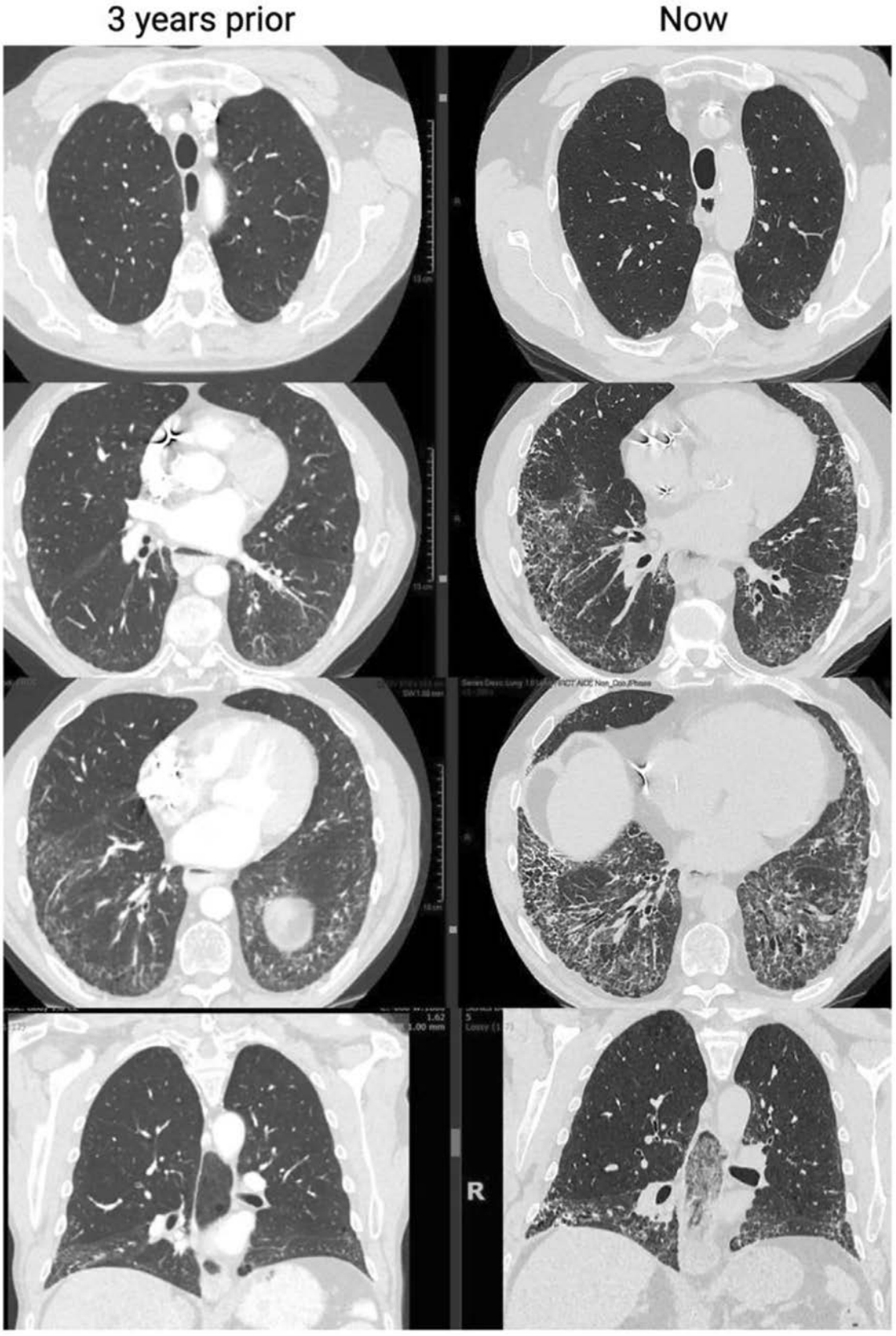


Figure 1. Legend on next page.

progressed from a previous scan three years prior. Following review of the investigations, a decision was made to treat with rituximab, a monoclonal antibody targeting CD20, leading to B cell depletion.

Evidence for the role of B cells in SSc-ILD

Dysregulated B cell function has been increasingly implicated in SSc and SSc-ILD pathogenesis (Figure 2). Patients with SSc are consistently found to have hypergammaglobulinemia, with SSc-specific autoantibodies found in the majority of patients.[11,29] As discussed previously, these autoantibodies strongly correlate with clinical phenotypes. However, levels of the autoantibodies have not been found to consistently be important in disease severity or trajectory, and their role in the direct pathogenesis of disease continues to be debated.[8,30,31] Autoantibodies have been demonstrated to precede symptom onset.[32] SSc-specific antibodies (ACA, anti–RNAP III, and ATA) were the best predictive marker of progression to SSc in the Very Early Diagnosis of Systemic Sclerosis (VEDOSS) cohort.[33]

Non–ANA-targeting autoantibodies have been described as functional autoantibodies; however, these are not routinely used in clinical practice. Antifibroblast antibodies found in patients with SSc have been demonstrated to up-regulate profibrotic chemokines from fibroblasts.[34] Autoantibodies targeting G protein–coupled receptors are thought to play a role in both immune response regulation in health and pathogenesis of autoimmune diseases, including SSc. For example, high concentrations of anti–endothelin 1 type A receptor and anti–angiotensin II type I receptor antibodies have been associated with a poor prognosis in SSc, and autoantibodies from patients with SSc promote vasoconstriction, fibrosis, and inflammation.[35,36]

The strongest evidence of the pathogenicity of ANAs comes from ATA. Identification of several B cell immunodominant epitopes elegantly links the ATA response with development of SSc-ILD.[37] Immunization of mice with topoisomerase I and Freund's complete adjuvant has been demonstrated to induce skin and lung fibrosis; however, it remains unclear the contribution of ATA in this model and how ATA exerts its affect.[38,39] One proposed mechanism by which ATA directly causes fibrosis is that the autoantigen topoisomerase I, released following endothelial cell damage, binds to fibroblast surfaces, leading to recruitment of ATA. This subsequently leads to monocyte adhesion, proinflammatory cytokine release, and fibroblast activation.[40] Within ATA-producing ATA-positive $CD27^+$ B cells, cells with a high affinity for topoisomerase I have been found to have high production of proinflammatory cytokines, including IL-6, and those with a low affinity for topoisomerase I have been found to have higher production of anti-inflammatory cytokines, suggesting antibody affinity may be important in SSc pathogenesis.[41] Anti–topoisomerase I IgM is not routinely measured, but in a cohort of ATA-positive patients with SSc, patients with progressive disease were more often positive for anti–topoisomerase I IgM than nonprogressors, suggesting that an ongoing autoreactive B cell response is associated with disease progression.[42] However, mixed data have been found in SSc regarding depletion of ATA levels and correlation with clinical outcomes for autologous hematopoietic stem cell transplantation (AHSCT) and B cell depletion.[43–46]

B cell infiltration has been demonstrated in SSc, with work predominantly performed in skin due to the increased tissue accessibility compared to lungs. A small study demonstrated high expression of B lymphocyte signatures in both affected and unaffected skin in patients with early dcSSc.[47] Another study of skin in early diffuse SSc found the majority of biopsy samples had B cell signatures (69%) along with other adaptive and innate immune cell signatures.[48] B cell infiltration of lung tissue has been found in patients with both nonspecific interstitial pneumonia (NSIP) and usual interstitial pneumonia–pattern SSc-ILD.[49,50] B cells may also organize into bronchus-associated lymphoid tissue in SSc.[51] Lymphocytosis has been found in bronchoalveolar lavage (BAL) fluid in some patients with SSc-ILD, and a higher CD19 level has been associated with increased progression of ILD.[52] However, the utility of BAL fluid lymphocytosis to predict progression has not been consistently demonstrated.[53]

Although absolute B cell numbers are not found to be altered in SSc, abnormal B cell homeostasis is found in disease. Increased naïve B cells and transitional B cells and reduced memory B cells have been demonstrated.[54,55] It is thought that both central and peripheral B cell tolerance are defective in SSc, resulting in autoreactive B cells involved in disease pathogenesis.[56]

B cells in SSc are thought to have a lower activation threshold, with increased presence of CD19, the positive B cell regulator, and reduction of the B cell inhibitor CD22 found in dcSSc but not lcSSc.[57–59] BAFF and APRIL both play roles in B cell survival and activation. Inhibition of BAFF reduces skin and lung fibrosis and IL-6–producing B cells in the bleomycin mouse model, a commonly used model of SSc.[60] Elevated BAFF levels have been found in some studies of patients with SSc, with higher levels of

Figure 1. Chest high-resolution computed tomography demonstrated progression from the scan three years prior. The scan remained in keeping in with NSIP, with ground glass opacification and reticulation alongside traction bronchiectasis and some freestanding bronchiectasis. The straight-edge sign associated with NSIP was present. This radiologic feature is commonly reported with NSIP and less so with UIP interstitial lung disease.[27] Diffuse dendriform pulmonary ossification, which is often associated with UIP and can also be found in NSIP, was seen.[28] Subpleural sparing present on the previous scan was lost due to progressive fibrosis, with increased volume loss and progression of the morphology of fibrosis. Esophageal dilatation with fluid level is also demonstrated. NSIP, nonspecific interstitial pneumonia; UIP, usual interstitial pneumonia.

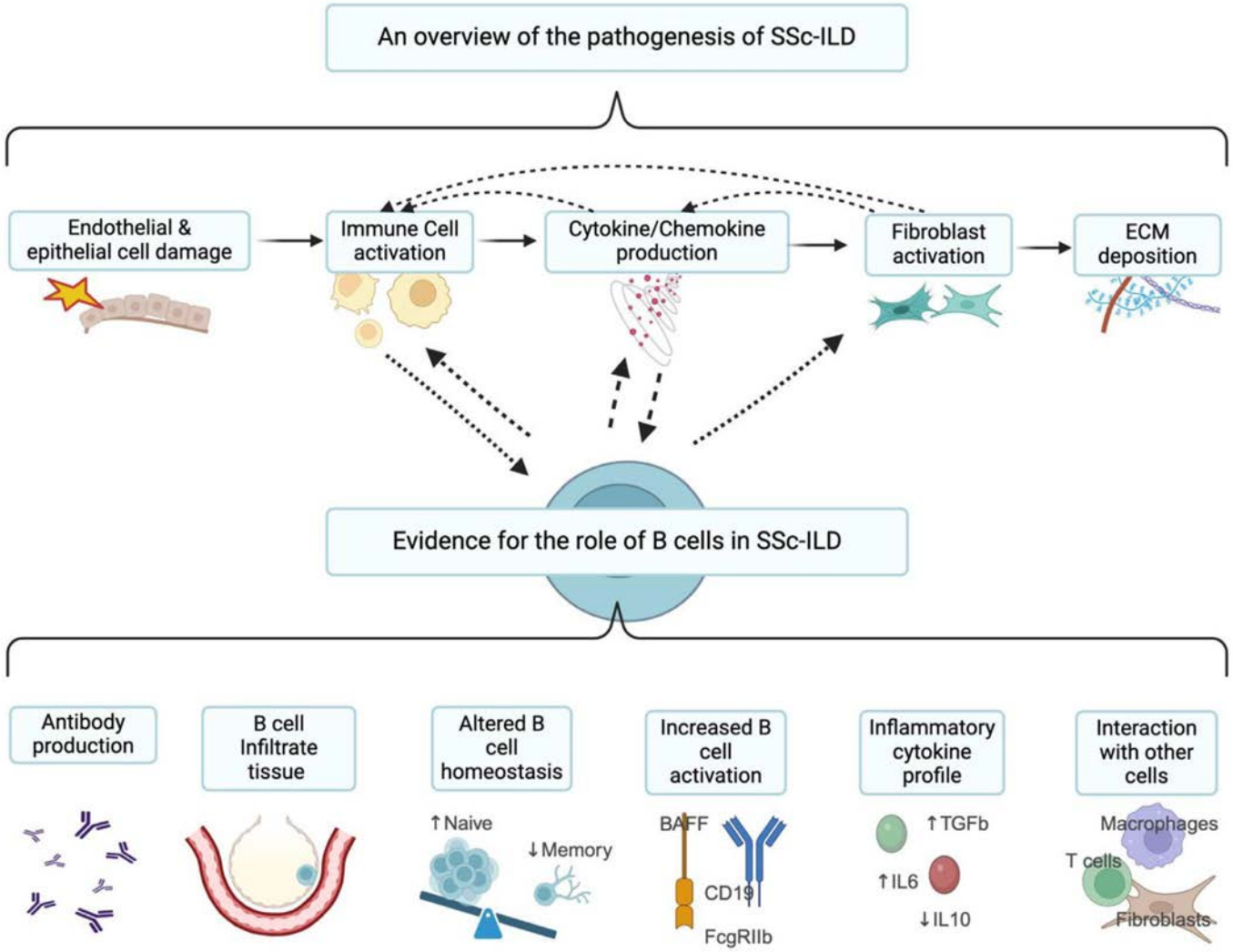


Figure 2. A summary of the pathogenesis of SSc-ILD alongside the role of B cells in SSc-ILD. ECM, extracellular matrix; ILD, interstitial lung disease; SSc, systemic sclerosis.

both correlated with both dcSSc and ILD.[60,61] However other studies have not replicated these results.[56]

In SSc, B cells have a more proinflammatory and profibrotic phenotype. Patients with SSc have been found to have increased IL-6–producing B cells, and although not statistically significant, a trend toward increased IL-6–positive B cells in patients with ATA positivity and severe lung fibrosis has been seen.[62,63] Reduced production of the anti-inflammatory cytokine IL-10 by B cells is also reported in SSc.[64]

B cells interact directly and indirectly with other immune and stromal cells to promote fibrosis in SSc (Figure 2). B cell supernatant from patients with SSc and B cell production of antifibroblast antibody results in fibroblast activation.[34,62] T and B cell interaction has been demonstrated to be essential in ATA production in SSc.[65] Increased circulating follicular helper T cells have been found in SSc, and these cells are important in the maturation and development of long-lived plasma cells.[66] B cells have also been demonstrated to induce profibrotic macrophages in a mouse model of SSc.[67] Evidence for the role of B cells in the pathogenesis of SSc is compelling, and further evidence has also arisen from the development of treatments for SSc and SSc-ILD, which are considered in the following section.

Treatment options for SSc-ILD

Several guidelines and recommendations have recently been published regarding the management of SSc and SSc-ILD.[68,69] This reflects the changing landscape of available treatments. The recent ACR/CHEST guideline includes recommendations on screening, monitoring, and treatment of selected systemic autoimmune rheumatic diseases (SARDs) at risk of developing into ILD.[18] In contrast, the American Thoracic Society (ATS) and the recently updated British Society for Rheumatology (BSR) guidelines were focused on SSc-ILD and SSc, respectively.[70] The

treatment strategies in these guidelines were broadly similar, with recommendation on immunosuppression as the mainstay approach and use of antifibrotics if disease progresses despite initial management. ACR/CHEST guideline was cognizant of the variable access to therapies in different countries, and therapeutic recommendations were divided into first-line therapies and progression after initial therapy. The ACR/CHEST guideline conditionally recommends against upfront combination of antifibrotics with MMF over MMF alone as a first-line treatment for SARD-ILD. In contrast, combination MMF and nintedanib was conditionally recommended in the ATS guideline. Specific comments on individual therapy are summarised[68–76] in Table 2.

MMF is recommended as a first-line therapy, with evidence from the Scleroderma Lung Study II showing equivalence of MMF to cyclophosphamide but an improved side effect profile with MMF.[76] Cyclophosphamide can be used as a rescue or alternative to MMF. Given the significant reduction in ILD progression, MMF was recommended as a first-line therapy in both the ATS and the BSR guidelines (Table 2). Prednisolone, however, is not recommended in SSc-ILD due to increased risk of scleroderma renal crisis. The ACR/CHEST guideline strongly recommends against steroid use, and both the ATS and the ACR/CHEST guidelines advise that the daily dose should not exceed the equivalent of 15 mg of prednisone. The risk and benefit of steroid use need to be carefully considered in the context of coexisting inflammatory disease, such as myositis or arthritis, and alternative use of immunosuppression should be considered if the patient is at risk of developing renal crisis.

Tocilizumab, the monoclonal antibody against IL-6, has been demonstrated to stabilize FVC in a randomized controlled trial (RCT) of SSc-ILD in early diffuse inflammatory SSc.[75,77] More recently, rituximab in the RECITAL and DESIRES RCTs has also been found to stabilize lung function.[78–80] The EVER-ILD trial of NSIP-pattern ILD, which included a small cohort of patients with SSc-ILD, suggested that combination rituximab and MMF was superior to MMF alone in improvement of FVC.[81] The benefit of combination therapy on FVC response was not demonstrated to be sustained at 12 months; however, progression-free survival was better with combination therapy, and, of note, the majority of patients did not receive maintenance rituximab therapy.[82] In the BSR guideline, rituximab and tocilizumab are recommended as an addition to MMF as rescue therapy. Tocilizumab is suggested as a possible first-line therapy in patients with early dcSSc and raised levels of markers of inflammation and particularly those who are ATA positive.[70] The ACR/CHEST guideline also recommends rituximab and tocilizumab as possible alternative first-line options to MMF.[69] A recent retrospective analysis suggests that ATA-positive patients have a particular benefit from tocilizumab therapy compared to ATA-negative patients.[83] Interestingly our data also suggest that a more diverse cohort of patients continue to benefit from rituximab and tocilizumab therapy than were included in clinical trials, including those who remain refractory to standard immunosuppression irrespective of disease duration, inflammatory response, and skin subset.

Robust data from clinical trials support use of AHSCT in specialist centers for dcSSc, and this can be used for SSc-ILD if disease progresses despite initial management; however, careful patient selection must be performed.[84] Severe internal organ involvement precludes AHSCT due to high treatment-related toxicity. Further evidence is required regarding both the outcomes with early use of AHSCT, which the ongoing UPSIDE trial is addressing, and comparisons of AHSCT with early combination immunosuppression.[85]

Nintedanib has been demonstrated to be effective in attenuating lung function decline in SSc-ILD in patients with a progressive fibrosing phenotype.[86,87] The ACR/CHEST and BSR guidelines preferentially recommend nintedanib as an additional therapy after immunosuppression; however, the ACR/CHEST guideline acknowledges it could be a first-line therapy in certain patients, and the BSR guideline recommends that it could be used in upfront combination with MMF for patients with extensive ILD at disease onset.[70]

Evidence for the B cell therapies

The success of B cell therapies, including rituximab, in improving outcomes in SSc-ILD provides support for the key role of B cells in the SSc pathogenesis, and an increasing number of studies that target B cells are currently in progress in SSc-ILD (Table 3). Rapid peripheral B cell depletion at two weeks after rituximab treatment has been correlated with treatment response for SSc-ILD, and a small retrospective study found that complete B cell depletion was associated with improved response to rituximab in connective tissue disease associated ILD (CTD-ILD).[91,92]

Rituximab, however, does not effectively deplete all B cells. For example, B cell precursors and plasmablasts that lack CD20 expression are not targeted, and memory B cells in lesional tissue can be relatively protected.[93] It has been considered whether this failure to deplete all B cells and lack of change in autoantibody titers after rituximab treatment result in failure or incomplete treatment response.[45,46]

AHSCT results in a shift of B cell homeostasis, with an increase in naïve B cells and a reduction in memory B cells. The B cell compartment appears to shift to a more Breg cell phenotype, with an increase in IL-10 levels and a reduction in IL-6–producing B cells.[94] However, the high treatment-related toxicity of AHSCT has led to the consideration of CD19 chimeric antigen receptor (CAR) T cells as an alternative cellular therapy that may have an advantageous safety profile.[95,96]

CAR T cell therapy that targets CD19 potentially may overcome the limitations of CD20-targeted strategies, including rituximab, and may target lesional pathogenic antibodies, conferring a deep depletive effect. This is supported by individual case studies and case series of autologous CD19 CAR T cell strategies, which

Table 2. Drugs used for the treatment of scleroderma-associated ILD*

Drug name	Dose	Immunologic basis for action	Clinical evidence for efficacy	Adverse effects	Comments
Cyclophosphamide	Oral: 2 mg/kg daily; intravenous: 600 mg/m^2 body surface area every 3–4 wk for a total of 6 doses	An alkylating agent that prevents replication of proliferating cells, especially bone marrow–derived leukocytes	Cyclophosphamide was associated with mean change in FVC% predicted of 2.8% at 12 mo. There was an improvement in FVC% predicted at 12 mo in a greater proportion of patients receiving cyclophosphamide compared with placebo (49.3% vs 26.4%, respectively).	Leukopenia and thrombocytopenia with increased risk of infections	Based on efficacy data, this is recommended as one of the first-line options and for those with ILD progression in the ACR/CHEST guideline. However, the adverse effect profile led to cyclophosphamide being considered an "additional option" rather than a preferred first-line treatment. In the ATS and BSR guidelines, this is an alternative to mycophenolate. Benefit wanes a year after cessation of cyclophosphamide.
MMF	2 g daily (maximum 3 g daily)	An inhibitor of lymphocyte guanine nucleotide synthesis that targets the enzyme inosine monophosphate dehydrogenase; some evidence that effect of MMF may be attributable, at least in part, to macrophage viability and altered alternative activation	Compared to placebo, MMF use confers a 5% difference in mean FVC% predicted of 2.8% with improvement in breathlessness (assessed with TDI score) at 24 mo. Improvement in adjusted FVC% predicted over 24 mo was observed among patients with dcSSc and early disease duration (<24 mo). There was no difference between MMF and cyclophosphamide in improvement in FVC% predicted.	Anemia, leukopenia, infection; more favorable side effect profile compared with cyclophosphamide	Recommended as a first-line therapy in the ATS and BSR guidelines and as one of the preferred first-line options in the ACR/CHEST guideline. MMF is better tolerated than azathioprine. Mycophenolic acid can be considered for those intolerant to MMF.
RTX	1,000 mg intravenously repeated twice at an interval of 2 wk; regimen can be repeated every 6 mo if necessary	CD20 inhibitor resulting in B cell depletion; reduction in circulating IL-6 levels was reported at 24 wk post RTX treatment	At 24–48 wk, compared with placebo, RTX reduced the decline in FVC% predicted by 3.3%. Upfront combination MMF and RTX improves FVC over 52 wk.	Infection, including hepatitis B reactivation in particular, in combination with MMF; cytopenia	Conditionally recommended as a first-line therapy or for those who progressed despite first-line therapy in the ACR/CHEST guideline. Numerically fewer adverse events were reported among patients taking RTX compared to cyclophosphamide in the RECITAL study.

(Continued)

Table 2. (*Cont'd*)

Drug name	Dose	Immunologic basis for action	Clinical evidence for efficacy	Adverse effects	Comments
NTD	100–150 mg twice daily	As a tyrosine kinase inhibitor that attenuates PDGF, fibroblast-derived growth factor, and vascular endothelial growth factor receptor signaling in fibrogenesis. There is inhibition of the profibrotic M2 phenotype via reduction of gene expression and protein synthesis of M2 cell surface markers (including CD204, CD206, and CD163) with reduction of TGFβ1 expression. NTD inhibits the mTOR-dependent signals, leading to the expression of cytokines and matrix metalloproteinase 7 as well as B cell proliferation and IgM secretion mediators.	Compared to placebo, the adjusted annual rate of FVC decline was 40.9 mL less over 52 wk. NTD in combination with MMF provided greater numerical preservation of lung function than MMF with placebo, but this was not statistically significant.	Increased nausea, vomiting, diarrhea, weight loss, increased risk of CV events, potential increased risk of bleeding	Conditionally recommended as one of the first-line options and when the patient meets the criteria for progression in the ACR/CHEST guideline. The ATS guideline broadens the recommendation beyond patients with progressive lung disease as well as in combination with MMF. Combination therapy with mycophenolate was not randomized in SENSCIS, and those patients receiving background mycophenolate had several differences in demographics compared with patients taking NTD alone.
Pirfenidone	Days 1–7: 267 mg 3 times per day; days 8–4: 534 mg 3 times per day; day 15 and after: 801 mg 3 times per day	Its precise mechanism of action is unclear: as an oral pyridine derivative that regulates cytokines, including TGFβ and TNF, it has antifibrotic, anti-inflammatory, and antioxidant properties. In contrast to NTD, pirfenidone inhibits mTOR-independent pathways in its effects on microbial antigens on B cells and their interaction with lung fibroblasts.	No statistically significant difference between treatment groups was demonstrated in the clinical studies. Small sample size did not permit subgroup analysis for the SSc-ILD cohort.	Gastrointestinal disturbances (diarrhea, nausea, abdominal pain, weight loss), skin rash, photosensitivity	This was not recommended in current guidelines. The ATS guideline advises further research into pirfenidone alone or in combination with MMF.

(Continued)

Table 2. (*Cont'd*)

Drug name	Dose	Immunologic basis for action	Clinical evidence for efficacy	Adverse effects	Comments
Tocilizumab	162 mg subcutaneously weekly	Monoclonal antibody that acts as an IL-6R antagonist, thus inhibiting IL-6 activity. Reduced levels of the circulating M2-macrophage chemokine CCL18 was reported with tocilizumab. It indirectly and directly targets B cells in SSc via inhibition of IL-6 function. There are no studies that evaluate the effect of tocilizumab specifically on the antinuclear B cell response in SSc.	The difference in mean change from baseline to 48 wk in FVC% predicted was 6.5% less in the tocilizumab cohort.	Increased infections, hyperlipidemia, abnormal liver function	The ACR/CHEST guideline recommends this as one of the first-line options as well as for those who progressed. Due to its shared clinical phenotype, MCTD was included in the ACR/CHEST recommendation. The ATS has a conditional recommendation. Both the ACR/CHEST and the BSR recommend this as a first-line treatment in early dcSSc with raised markers of inflammation and ATA positivity, independent of the extent of ILD on CT. None of the guidelines recommends combination therapy with MMF.

* ACR, American College of Rheumatology; ATA, anti–topoisomerase I antibody; ATS, American Thoracic Society; BSR, British Society for Rheumatology; CHEST, American College of Chest Physicians; CT, computed tomography; CV, cardiovascular; dcSSc, diffuse cutaneous systemic sclerosis; FVC, force vital capacity; IL, interleukin; ILD, interstitial lung disease; MCTD, mixed connective tissue disease; MMF, mycophenolate mofetil; mTOR, mechanistic target of rapamycin; NTD, nintedanib; PDGF, platelet-derived growth factor; RTX, rituximab; SSc, systemic sclerosis; TDI, transition dyspnea index; TGF, transforming growth factor; TNF, tumor necrosis factor.

Table 3. B cell therapies in SSc*

Target	Agent	Study design	Outcome	Reference	ClinicalTrials.gov identifier
Completed B cell–directed therapies in SSc					
CD20	RTX	RCT (phase 2/3)	RTX improved mRSS and lung function in SSc	Ebata et al[79]	NCT04274257
CD20	RTX	RCT (phase 2)	No statistical improvement but potential treatment for SSc-PAH	Zamanian et al[88]	NCT01086540
CD20	RTX	RCT (phase 2/3)	RTX increased FVC at 24 wk for CTD-ILD, including SSc, but was not superior to cyclophosphamide. Improvement in mRSS. Steroid-sparing effect was noted.	Maher et al[78]	NCT0182926
CD20	RTX	RCT (phase 3)	Upfront RTX and MMF combination stabilized lung function at 6 mo	Mankikian et al[81]	NCT02990286
CD19	Inebilizumab	RCT (phase 1)	Single escalating dose of MEDI-551 was well tolerated and safe	Schiopu et al[89]	NCT00946699
BAFF	Belimumab	RCT (phase 2)	No significant effect on mRSS at 52 wk with belimumab–MMF compared to MMF alone	Gordon et al[90]	NCT01670565
Enrolling trials investigating B cell–targeted therapies for SSc					
BAFF and CD20	Belimumab and RTX	RCT (phase 2)	Ongoing; primary outcome evaluating change in the ACR rCRISS at 12 mo	–	NCT03844061
BAFF	Belimumab with standard therapy	RCT (Bliss-ILD phase 2/3)	Ongoing; primary outcome on lung function at week 52	–	NCT05878717
BAFF and IL-17	Tibulizumab	RCT (TibuSURE phase 2)	Not recruiting yet; primary outcome on skin at week 24	–	NCT06843239
BAFF receptor	Ianalumab	RCT (phase 2)	Recruiting; primary outcome evaluating rCRISS response at week 52	–	NCT06470048
CD20	Divozilimab	RCT (phase 3)	Not recruiting; primary outcome on skin at week 24	–	NCT05726630

* ACR, American College of Rheumatology; CTD, connective tissue disease; FVC, force vital capacity; IL, interleukin; ILD, interstitial lung disease; MMF, mycophenolate mofetil; mRSS, modified Rodnan skin score; PAH, pulmonary arterial hypertension; rCRISS, revised Composite Response Index in Systemic Sclerosis; RCT, randomized controlled trial; RTX, rituximab; SSc, systemic sclerosis.

have shown improvement in FVC (median 195 mL) and a 4% median reduction in ILD extent on HRCT, in addition to improvement in the mRSS.[97] Using the CRISPR/Cas9 strategy with a healthy donor–derived allogeneic CD19 CAR T cell product in two patients with dcSSc, improvement in ILD without graft-versus-host disease–related adverse effects was recently reported.[98] Bispecific autoantigen T cell–engaging antibodies are a new class of "off-the-shelf" B cell–depletion approaches that do not require personalized manufacturing. Blinatumomab was reported to improve the mRSS and result in profound B cell depletion in a patient with progressive dcSSc.[99] Similar preclinical progress on engineering CARs for expression on Treg cells to recognize particular self-antigens is under evaluation. It is envisaged that the pace of innovation in CAR T cell development for autoantibody-mediated diseases, including SSc, will continue to accelerate (ClinicalTrials.gov identifiers NCT06400303, NCT06347718, and NCT06328777).

As discussed previously, data regarding BAFF in SSc are mixed. Elevated BAFF levels have been found following treatment with rituximab in SSc, and BAFF levels have been associated with

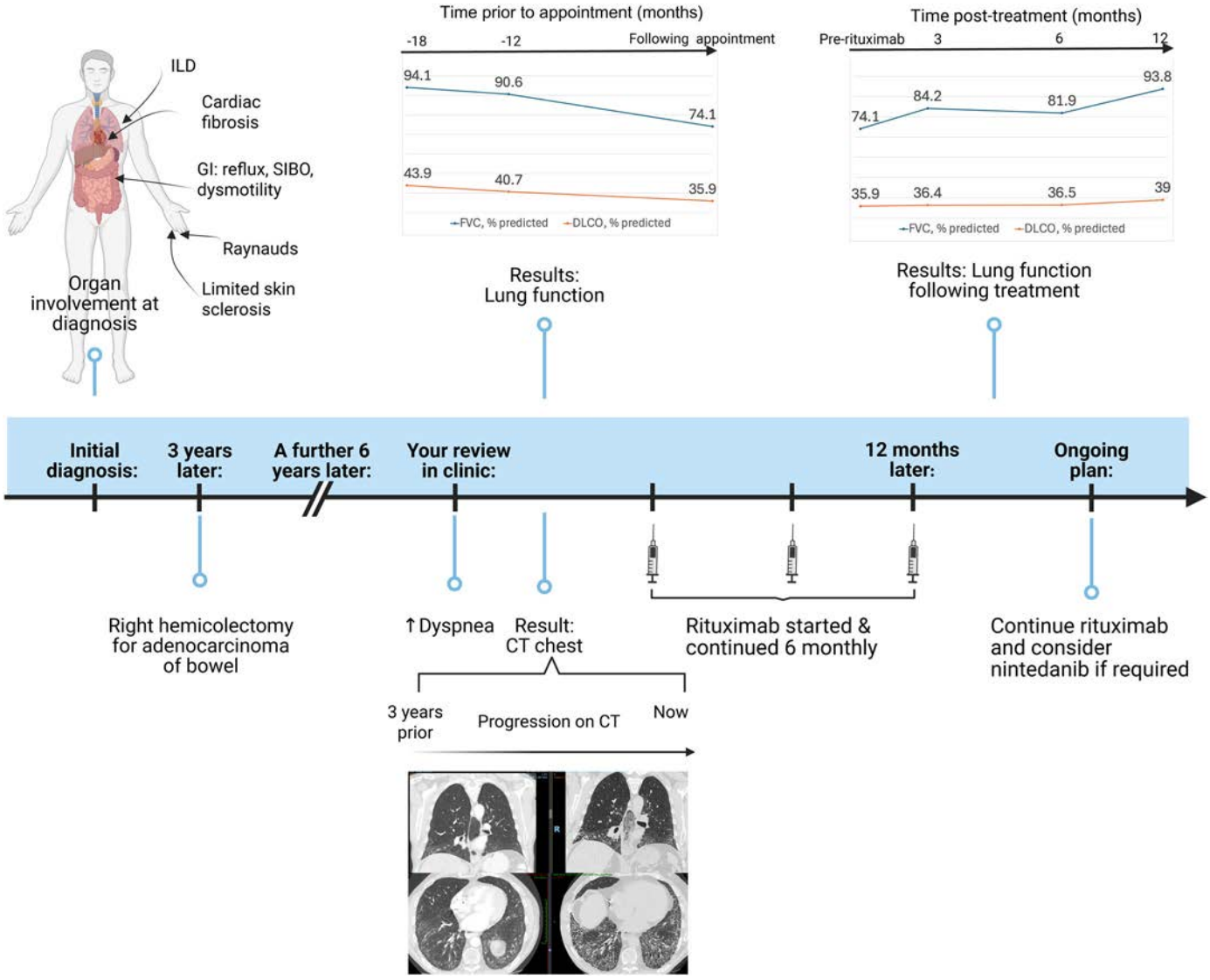


Figure 3. A pictorial summary of the case. CT, computed tomography; DLCO, diffusing capacity for carbon monoxide; FVC, forced vital capacity; GI, gastrointestinal; ILD, interstitial lung disease; SIBO, small instestinal bacterial overgrowth. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.43326/abstract.

flares following rituximab treatment in systemic lupus erythematosus.[73,100] A small trial of belimumab, a monoclonal antibody targeting BAFF, demonstrated no significant improvement in skin when administered in addition to MMF for early dcSSc, and a phase 2/3 trial of belimumab for SSc-ILD is ongoing.[90,101] Ianalumab is an anti-BAFF receptor monoclonal antibody that has a dual mechanism of action with blockade of BAFF receptor–mediated signaling and B cell depletion by enhanced antibody-dependent cellular cytotoxicity.[102] A phase 2 clinical trial of ianalumab in early dcSSc has recently begun[103] (Table 3).

Although IL-6 is produced by a range of cells, in SSc, it has been suggested[62] that B cells may be the predominant source of increased levels of IL-6. Rituximab has been demonstrated to reduce IL-6 levels in SSc, and the interaction of tocilizumab with B cells is likely important in the mechanism of action of this treatment.[73]

Infection risk with immunosuppression

Concerns regarding the use of rituximab and other biologics, especially in combination with other immunosuppressants, often revolve around risk of infection. Infection risk may also increase over time with hypogammaglobulinemia from B cell depletion antibody therapy.[104,105]

In the RECITAL trial, there was reduced steroid exposure in patients with CTD-ILD taking rituximab compared to cyclophosphamide, and infection rates were similar across groups; however, rituximab was not used in combination with other immunosuppression.[78] In the EVER-ILD trial, in which rituximab was given in combination with 2 g daily of MMF, infection rates were higher in the combination therapy group compared to those receiving MMF alone; however, increased infections were predominantly nonserious viral infections.[81]

An increased risk of severe infection, hospitalization, and death from COVID-19 in patients receiving rituximab for

immune-mediated disease compared to the general population has been found.[106] It has also been shown that there is an increased risk of COVID-19 infection with rituximab in patients with inflammatory rheumatic diseases receiving rituximab compared to those receiving other disease-modifying antirheumatic drugs and nonrituximab biologic agents.[106,107] This risk is of greater concern with additional risk factors for severe COVID-19 infection present, such as ILD.[108]

Rituximab treatment also results in a reduction in the humoral response to the COVID-19 vaccine, although the functional T cell response is unaltered to the messenger RNA vaccine.[109] A single-center study of patients with autoimmune disease treated with rituximab in the United Kingdom found that full vaccination significantly reduced rates of COVID-19 infection, and although breakthrough infection rates were high after vaccination, most cases were mild in fully vaccinated individuals).[110] Hypogammaglobulinemia before rituximab treatment was found to be associated with moderate and severe COVID-19 infection. However, of note in this study, 74% (261 of 352) received their first vaccine dose more than six months after rituximab treatment. Current ACR guidelines recommend that COVID-19 vaccination timing be carefully considered in relation to rituximab dosing, either using CD19 B cell measurements to time vaccination or by giving booster doses two to four weeks before the next anticipated dose.[111]

Due to the B cell–targeting action of rituximab, there is no evidence of increased risk of tuberculosis with rituximab.[112] Risk of hepatitis B reactivation with rituximab treatment is clear, and screening is recommended before rituximab initiation.[113] *Pneumocystis jirovecii* pneumonia (PJP) prevention is also recommended for patients with ILD taking rituximab.

Case follow-up

Following vaccination against COVID-19 and after commencing PJP prophylaxis, the patient was treated with rituximab (two doses of 1 g with dosing interval of two weeks) and continued with MMF 2 g daily. He was given further rituximab treatment at the same dose six months later. No infections requiring antibiotic therapy occurred in the first year following rituximab initiation. A repeat lung function test was performed for monitoring after rituximab treatment, and results are shown in Supplemental Table 1. A decision was made to continue rituximab therapy every six months and keep nintedanib in reserve if SSc-ILD progresses further (Figure 3).

Rituximab treatment appears to have stabilized SSc-ILD in this patient with long-standing SSc and normal markers of inflammation. This case highlights the diverse range of patients with progressive SSc-ILD, for whom additional biologic therapy can be considered, as well as the need to be cognizant of non-ATA high-risk SSc-ILD antibodies. Although continued rituximab treatment is planned, the durability of long-term response beyond the observed benefit with initial treatment is also unknown.

Conclusions

Future studies will determine whether CD19-targeting CAR T cells or the CD20 antibody obinutuzumab may be more effective or have an unacceptable infection risk compared to rituximab. The clear message is that B cells may have a key role in SSc-ILD, and further development of both cell- and antibody-based B cell–depleting therapies are justified.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr Denton confirms that all authors have provided the final approval of the version to be published and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. LeRoy EC, Black C, Fleischmajer R, et al. Scleroderma (systemic sclerosis): classification, subsets and pathogenesis. J Rheumatol 1988;15(2):202–205.
2. Elhai M, Meune C, Boubaya M, et al; EUSTAR group. Mapping and predicting mortality from systemic sclerosis. Ann Rheum Dis 2017; 76(11):1897–1905.
3. Steen VD, Medsger TA. Changes in causes of death in systemic sclerosis, 1972-2002. Ann Rheum Dis 2007;66(7):940–944.
4. Harrison NK, Myers AR, Corrin B, et al. Structural features of interstitial lung disease in systemic sclerosis. Am Rev Respir Dis 1991; 144(3 Pt 1):706–713.
5. Truchetet ME, Brembilla NC, Chizzolini C. Current concepts on the pathogenesis of systemic sclerosis. Clin Rev Allergy Immunol 2023; 64(3):262–283.
6. Liu X, Dai K, Zhang X, et al. Multiple fibroblast subtypes contribute to matrix deposition in pulmonary fibrosis. Am J Respir Cell Mol Biol 2023;69(1):45–56.
7. Fanning GC, Welsh KI, Bunn C, et al. HLA associations in three mutually exclusive autoantibody subgroups in UK systemic sclerosis patients. Br J Rheumatol 1998;37(2):201–207.
8. Nihtyanova SI, Sari A, Harvey JC, et al. Using autoantibodies and cutaneous subset to develop outcome-based disease classification in systemic sclerosis. Arthritis Rheumatol 2020;72(3):465–476.
9. Perelas A, Silver RM, Arrossi AV, et al. Systemic sclerosis-associated interstitial lung disease. Lancet Respir Med 2020;8(3):304–320.
10. De Lauretis A, Sestini P, Pantelidis P, et al. Serum interleukin 6 is predictive of early functional decline and mortality in interstitial lung disease associated with systemic sclerosis. J Rheumatol 2013;40(4): 435–446.
11. Steen VD. Autoantibodies in systemic sclerosis. Semin Arthritis Rheum 2005;35(1):35–42.
12. Acosta-Herrera M, Kerick M, Lopéz-Isac E, et al; International SSc Group; Australian Scleroderma Interest Group (ASIG).

Comprehensive analysis of the major histocompatibility complex in systemic sclerosis identifies differential HLA associations by clinical and serological subtypes. Ann Rheum Dis 2021;80(8):1040–1047.

13. Nihtyanova SI, Denton CP. Autoantibodies as predictive tools in systemic sclerosis. Nat Rev Rheumatol 2010;6(2):112–116.
14. Yamashita Y, Yamano Y, Muro Y, et al. Clinical significance of anti-NOR90 antibodies in systemic sclerosis and idiopathic interstitial pneumonia. Rheumatology (Oxford) 2022;61(4):1709–1716.
15. Fritzler MJ, Hudson M, Choi MY, et al; Canadian Scleroderma Research Group. Bicaudal D2 is a novel autoantibody target in systemic sclerosis that shares a key epitope with CENP-A but has a distinct clinical phenotype. Autoimmun Rev 2018;17(3):267–275.
16. Adler BL, Boin F, Wolters PJ, et al. Autoantibodies targeting telomere-associated proteins in systemic sclerosis. Ann Rheum Dis 2021;80(7):912–919.
17. Kuwana M, Gil-Vila A, Selva-O'Callaghan A. Role of autoantibodies in the diagnosis and prognosis of interstitial lung disease in autoimmune rheumatic disorders. Ther Adv Musculoskelet Dis 2021;13:1759720X211032457.
18. Johnson SR, Bernstein EJ, Bolster MB, et al. 2023 American College of Rheumatology (ACR)/American College of Chest Physicians (CHEST) guideline for the screening and monitoring of interstitial lung disease in people with systemic autoimmune rheumatic diseases. Arthritis Rheumatol 2024;76(8):1201–1213.
19. Fertig N, Domsic RT, Rodriguez-Reyna T, et al. Anti-U11/U12 RNP antibodies in systemic sclerosis: a new serologic marker associated with pulmonary fibrosis. Arthritis Rheum 2009;61(7):958–965.
20. Fritzler MJ, Bentow C, Beretta L, et al. Anti-U11/U12 antibodies as a rare but important biomarker in patients with systemic sclerosis: a narrative review. Diagnostics (Basel) 2023;13(7):1257.
21. Callejas-Moraga EL, Guillén-Del-Castillo A, Perurena-Prieto J, et al. Anti-RNPC-3 antibody predicts poor prognosis in patients with interstitial lung disease associated to systemic sclerosis. Rheumatology (Oxford) 2021;61(1):154–162.
22. Shah AA, Xu G, Rosen A, et al. Brief report: anti-RNPC-3 antibodies as a marker of cancer-associated scleroderma. Arthritis Rheumatol 2017;69(6):1306–1312.
23. Will CL, Schneider C, Hossbach M, et al. The human 18S U11/U12 snRNP contains a set of novel proteins not found in the U2-dependent spliceosome. RNA 2004;10(6):929–941.
24. Gilliam AC, Steitz JA. Rare scleroderma autoantibodies to the U11 small nuclear ribonucleoprotein and to the trimethylguanosine cap of U small nuclear RNAs. Proc Natl Acad Sci USA 1993;90(14):6781–6785.
25. McMahan ZH, Domsic RT, Zhu L, et al. Anti-RNPC-3 (U11/U12) antibodies in systemic sclerosis in patients with moderate-to-severe gastrointestinal dysmotility. Arthritis Care Res (Hoboken) 2019;71(9):1164–1170.
26. Mahler M, Roup F, Bentow C, et al. Anti-RNPC-3 antibodies are associated with nuclear speckled immunofluorescence pattern and enriched in triple negative systemic sclerosis patients [abstract]. Arthritis Rheumatol 2019;71(suppl 10). https://acrabstracts.org/abstract/anti-rnpc-3-antibodies-are-associated-with-nuclear-speckled-immunofluorescence-pattern-and-enriched-in-triple-negative-systemic-sclerosis-patients/
27. Zhan X, Koelsch T, Montner SM, et al. Differentiating usual interstitial pneumonia from nonspecific interstitial pneumonia using high-resolution computed tomography: the "straight-edge sign". J Thorac Imaging 2018;33(4):266–270.
28. Egashira R, Jacob J, Kokosi MA, et al. Diffuse pulmonary ossification in fibrosing interstitial lung diseases: prevalence and associations. Radiology 2017;284(1):255–263.
29. Fleischmajer R, Perlish JS, Reeves JRT. Cellular infiltrates in scleroderma skin. Arthritis Rheum 1977;20(4):975–984.
30. Hu PQ, Fertig N, Medsger TA Jr, et al. Correlation of serum anti-DNA topoisomerase I antibody levels with disease severity and activity in systemic sclerosis. Arthritis Rheum 2003;48(5):1363–1373.
31. Tramposch HD, Smith CD, Senecal JL, et al. A long-term longitudinal study of anticentromere antibodies. Arthritis Rheum 1984;27(2):121–124.
32. Burbelo PD, Gordon SM, Waldman M, et al. Autoantibodies are present before the clinical diagnosis of systemic sclerosis. PLoS One 2019;14(3):e0214202.
33. Minier T, Guiducci S, Bellando-Randone S, et al; EUSTAR coworkers; EUSTAR co-workers. Preliminary analysis of the very early diagnosis of systemic sclerosis (VEDOSS) EUSTAR multicentre study: evidence for puffy fingers as a pivotal sign for suspicion of systemic sclerosis. Ann Rheum Dis 2014;73(12):2087–2093.
34. Fineschi S, Goffin L, Rezzonico R, et al. Antifibroblast antibodies in systemic sclerosis induce fibroblasts to produce profibrotic chemokines, with partial exploitation of toll-like receptor 4. Arthritis Rheum 2008;58(12):3913–3923.
35. Kill A, Tabeling C, Undeutsch R, et al. Autoantibodies to angiotensin and endothelin receptors in systemic sclerosis induce cellular and systemic events associated with disease pathogenesis. Arthritis Res Ther 2014;16(1):R29.
36. Akbarzadeh R, Müller A, Humrich JY, et al. When natural antibodies become pathogenic: autoantibodies targeted against G protein-coupled receptors in the pathogenesis of systemic sclerosis. Front Immunol 2023;14:1213804.
37. Rizou C, Ioannidis JPA, Panou-Pomonis E, et al. B-cell epitope mapping of DNA topoisomerase I defines epitopes strongly associated with pulmonary fibrosis in systemic sclerosis. Am J Respir Cell Mol Biol 2000;22(3):344–351.
38. Liem SIE, Neppelenbroek S, Fehres CM, et al. Autoreactive B cell responses targeting nuclear antigens in systemic sclerosis: implications for disease pathogenesis. Semin Arthritis Rheum 2023;58:152136.
39. Yoshizaki A, Yanaba K, Ogawa A, et al. Immunization with DNA topoisomerase I and Freund's complete adjuvant induces skin and lung fibrosis and autoimmunity via interleukin-6 signaling. Arthritis Rheum 2011;63(11):3575–3585.
40. Senécal JL, Hoa S, Yang R, et al. Pathogenic roles of autoantibodies in systemic sclerosis: current understandings in pathogenesis. J Scleroderma Relat Disord 2020;5(2):103–129.
41. Fukasawa T, Yoshizaki A, Ebata S, et al. Single-cell-level protein analysis revealing the roles of autoantigen-reactive B lymphocytes in autoimmune disease and the murine model. eLife 2021;10:e67209.
42. Boonstra M, Bakker JA, Grummels A, et al. Association of anti-topoisomerase I antibodies of the IgM isotype with disease progression in anti-topoisomerase I-positive systemic sclerosis. Arthritis Rheumatol 2020;72(11):1897–1904.
43. Farge D, Henegar C, Carmagnat M, et al. Analysis of immune reconstitution after autologous bone marrow transplantation in systemic sclerosis. Arthritis Rheum 2005;52(5):1555–1563.
44. Henes J, Glaeser L, Kötter I, et al. Analysis of anti-topoisomerase I antibodies in patients with systemic sclerosis before and after autologous stem cell transplantation. Rheumatology (Oxford) 2017;56(3):451–456.
45. Bonroy C, Smith V, Deschepper E, et al. Specific antinuclear antibody level changes after B cell depletion therapy in systemic sclerosis are associated with improvement of skin thickening. J Rheumatol 2016;43(3):681.

46. Lafyatis R, Kissin E, York M, et al. B cell depletion with rituximab in patients with diffuse cutaneous systemic sclerosis. Arthritis Rheum 2009;60(2):578–583.
47. Whitfield ML, Finlay DR, Murray JI, et al. Systemic and cell type-specific gene expression patterns in scleroderma skin. Proc Natl Acad Sci USA 2003;100(21):12319–12324.
48. Skaug B, Khanna D, Swindell WR, et al. Global skin gene expression analysis of early diffuse cutaneous systemic sclerosis shows a prominent innate and adaptive inflammatory profile. Ann Rheum Dis 2020; 79(3):379–386.
49. Lafyatis R, O'Hara C, Feghali-Bostwick CA, et al. B cell infiltration in systemic sclerosis-associated interstitial lung disease. Arthritis Rheum 2007;56(9):3167–3168.
50. Arvelo Castro L, Van Keulen V, Ali M, et al. B-cell infiltration varies among patients with interstitial lung diseases [abstract]. Eur Respir J 2016;48(suppl 60):PA3099.
51. Popper H, Stacher-Priehse E, Brcic L, et al. Lung fibrosis in autoimmune diseases and hypersensitivity: how to separate these from idiopathic pulmonary fibrosis. Rheumatol Int 2022;42(8):1321–1330.
52. De Santis M, Bosello SL, Peluso G, et al. Bronchoalveolar lavage fluid and progression of scleroderma interstitial lung disease. Clin Respir J 2012;6(1):9–17.
53. Goh NSL, Veeraraghavan S, Desai SR, et al. Bronchoalveolar lavage cellular profiles in patients with systemic sclerosis-associated interstitial lung disease are not predictive of disease progression. Arthritis Rheum 2007;56(6):2005–2012.
54. Sato S, Fujimoto M, Hasegawa M, et al. Altered blood B lymphocyte homeostasis in systemic sclerosis: expanded naive B cells and diminished but activated memory B cells. Arthritis Rheum 2004; 50(6):1918–1927.
55. Forestier A, Guerrier T, Jouvray M, et al. Altered B lymphocyte homeostasis and functions in systemic sclerosis. Autoimmun Rev 2018;17(3):244–255.
56. Glauzy S, Olson B, May CK, et al. Defective early B cell tolerance checkpoints in patients with systemic sclerosis allow the production of self antigen-specific clones. Arthritis Rheumatol 2022;74(2): 307–317.
57. Soto L, Ferrier A, Aravena O, et al. Systemic sclerosis patients present alterations in the expression of molecules involved in B-cell regulation. Front Immunol 2015;6:496.
58. Sato S, Hasegawa M, Fujimoto M, et al. Quantitative genetic variation in CD19 expression correlates with autoimmunity. J Immunol 2000;165(11):6635–6643.
59. Melissaropoulos K, Liossis SN. Decreased CD22 expression and intracellular signaling aberrations in B cells of patients with systemic sclerosis. Rheumatol Int 2018;38(7):1225–1234.
60. Matsushita T, Hasegawa M, Yanaba K, et al. Elevated serum BAFF levels in patients with systemic sclerosis: enhanced BAFF signaling in systemic sclerosis B lymphocytes. Arthritis Rheum 2006;54(1): 192–201.
61. Matsushita T, Fujimoto M, Hasegawa M, et al. Elevated serum APRIL levels in patients with systemic sclerosis: distinct profiles of systemic sclerosis categorized by APRIL and BAFF. J Rheumatol 2007; 34(10):2056–2062.
62. Dumoitier N, Chaigne B, Régent A, et al. Scleroderma peripheral B lymphocytes secrete interleukin-6 and transforming growth factor β and activate fibroblasts. Arthritis Rheumatol 2017;69(5):1078–1089.
63. Taher TE, Ong VH, Bystrom J, et al. Association of defective regulation of autoreactive interleukin-6-producing transitional B lymphocytes with disease in patients with systemic sclerosis. Arthritis Rheumatol 2018;70(3):450–461.
64. Matsushita T, Hamaguchi Y, Hasegawa M, et al. Decreased levels of regulatory B cells in patients with systemic sclerosis: association with autoantibody production and disease activity. Rheumatology (Oxford) 2016;55(2):263–267.
65. Kuwana M, Medsger TA Jr, Wright TMT. T and B cell collaboration is essential for the autoantibody response to DNA topoisomerase I in systemic sclerosis. J Immunol 1995;155(5):2703–2714.
66. Ricard L, Jachiet V, Malard F, et al. Circulating follicular helper T cells are increased in systemic sclerosis and promote plasmablast differentiation through the IL-21 pathway which can be inhibited by ruxolitinib. Ann Rheum Dis 2019;78(4):539–550.
67. Numajiri H, Kuzumi A, Fukasawa T, et al. B cell depletion inhibits fibrosis via suppression of profibrotic macrophage differentiation in a mouse model of systemic sclerosis. Arthritis Rheumatol 2021; 73(11):2086–2095.
68. Raghu G, Montesi SB, Silver RM, et al. Treatment of systemic sclerosis-associated interstitial lung disease: evidence-based recommendations. an official American Thoracic Society clinical practice guideline. Am J Respir Crit Care Med 2024;209(2):137–152.
69. Johnson SR, Bernstein EJ, Bolster MB, et al. 2023 American College of Rheumatology (ACR)/American College of Chest Physicians (CHEST) guideline for the treatment of interstitial lung disease in people with systemic autoimmune rheumatic diseases. Arthritis Rheumatol 2024;76(8):1182–1200.
70. Denton CP, De Lorenzis E, Roblin E, et al. The 2024 British Society for Rheumatology guideline for management of systemic sclerosis. Rheumatology (Oxford) 2024;63(11):2956–2975.
71. Ali MF, Egan AM, Shaughnessy GF, et al. Antifibrotics modify B-cell-induced fibroblast migration and activation in patients with idiopathic pulmonary fibrosis. Am J Respir Cell Mol Biol 2021;64(6):722–733.
72. Morris EA, Parvizi R, Orzechowski NM, et al. Mycophenolate mofetil directly modulates myeloid viability and pro-fibrotic activation of human macrophages. Rheumatology (Oxford) 2025;64(5):3125–3133.
73. Bosello S, De Santis M, Lama G, et al. B cell depletion in diffuse progressive systemic sclerosis: safety, skin score modification and IL-6 modulation in an up to thirty-six months follow-up open-label trial. Arthritis Res Ther 2010;12(2):R54.
74. Soldano S, Smith V, Montagna P, et al. Nintedanib downregulates the profibrotic M2 phenotype in cultured monocyte-derived macrophages obtained from systemic sclerosis patients affected by interstitial lung disease. Arthritis Res Ther 2024;26(1):74.
75. Khanna D, Denton CP, Jahreis A, et al. Safety and efficacy of subcutaneous tocilizumab in adults with systemic sclerosis (faSScinate): a phase 2, randomised, controlled trial. Lancet 2016;387(10038): 2630–2640.
76. Tashkin DP, Roth MD, Clements PJ, et al; Sclerodema Lung Study II Investigators. Mycophenolate mofetil versus oral cyclophosphamide in scleroderma-related interstitial lung disease (SLS II): a randomised controlled, double-blind, parallel group trial. Lancet Respir Med 2016;4(9):708–719.
77. Khanna D, Lin CJF, Furst DE, et al; focuSSced investigators. Tocilizumab in systemic sclerosis: a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet Respir Med 2020;8(10):963–974.
78. Maher TM, Tudor VA, Saunders P, et al; RECITAL Investigators. Rituximab versus intravenous cyclophosphamide in patients with connective tissue disease-associated interstitial lung disease in the UK (RECITAL): a double-blind, double-dummy, randomised, controlled, phase 2b trial. Lancet Respir Med 2023;11(1):45–54.
79. Ebata S, Yoshizaki A, Oba K, et al. Safety and efficacy of rituximab in systemic sclerosis (DESIRES): a double-blind, investigator-initiated, randomised, placebo-controlled trial. Lancet Rheumatol 2021;3(7): e489–e497.
80. Ebata S, Yoshizaki A, Oba K, et al. Safety and efficacy of rituximab in systemic sclerosis (DESIRES): open-label extension of a double-

blind, investigators-initiated, randomised, placebo-controlled trial. Lancet Rheumatol 2022;4(8):e546–e555.
81. Mankikian J, Caille A, Reynaud-Gaubert M, et al; EVER-ILD investigators and the OrphaLung network. Rituximab and mycophenolate mofetil combination in patients with interstitial lung disease (EVER-ILD): a double-blind, randomised, placebo-controlled trial. Eur Respir J 2023;61(6):2202071.
82. Mansy L, Caille A, Reynaud-Gaubert M, et al; EVER-ILD investigators and the OrphaLung network. Rituximab and mycophenolate mofetil in interstitial lung disease (EVER-ILD): 1-year follow-up results of a randomised controlled trial. Eur Respir J 2024;64(3):2401368.
83. Goldman NR, Nihtyanova SI, Beesley CF, et al. Tocilizumab and rituximab for systemic sclerosis interstitial lung disease: a real-world cohort analysis. Rheumatology (Oxford) Published online January 3, 2025. https://doi.org/10.1093/rheumatology/keaf006
84. Bruera S, Sidanmat H, Molony DA, et al. Stem cell transplantation for systemic sclerosis. Cochrane Database Syst Rev. 2022;(7): CD011819.
85. Spierings J, van Rhenen A, Welsing PMW, et al. A randomised, open-label trial to assess the optimal treatment strategy in early diffuse cutaneous systemic sclerosis: the UPSIDE study protocol. BMJ Open 2021;11(3):e044483.
86. Flaherty KR, Wells AU, Cottin V, et al; INBUILD Trial Investigators. Nintedanib in progressive fibrosing interstitial lung diseases. N Engl J Med 2019;381(18):1718–1727.
87. Distler O, Highland KB, Gahlemann M, et al; SENSCIS Trial Investigators. Nintedanib for systemic sclerosis-associated interstitial lung disease. N Engl J Med 2019;380(26):2518–2528.
88. Zamanian RT, Badesch D, Chung L, et al. Safety and efficacy of B-cell depletion with rituximab for the treatment of systemic sclerosis-associated pulmonary arterial hypertension: a multicenter, double-blind, randomized, placebo-controlled trial. Am J Respir Crit Care Med 2021;204(2):209–221.
89. Schiopu E, Chatterjee S, Hsu V, et al. Safety and tolerability of an anti-CD19 monoclonal antibody, MEDI-551, in subjects with systemic sclerosis: a phase I, randomized, placebo-controlled, escalating single-dose study. Arthritis Res Ther 2016;18(1):131.
90. Gordon JK, Martyanov V, Franks JM, et al. Belimumab for the treatment of early diffuse systemic sclerosis: results of a randomized, double-blind, placebo-controlled, pilot trial. Arthritis Rheumatol 2018;70(2):308–316.
91. Ebata S, Yoshizaki A, Fukasawa T, et al. Percentage of residual B cells after 2 weeks of rituximab treatment predicts the improvement of systemic sclerosis-associated interstitial lung disease. J Dermatol 2022;49(1):179–183.
92. Al Tabaa O, Ghossan R, Combier A, et al. Sustained complete B cell depletion is associated with rituximab efficacy in connective tissue disorders-associated interstitial lung disease. RMD Open 2023;9(1): e002832.
93. Hofmann K, Clauder AK, Manz RA. Targeting B cells and plasma cells in autoimmune diseases Front Immunol 2018;9:835.
94. Lima-Júnior JR, Arruda LCM, Gonçalves MS, et al. Autologous haematopoietic stem cell transplantation restores the suppressive capacity of regulatory B cells in systemic sclerosis patients. Rheumatology (Oxford) 2021;60(12):5538–5548.
95. Bergmann C, Müller F, Distler JHW, et al. Treatment of a patient with severe systemic sclerosis (SSc) using CD19-targeted CAR T cells. Ann Rheum Dis 2023;82(8):1117–1120.
96. Merkt W, Freitag M, Claus M, et al. Third-generation CD19.CAR-T cell-containing combination therapy in Scl70+ systemic sclerosis. Ann Rheum Dis 2024;83(4):543–546.
97. Auth J, Müller F, Völkl S, et al. CD19-targeting CAR T-cell therapy in patients with diffuse systemic sclerosis: a case series. Lancet Rheumatol 2025;7(2):e83–e93.
98. Wang X, Wu X, Tan B, et al. Allogeneic CD19-targeted CAR-T therapy in patients with severe myositis and systemic sclerosis. Cell 2024;187(18):4890–4904.e9.
99. Subklewe M, Magno G, Gebhardt C, et al. Application of blinatumomab, a bispecific anti-CD3/CD19 T-cell engager, in treating severe systemic sclerosis: a case study. Eur J Cancer 2024;204:114071.
100. Carter LM, Isenberg DA, Ehrenstein MR. Elevated serum BAFF levels are associated with rising anti-double-stranded DNA antibody levels and disease flare following B cell depletion therapy in systemic lupus erythematosus. Arthritis Rheum 2013;65(10):2672–2679.
101. Denton CP, Spiera R, Jörg D, et al. Belimumab for the treatment of interstitial lung disease associated with systemic sclerosis: a phase 2/3, randomised, double-blind, placebo-controlled trial [abstract]. Ann Rheum Dis 2023;82(suppl 1):1668.
102. Dörner T, Posch MG, Li Y, et al. Treatment of primary Sjögren's syndrome with ianalumab (VAY736) targeting B cells by BAFF receptor blockade coupled with enhanced, antibody-dependent cellular cytotoxicity. Ann Rheum Dis 2019;78(5):641–647.
103. A clinical study to evaluate ianalumab in participants with diffuse cutaneous systemic sclerosis. ClinicalTrials.gov identifier: NCT06470048. Updated August 7, 2024. Accessed September 20, 2024. https://clinicaltrials.gov/study/NCT06470048
104. van Vollenhoven RF, Emery P, Bingham CO III, et al. Long-term safety of rituximab in rheumatoid arthritis: 9.5-year follow-up of the global clinical trial programme with a focus on adverse events of interest in RA patients. Ann Rheum Dis 2013;72(9):1496–1502.
105. Nie Y, Zhang N, Li J, et al. Hypogammaglobulinemia and infection events in patients with autoimmune diseases treated with rituximab: 10 years real-life experience. J Clin Immunol 2024;44(8):179.
106. Patel NJ, D'Silva KM, Hsu TYT, et al. Coronavirus disease 2019 outcomes among recipients of Anti-CD20 monoclonal antibodies for immune-mediated diseases: a comparative cohort study. ACR Open Rheumatol 2022;4(3):238–246.
107. Felten R, Duret PM, Bauer E, et al. B-cell targeted therapy is associated with severe COVID-19 among patients with inflammatory arthritides: a 1-year multicentre study in 1116 successive patients receiving intravenous biologics. Ann Rheum Dis 2022;81(1): 143–145.
108. Gao YD, Ding M, Dong X, et al. Risk factors for severe and critically ill COVID-19 patients: a review Allergy 2021;76(2):428–455.
109. Bitoun S, Henry J, Desjardins D, et al. Rituximab impairs B cell response but not T cell response to COVID-19 vaccine in autoimmune diseases. Arthritis Rheumatol 2022;74(6):927–933.
110. Md Yusof MY, Arnold J, Saleem B, et al. Breakthrough SARS-CoV-2 infections and prediction of moderate-to-severe outcomes during rituximab therapy in patients with rheumatic and musculoskeletal diseases in the UK: a single-centre cohort study. Lancet Rheumatol 2023;5(2):e88–e98.
111. Curtis JR, Johnson SR, Anthony DD, et al. American College of Rheumatology guidance for COVID-19 vaccination in patients with rheumatic and musculoskeletal diseases: version 5. Arthritis Rheumatol 2023;75(1):E1–E16.
112. Cantini F, Nannini C, Niccoli L, et al. Risk of tuberculosis reactivation in patients with rheumatoid arthritis, ankylosing spondylitis, and psoriatic arthritis receiving non-anti-TNF-targeted biologics. Mediators Inflamm 2017;2017:8909834.
113. Varley CD, Winthrop KL. Long-term safety of rituximab (risks of viral and opportunistic infections). Curr Rheumatol Rep 2021;23(9):74.

Arthritis & Rheumatology
Vol. 78, No. 3, March 2026, pp 582–591
DOI 10.1002/art.43447

American College of Rheumatology
Empowering Rheumatology Professionals

The Impact of Nonsteroidal Anti-Inflammatory Drugs on Radiographic Spinal Progression in Patients With Axial Spondyloarthritis: 10-Year Results From an Inception Cohort

Murat Torgutalp,[1,2,3,4] Valeria Rios Rodriguez,[1] Fabian Proft,[1] Mikhail Protopopov,[1] Judith Rademacher,[1,5] Hildrun Haibel,[1] Joachim Sieper,[1] Martin Rudwaleit,[6] and Denis Poddubnyy[1,7]

Objective. This study aims to investigate the impact of nonsteroidal anti-inflammatory drug (NSAID) intake on radiographic spinal progression in axial spondyloarthritis (axSpA), considering different NSAID types (COX-2 inhibitors [COX2i] and nonselective NSAIDs [ns-NSAIDs]) and disease subgroups (radiographic [r-axSpA] and nonradiographic [nr-axSpA]).

Methods. Leveraging data from the German Spondyloarthritis Inception Cohort (GESPIC), we conducted analyses on 252 patients with axSpA (139 with nr-axSpA and 113 with r-axSpA), who had minimum two sets of spinal radiographs. The outcome was progression in modified Stoke Ankylosing Spondylitis Spine Score (mSASSS) in two-year intervals. We fitted sequential conditional mean models by using generalized estimating equations and adjusting for longitudinal repeated measures of exposure and time-dependent confounders. We report β-coefficients with 95% confidence intervals (CIs) for outcomes reflecting the progression in mSASSS per 10-point increase in NSAID intake score.

Results. At baseline, 201 (80.0%) patients were under NSAID treatment, with 46 (18%) taking COX2i and 156 (62%) taking ns-NSAIDs, and mean total NSAID intake score was 38.3 ± 35.5. A 10-point increase in NSAID intake score was associated with retardation of radiographic progression (β = –0.052, 95% CI: –0.097 to –0.007), with this effect being most pronounced in patients with r-axSpA (β = –0.077, 95% CI: –0.152 to –0.003). COX2i showed a slightly lower point estimate (although not statistically significant) in progression compared to ns-NSAIDs among all patients with axSpA (β = –0.061 and –0.045, respectively).

Conclusion. Our findings suggest a beneficial effect of higher NSAID intake, particularly COX2i, on slowing radiographic progression in axSpA. These findings may help inform therapeutic strategies, particularly in r-axSpA, although further research is needed for nr-axSpA.

INTRODUCTION

Axial spondyloarthritis (axSpA) is a chronic inflammatory disease that primarily affects the axial skeleton, including the sacroiliac joints and spine.[1] This inflammation causes pain, stiffness, and structural damage in the affected joints. If left uncontrolled, inflammation may lead to radiographic progression in the spine characterized by new bone formation, which significantly contributes to functional limitations and impaired mobility.[2,3]

The German Spondyloarthritis Inception Cohort (GESPIC) was supported by the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung [BMBF]). Because funding by the BMBF was reduced according to schedule in 2005 and stopped in 2007, complementary financial support was also obtained from Abbott/AbbVie, Amgen, Centocor, Schering-Plough, and Wyeth. Since 2010, the GESPIC was supported by AbbVie and A Network on Clinics and Pathophysiology of Osteophytes and Ankylosis (grant FKZ 01EC1002D), ArthroMark (grants FKZ 01EC1009A and FKZ 01EC1401A), and Metabolic impact on joint and bone disease (grant FKZ 01EC1407A), which are projects funded by the BMBF.

[1]Department of Gastroenterology, Infectiology and Rheumatology (including Nutrition Medicine), Charité – Universitätsmedizin Berlin, Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany; [2]Department of Rheumatology and Clinical Immunology, Charité – Universitätsmedizin Berlin, Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany; [3]Fraunhofer Institute for Translational Medicine and Pharmacology, Berlin, Germany; [4]Institute of Public Health, Charité – Universitätsmedizin Berlin, Berlin, Germany; [5]Berlin Institute of Health, Charité – Universitätsmedizin Berlin, Berlin, Germany; [6]Department of Internal Medicine and Rheumatology, Klinikum Bielefeld Rosenhöhe, Bielefeld, Germany; [7]Division of Rheumatology, University of Toronto and Schroeder Arthritis Institute, University Health Network, Toronto, Ontario, Canada.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/art.43447).

Author disclosures and graphical abstract are available at https://onlinelibrary.wiley.com/doi/10.1002/art.43447.

Address correspondence via email to Murat Torgutalp, MD, MSc, at murat.torgutalp@charite.de.

Submitted for publication May 29, 2025; accepted in revised form October 10, 2025.

Over the past two decades, the pharmaceutical treatment of axSpA has expanded significantly with the introduction of biologic therapies targeting tumor necrosis factor (TNF), interleukin-17, and JAK.[4] Despite these advancements, current guidelines recommend nonsteroidal anti-inflammatory drugs (NSAIDs) as the initial therapy for the management of axSpA symptoms.[5,6] Beyond symptomatic pain relief, NSAIDs have also been investigated for their disease-modifying impacts on the course of axSpA, particularly with regard to radiographic spinal progression. The potential disease-modifying effects of NSAIDs are attributed to the inhibition of cyclooxygenase (COX) enzymes, leading to reduced production of prostaglandins that play a key role in inflammation and osteoblastogenesis.[7,8]

Despite these potential benefits, existing literature on the impact of NSAIDs on radiographic spinal progression in axSpA presents conflicting findings.[9–13] This inconsistency is primarily due to variations in the types of NSAIDs used, the populations studied, and the study designs. Although some research suggests that NSAIDs, particularly COX2 inhibitors (COX2i), may reduce structural damage,[9,10] other studies, like the ENRADAS and CONSUL trials, present a more cautious view or have not found such a clear benefit.[12,13] Moreover, the majority of these studies have targeted patients with established radiographic disease (radiographic axSpA [r-axSpA], previously referred to as ankylosing spondylitis), with only one cohort study involving patients with nonradiographic axSpA (nr-axSpA).[11]

Given the mixed results in previous studies, the present study aims to investigate the association between NSAID intake and radiographic spinal progression in patients with axSpA. The secondary objectives include assessing the differential impacts of NSAID types (COX2i and nonselective NSAIDs [ns-NSAIDs]) and examining whether this relationship varies among disease subgroups categorized by their classification status (r-axSpA and nr-axSpA).

METHODS

Cohort design and participants. We performed analyses on the data from patients enrolled in German Spondyloarthritis Inception Cohort (GESPIC), an ongoing longitudinal study designed to investigate clinical and radiographic outcomes of patients with axSpA. The study design, methodology, and eligibility criteria were documented in detail in previous publications.[14,15] In summary, inclusion criteria required patients to have either r-axSpA meeting the modified New York criteria with a symptom duration of less than 10 years or nr-axSpA meeting the European Spondyloarthropathy Study Group criteria (with a slight modification) with a symptom duration of less than five years. The r-axSpA or nr-axSpA classification was based on central reading of baseline sacroiliac radiographs as described previously; when central reading results were not available, the local rheumatologist's assessment served as the basis for classification.[15,16] Due to observational nature of the cohort, there were no treatment restrictions, allowing physicians to prescribe treatments based on their clinical judgment. Data on demographics, disease-relevant characteristics (such as activity, function, mobility, and presence of extra musculoskeletal manifestations), laboratory (C-reactive protein [CRP], and HLA-B27), and treatments (NSAIDs, TNF inhibitors [TNFi]) were systematically collected at baseline, every six months for the first two years, and then annually through year 10. Disease activity was assessed by using the Bath Ankylosing Spondylitis Disease Activity Index, CRP levels, and patient global assessment of disease activity. The Axial Spondyloarthritis Disease Activity Score (ASDAS) was also calculated.[17,18]

The study protocol was approved by the ethics committee of the coordinating center (Charité—Universitätsmedizin Berlin, Germany/ethics approval number 188-19) and by local ethical committees of participating centers. Participants gave informed consent to participate in the study before taking part. Patients and/or the public were not involved in the design, conduct, reporting, or dissemination plans of this research. The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Exposure definition: NSAID types and NSAID intake score calculation. During the 10-year follow-up, data on NSAID types, their daily dosage, frequency of intake, and any side effect of NSAIDs were recorded at each visit. The most used COX2i were celecoxib and etoricoxib, whereas diclofenac, ibuprofen, naproxen, and indomethacin were the most used ns-NSAIDs.

We assessed NSAID intake using the score previously proposed by the Assessment of Spondyloarthritis International Society, which has been designed for the quantification of NSAID intake.[19] This score primarily incorporates two key elements: daily dose of NSAID, determined based on the equivalent dose conversion for each NSAID, and duration of NSAID consumption over the relevant time period. Scores for NSAID intake range from 0 to 100, with 0 representing no NSAID use and 100 representing of a full dose NSAID use. For the current analysis, we calculated the total NSAID intake score at each visit for each type and averaged these values within each two-year radiographic interval. This approach allowed us to capture longitudinal exposure, including any discontinuation or switching of NSAIDs during follow-up. We used the total NSAID intake score as the primary exposure of interest for the main analyses. For analyses of NSAID subtypes, we classified patients according to the type of NSAID (COX2i or ns-NSAID) that they used for at least one year within the interval. To help readers better understand and interpret the findings, we reported the results as a 10-point increase in the total NSAID intake score. We also assessed the total causal effect estimate for each

type of NSAIDs to explore their associations with radiographic progression (details are provided in the Statistical analysis section).

Outcome definition: scoring of spinal radiographs and outcomes. Lateral cervical and lumbar spinal radiographs were obtained at baseline and every two years thereafter with a maximum six time points (baseline and years 2, 4, 6, 8, and 10), as specified in the protocol. Patients with at least two consecutive sets of cervical and lumbar spinal radiographs were eligible for analysis, resulting in a final sample of 252 participants.

Three trained and calibrated readers evaluated spine radiographs by using the modified Stoke Ankylosing Spondylitis Spine Score (mSASSS).[20] Readers were blinded to all clinical data except for the chronological order of the images. A total mSASSS, ranging from 0 to 72, was calculated for each reader. In cases in which radiographic data were missing at a given time point, we used the following imputation method: (1) if the mSASSS was missing at a time point, and if the scores were identical for the previous and subsequent time points, the missing score was imputed with the available value; (2) for individual missing vertebral corners, a score of 3 (indicating a total ankylosis) at an earlier time point was used to impute the missing score at a later time point; and similarly, but in reverse, a score of 0 (indicating no abnormality) at a later time point was used to impute the missing score at an earlier time point.

We made no further imputations and calculated the final mSASSS for each patient and time point as the mean scores of three readers. We assessed the reliability of the assessments by the intraclass correlation coefficient for status and change scores of mSASSS among readers. Our outcome was the progression in mSASSS in a two-year interval, which was determined as a difference in final mSASSS between the two time points, modeled as a dependent variable.

Covariates. We used directed acyclic graphs (DAGs) to formalize causal structures and determine the total effect of NSAID intake on radiographic spinal progression.[21] Supplementary Figure 1 represents the DAG, illustrating the exposure (a 10-point increase in the total NSAID intake score over a two-year interval) and outcome (progression in mSASSS); furthermore, this figure denotes the temporal sequence of variables by "t – 1" and "t" for previous and current time points, respectively. We identified covariates from the published literature and expert knowledge to define the exposure-outcome relationship conceptually. These covariates included status and time-averaged variables. Status variables, measured at the beginning of each interval, were sex, age, symptom duration, smoking status, classification status (classification as r-axSpA or nr-axSpA), HLA-B27 positivity, body mass index, and mSASSS; whereas time-averaged variables, calculated over the two-year interval, were ASDAS, side effects due to NSAID intake, conventional disease-modifying antirheumatic drugs (csDMARD) use, and TNFi exposure. Although data regarding inflammation on magnetic resonance imaging (MRI), history of cardiovascular disease, and other potential contributors of mechanical pain (eg, degenerative disc disease, osteoarthritis) were not collected in the cohort's case report form, we nonetheless included them into our DAG because of their potential causal relationships with exposure or outcome.

Statistical analyses. For the descriptive data, we used mean ± SD and number (percentage) for numerical and categorical variables, respectively. The statistical model structure we used is as follows: We used a per-10-point increase in the mean NSAID intake score over the two-year interval as the main exposure variable (independent variable of interest). The β-coefficients for the primary outcome reflected the progression in mSASSS per 10-point increase in NSAID intake.

We analyzed longitudinal association between NSAID intake and radiographic spinal progression using generalized estimating equations (GEEs) with an independent working correlation structure. Given the design of the study, which included longitudinal repeated measures of exposure and outcome and time-dependent confounders, we used sequential conditional mean models (SCMMs) to minimize potential GEE bias as described by Keogh et al.[22] Briefly this approach addresses biases that may arise in traditional regression analyses due to time-varying variables affected by prior exposures or covariates and can be incorporated with propensity score (PS) adjustment.

We built three different models including adjustments for both prior covariates (ASDAS or TNFi exposure) and prior exposure (total NSAID intake score) and implemented SCMMs to determine the total causal effect of NSAID intake on radiographic spinal progression. To explore the robustness of our findings, we conducted sensitivity analyses incorporating PS adjustments within the SCMMs. The PS represents the estimated probability of an individual receiving the exposure at a specific time point ("t"), given their prior and current characteristics. For our continuous exposure variable, we fitted a linear regression model to estimate PS across all time points combined. This model accounted for the influence of relevant variables on exposure or outcome. Subsequently, the estimated PS was integrated into the SCMM for further sensitivity analysis. This step allowed us to control potential confounding factors that might have biased the initial estimates of the association between NSAID intake and radiographic progression. We also tested for all relevant interactions between mSASSS and NSAID exposure, and other covariates.

To assess the effect of two NSAIDs types over two-year radiographic intervals, we also assessed the following two variables as exposure variables: COXi intake score and ns-NSAID intake score, per 10-point increase for both variables. Considering that our main exposure variable was the total NSAID intake score, a compositional (also described as comparative) data set, which consists of these two distinct components, we, therefore, used all-component models to calculate the effect of each NSAID

types on radiographic spinal progression. Briefly, this model targets the total causal effects of two NSAID types separately and does not include the total NSAID intake score (main exposure variable), and, therefore, does not allow opening a closed path by a collider. The estimated coefficient is expected to provide an unbiased total causal effect estimate of the individual NSAID types as described by Tomova et al.[23]

We used DAGitty version 3.1, a web application, to formalize causal structures.[21] For main statistical analyses, we used the "*geepack*" package in R software version 4.3.3 (R Project for Statistical Computing).[24,25] We calculated and reported parameter estimates (β) with 95% confidence intervals (CIs).

RESULTS

Baseline characteristics and NSAID intake. The GESPIC core cohort enrolled 525 patients with axSpA. For our analysis, we focused on 252 patients who had at least two consecutive radiographs (113 r-axSpA and 139 nr-axSpA). Table 1 presents a comparison of the baseline characteristics of included and excluded patients.

The included and excluded groups had similar characteristics, with a mean age of 35.7 ± 10.3 years and a mean symptom duration of 3.9 ± 2.7 years. However, included patients were less frequently male (50% vs 59%) and had a lower mean ASDAS (2.5 vs 2.7) than excluded patients. Additionally, included patients had slightly lower mean baseline mSASSS (2.6 vs 2.7). However, there was no difference in the proportion of patients with syndesmophytes at baseline or the proportion of patients with r-axSpA. Only 8 (3%) of the included patients were treated with a TNFi, whereas 68 (27%) were receiving a csDMARD.

At baseline, 201 (80.0%) of the included patients were receiving NSAID treatment, with 46 (18%) taking COX2i and 156 (62%) taking ns-NSAIDs. The mean total NSAID intake score was 38.3 ± 35.5, with scores of 9.1 ± 22.8 and 29.2 ± 35.1 for COX2i and ns-NSAIDs, respectively.

The effect of total NSAID intake on radiographic spinal progression. Adjusted multivariable GEE analyses showed a consistent association between increased NSAID intake and reduced radiographic progression in patients with whole axSpA in three different models (Table 2). Model

Table 1. Baseline characteristics of included and excluded patients*

Variable	Overall (N = 525)	Included (n = 252)	Excluded (n = 273)
Age at baseline, y	35.7 ± 10.3	35.9 ± 10.2	35.4 ± 10.4
Male sex, n (%)	286 (54)	125 (50)	161 (59)
Symptom duration, y	3.9 ± 2.7	3.9 ± 2.5	4.0 ± 2.9
Family history of SpA, n (%)	158 (30)	86 (34)	72 (26)
HLA-B27 positivity, n (%)	406 (78)	199 (80)	207 (76)
Ever smoked, n (%)	132 (25)	68 (27)	64 (23)
BMI at baseline, kg/m^2	24.9 ± 4.8	24.7 ± 4.8	25.1 ± 4.7
Ever uveitis, n (%)	86 (16)	46 (18)	40 (15)
Ever psoriasis, n (%)	53 (10)	29 (12)	24 (9)
Ever IBD, n (%)	14 (3)	7 (3)	7 (3)
CRP, mg/L	10.7 ± 17.3	8.8 ± 14.2	12.4 ± 19.6
ASDAS	2.6 ± 1.0	2.5 ± 1.0	2.7 ± 1.0
BASDAI (0–10 points NRS)	3.9 ± 2.1	3.8 ± 2.1	4.1 ± 2.0
BASFI (0–10 points NRS)	2.8 ± 2.4	2.7 ± 2.3	2.8 ± 2.4
BASMI (0–10 points NRS)	1.5 ± 1.6	1.6 ± 1.6	1.4 ± 1.7
Use of systemic glucocorticoids, n (%)	48 (9)	17 (7)	31 (11)
Use of csDMARDs, n (%)	132 (25)	68 (27)	64 (23)
Use of TNFi, n (%)	13 (2)	8 (3)	5 (2)
Use of NSAID, n (%)			
No NSAID	111 (21)	50 (20)	61 (22)
COX2i	103 (20)	46 (18)	57 (21)
ns-NSAID	311 (59)	156 (62)	155 (57)
Total NSAID intake score	38.3 ± 35.5	37.1 ± 33.8	39.4 ± 37.1
Mean COX2i intake score	9.1 ± 22.8	7.7 ± 19.3	10.4 ± 25.5
Mean ns-NSAID intake score	29.2 ± 35.1	29.4 ± 34.4	29.0 ± 35.8
Presence of syndesmophytes, n (%)	66 (16)	40 (16)	26 (16)
mSASSS, units	2.5 ± 5.9	2.5 ± 6.4	2.7 ± 4.9
Patients with r-axSpA, n (%)	248 (47)	113 (45)	135 (49)

* Values described as mean ± SD unless otherwise indicated. ASDAS, Axial Spondyloarthritis Disease Activity Score; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; BASMI, Bath Ankylosing Spondylitis Metrology Index; BMI, body mass index; COX2i, selective cyclooxygenase-2 inhibitor; CRP, C-reactive protein; csDMARD, conventional synthetic disease-modifying antirheumatic drug; IBD, inflammatory bowel disease; mSASSS, modified Stroke Ankylosing Spondylitis Spine Score; NRS, numeric rating scale; NSAID, nonsteroidal anti-inflammatory drug; ns-NSAID, nonselective NSAID; r-axSpA, axSpA, axial spondyloarthritis; SpA, spondyloarthritis; TNFi, tumor necrosis factor alpha inhibitor.

Table 2. Effect of per 10-point increase in total NSAID intake score on progression in mSASSS*

		All axSpA[c]		nr-axSpA		r-axSpA	
Model[a]	Exposure[b]	Adjusted β	95% CI	Adjusted β	95% CI	Adjusted β	95% CI
Model 1[d]	Total NSAID intake score	−0.052	−0.097 to −0.007	−0.014	−0.042 to 0.014	−0.077	−0.152 to −0.003
Model 1 with PS	Total NSAID intake score	−0.056	−0.101 to −0.010	−0.012	−0.041 to 0.017	−0.089	−0.164 to −0.014
Model 2[e]	Total NSAID intake score	−0.051	−0.097 to −0.006	−0.007	−0.035 to 0.021	−0.081	−0.157 to −0.005
Model 2 with PS	Total NSAID intake score	−0.056	−0.102 to −0.009	−0.009	−0.037 to 0.020	−0.092	−0.169 to −0.014
Model 3[f]	Total NSAID intake score	−0.051	−0.096 to −0.006	−0.007	−0.032 to 0.019	−0.075	−0.150 to −0.000
Model 3 with PS	Total NSAID intake score	−0.056	−0.103 to −0.009	−0.011	−0.037 to 0.015	−0.088	−0.165 to −0.012

* ASDAS, Axial Spondyloarthritis Disease Activity Score; axSpA, axial spondyloarthritis; CI, confidence interval; EMM, extra musculoskeletal manifestation; mSASSS, modified Stroke Ankylosing Spondylitis Spine Score; nr-axSpA, nonradiographic axSpA; NSAID, nonsteroidal anti-inflammatory drugs; PS, propensity score; r-axSpA, radiographic axSpA; TNFi, tumor necrosis factor alpha inhibitor
[a] Parameter estimates from all the multivariable models adjusted for sex, presence of EMMs, time-averaged ASDAS in the current interval, mSASSS at the beginning of the interval, and total NSAID intake score in the previous interval.
[b] Reference is per 10-point increase in NSAID intake score.
[c] Models in all axSpA groups also include classification as radiographic axSpA at the beginning of each interval in the adjustment set.
[d] Model 1 also adjusted for age at the beginning of the interval, smoking in the interval, and time-averaged ASDAS in the previous interval.
[e] Model 2 also adjusted for age at the beginning of the interval and TNFi exposure (≥12 months) in the previous interval.
[f] Model 3 also adjusted for symptom duration at the beginning of the interval and TNFi exposure (≥12 months) in the previous interval.

1 demonstrated an inverse association between total NSAID intake score and progression in mSASSS, with a moderate reduction of progression per 10-point NSAID intake (adjusted β = –0.052, 95% CI: –0.097 to –0.007), representing a borderline association. Models 2 and 3, which included different adjustment sets, yielded similar results, with a β of –0.051 for both. Similarly, including PS adjustments in the analyses further strengthened the observed association and provided comparable results (β's were –0.056 for all models with PS in Table 2).

Upon stratifying patients according to their classification status, the analyses revealed a trend toward a greater benefit of NSAID intake on radiographic progression between nr-axSpA and r-axSpA subsets. Specifically, we observed a numerically larger reduction in radiographic progression per 10-point increase in NSAID intake score in the r-axSpA group, with an adjusted β of –0.077 (95% CI: –0.152 to –0.003) in model 1 (Table 2). However, the effect was less pronounced in the nr-axSpA population, with a β of –0.014 (95% CI: –0.042 to 0.014). The results indicate that NSAID intake has a more significant impact on reducing radiographic progression in patients with r-axSpA than in those with nr-axSpA. This observed trend was consistent also in other models (βs were –0.081 and –0.075, for model 2 and 3, respectively) and PS adjusted models (βs were –0.092 and –0.088, for model 2 and 3, respectively), as shown in Table 2.

Figure 1 presents the cumulative probability plot, highlighting the divergence in mSASSS change at individual interval level between high (NSAID intake score ≥50) and low (NSAID intake score <50) NSAID intake groups. This trend appeared numerically stronger in the r-axSpA subgroup.

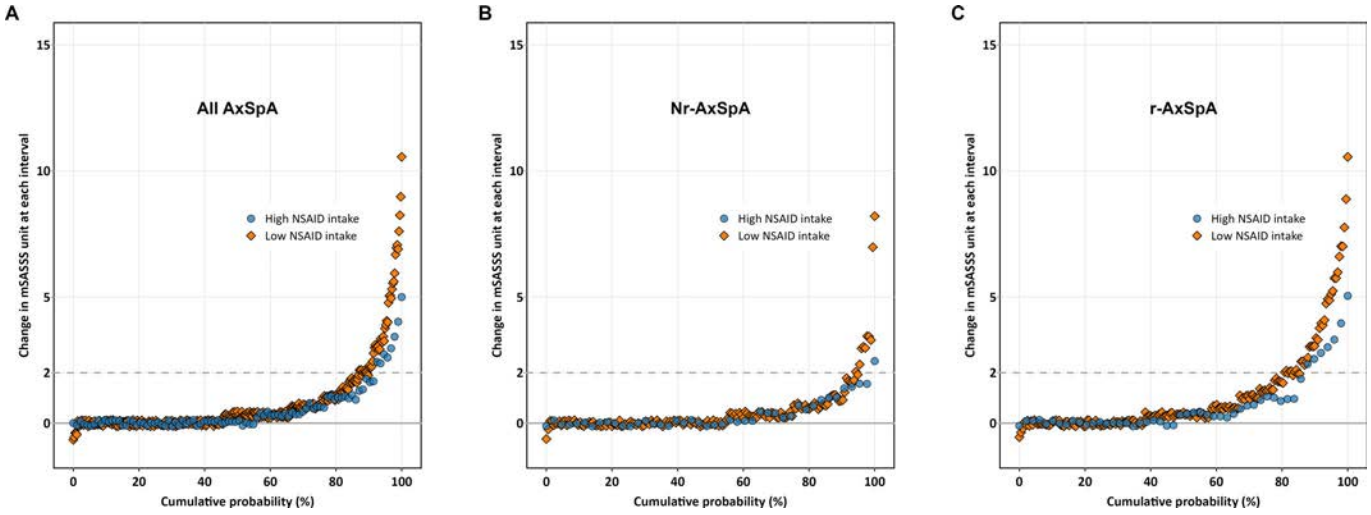


Figure 1. Cumulative probability plot of mSASSS changes over two-year intervals in relation to NSAID intake scores, stratified by high and low NSAID, in (A) patients with whole axSpA, (B) patients with nr-axSpA, (C) and patients with r-axSpA. axSpA, axial spondyloarthritis; mSASSS, modified Stoke Ankylosing Spondylitis Spinal Score; nr-axSpA, nonradiographic axSpA; NSAID, nonsteroidal anti-inflammatory drug; r-axSpA, radiographic axSpA.

Table 3. Effect of per 10-point increase in subtypes of NSAIDs (COX2i and ns-NSAID intake Score) on Progression in mSASSS*

		All axSpA[c]		nr-axSpA		r-axSpA	
Model[a]	Exposure[b]	Adjusted β	95% CI	Adjusted β	95% CI	Adjusted β	95% CI
Model 1[d]	COX2i intake score	−0.061	−0.138 to 0.015	−0.031	−0.064 to 0.003	−0.073	−0.191 to 0.045
Model 1[d]	ns-NSAID intake score	−0.045	−0.089 to −0.001	−0.010	−0.043 to 0.023	−0.076	−0.152 to −0.000
Model 1 with PS	COX2i intake score	−0.064	−0.141 to 0.012	−0.029	−0.061 to 0.004	−0.084	−0.203 to 0.035
Model 1 with PS	ns-NSAID intake score	−0.048	−0.093 to −0.004	−0.008	−0.042 to 0.026	−0.088	−0.165 to −0.012
Model 2[e]	COX2i intake score	−0.062	−0.137 to 0.014	−0.020	−0.053 to 0.013	−0.078	−0.204 to 0.048
Model 2[e]	ns-NSAID intake score	−0.044	−0.087 to −0.000	−0.004	−0.038 to 0.030	−0.078	−0.152 to −0.004
Model 2 with PS	COX2i intake score	−0.069	−0.148 to 0.011	−0.028	−0.061 to 0.005	−0.090	−0.217 to 0.037
Model 2 with PS	ns-NSAID intake score	−0.047	−0.091 to −0.002	−0.004	−0.038 to 0.029	−0.088	−0.164 to −0.012
Model 3[f]	COX2i intake score	−0.061	−0.137 to 0.015	−0.018	−0.049 to 0.013	−0.075	−0.203 to 0.054
Model 3[f]	ns-NSAID intake score	−0.044	−0.087 to −0.001	−0.005	−0.037 to 0.026	−0.071	−0.141 to −0.001
Model 3 with PS	COX2i intake score	−0.069	−0.149 to 0.012	−0.028	−0.060 to 0.003	−0.090	−0.220 to 0.041
Model 3 with PS	ns-NSAID intake score	−0.047	−0.092 to −0.002	−0.008	−0.038 to 0.023	−0.083	−0.156 to −0.011

* ASDAS, Axial Spondyloarthritis Disease Activity Score; axSpA, axial spondyloarthritis; CI, confidence interval; COX2i, selective cyclooxygenase-2 inhibitor; EMM, extra musculoskeletal manifestation; mSASSS, modified Stroke Ankylosing Spondylitis Spine Score; nr-axSpA, nonradiographic axSpA; NSAID, nonsteroidal anti-inflammatory drug; ns-NSAID, nonselective NSAID; PS, propensity score; r-axSpA, radiographic axSpA; TNFi, tumor necrosis factor alpha inhibitor.
[a] Parameter estimates from all the multivariable models adjusted for sex, presence of EMMs, time-averaged ASDAS in the current interval, mSASSS at the beginning of the interval, and total NSAID intake score in the previous interval.
[b] Models in all axSpA groups also include classification as radiographic axSpA at the beginning of each interval in the adjustment set.
[c] Reference is per 10-point increase in corresponding intake score for each exposure. All models include per 10-point increase in COX2i intake score and ns-NSAID intake score (these models did not adjust for total NSAID intake score).
[d] Model 1 also adjusted for age at the beginning of the interval, smoking in the interval, and time-averaged ASDAS in the previous interval.
[e] Model 2 also adjusted for age at the beginning of the interval and TNFi exposure (≥12 months) in the previous interval.
[f] Model 3 also adjusted for symptom duration at the beginning of the interval and TNF exposure (≥12 months) in the previous interval.

The effect of different NSAID types on radiographic spinal progression. Table 3 displays the effects of NSAID types on radiographic spinal progression. Our multivariable analyses suggested numerically larger reduction in progression with COX2i compared to ns-NSAIDs in all patients with axSpA. Model 1 showed a reduction of 0.061 points in mSASSS progression for each 10-point increase in COX2i use (β: –0.061, 95% CI: –0.138 to 0.015), compared with a reduction of 0.045 points for the same increase in ns-NSAID use (β: –0.045, 95% CI: –0.089 to –0.001). Although point estimates suggested a stronger effect for COX2i, the CIs overlapped, and the difference was not statistically significant. Consistent with main analyses, models 2 and 3, and all models incorporating PS adjustments to evaluate the total effect of NSAID types, showed similar trends, with numerically larger reductions in radiographic progression associated with COX2i compared to ns-NSAIDs in the whole axSpA population (Table 3).

After stratification by classification status, we extended our analysis to differentiate between COX2i and ns-NSAIDs. We performed multivariable analyses using the same covariate adjustments as in the models for the whole axSpA group and found similar reductions in radiographic spinal progression in the r-axSpA subgroup for both types of NSAIDs (Table 3). In patients with r-axSpA, each 10-point increase in COX2i intake score was associated with a 0.073-point reduction in mSASSS progression (β = –0.073, 95% CI: –0.191 to 0.045). Notably, a 10-point increase in ns-NSAIDs also corresponded to a similar reduction in mSASSS progression (β = –0.076, 95% CI: –0.152 to 0.000) in the r-axSpA group. However, like our analyses for total NSAID intake score, we observed no statistically significant association for either COX2i or ns-NSAID use in the nr-axSpA group in the nr-axSpA group (β: –0.031, 95% CI: –0.064 to 0.003 and β = –0.010, 95% CI: –0.043 to 0.023, respectively, in model 1). These observed trends were consistent across all other models used in stratified analyses (Table 3). Figure 2 shows a cumulative probability plot that illustrates a trend toward the lower probability of mSASSS progression over two years for intervals covered by COX2i compared to intervals covered by an ns-NSAID, at the level of all individual intervals.

DISCUSSION

Our results consistently showed an association between higher NSAID intake and a retardation in radiographic spinal progression in axSpA across different models. Specifically, a 10-point increase in the total NSAID intake score was associated with a moderate reduction in mSASSS progression (β ranged from –0.051 to –0.058). After stratification, this association was more evident in patients with established radiographic disease (r-axSpA). Additionally, the analyses revealed a stronger effect of COX2i compared to ns-NSAIDs on reducing radiographic spinal progression in whole axSpA population. We believe that these findings provide a robust estimation of the longitudinal impact of NSAIDs (selective COX2 and nonselective inhibitors) on radiographic spinal progression in axSpA (in both radiographic and nonradiographic forms).

The relationship between NSAID intake and radiographic spinal progression in axSpA remains a topic of considerable debate. Although some previous research supported the idea that

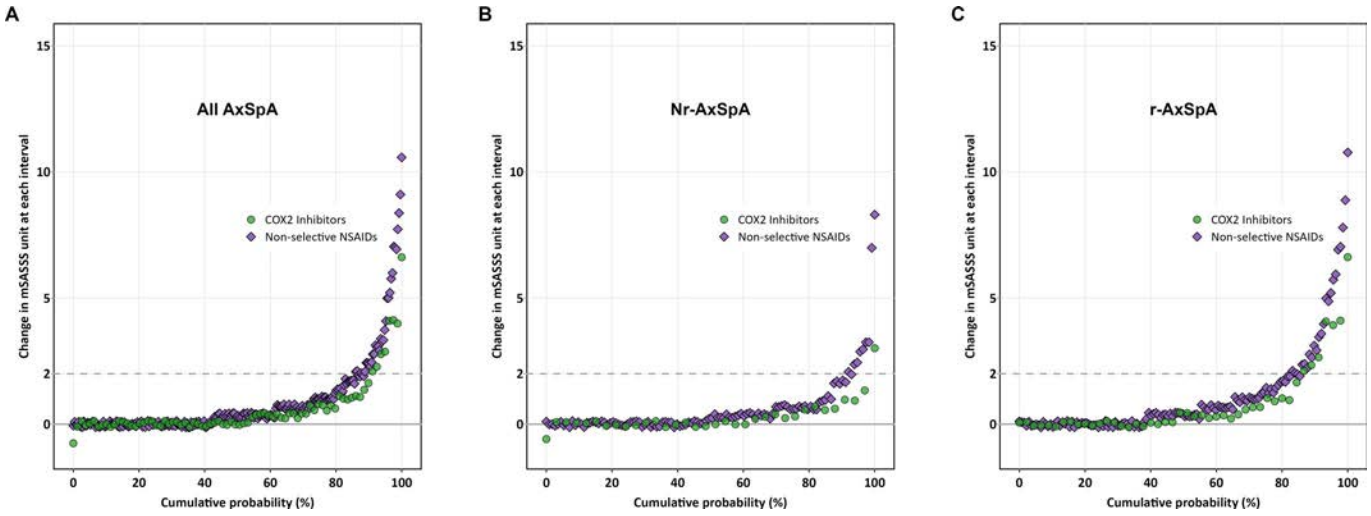


Figure 2. Cumulative probability plot of mSASSS changes over two-year intervals in relation to types of NSAIDs in (A) patients with whole axSpA, (B) patients with nr-axSpA, (C) and patients with r-axSpA. axSpA, axial spondyloarthritis; COX2, selective cyclooxygenase-2; mSASSS, modified Stoke Ankylosing Spondylitis Spinal Score; nr-axSpA, nonradiographic axSpA; NSAID, nonsteroidal anti-inflammatory drug; r-axSpA, radiographic axSpA. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.43447/abstract.

NSAIDs may slow radiographic progression, the magnitude of this effect and its applicability to different types of NSAIDs varied.[9–13] Similar to ours, some studies suggested that NSAIDs might have a beneficial effect in slowing structural damage. For instance, a 2005 randomized controlled trial (RCT) showed a reduction in radiographic spinal progression in patients with r-axSpA who received continuous treatment of celecoxib, a COX2i, compared to those treated on demand.[9] The effect of COX2i on slowing progression was more evident in continuous group, particularly in patients with high disease activity, indicating a potential disease-modifying effect.[10] The two-year results of GESPIC confirmed this finding and demonstrated that higher NSAID intake had a protective effect on radiographic progression, especially in the r-axSpA group, but the study did not differentiate between the effects of COX2i and ns-NSAID intake.[11] However, in 2015 the ENRADAS study showed no significant difference between continuous and on-demand treatment with diclofenac, an ns-NSAID, on structural progression in r-axSpA patients.[12] The most recent CONSUL study further complicated the narrative by questioning the added benefit of celecoxib in combination therapy with a TNFi.[13] Although this study demonstrated a numerical effect of combination with celecoxib on radiographic progression compared to TNFi alone in patients with r-axSpA with high risk for radiographic progression, this difference did not reach statistical significance. These discrepancies among studies can be explained by several factors, chief among which are methodologic differences between studies. Variations in results may be attributable to differences in study design (observational vs RCTs), sample size, type of NSAID used, dosage, and duration of treatment. The observed differences may also be attributed to the specific population studied. It is important to note that all RCTs discussed above included only patients with r-axSpA, and only the cohort study included the entire axSpA population.

These studies highlighted the complexity of understanding the impact of NSAIDs on structural progression in axSpA, although there is some evidence supporting their use in reducing radiographic progression, especially in patients with r-axSpA or with COX2i. The results of the present study suggest a trend toward stronger effects of COX2i compared to ns-NSAIDs in all patients with axSpA, which is consistent with previous research. Additionally, we found that ns-NSAIDs also had a protective effect on progression in r-axSpA group, in contrast to above-mentioned ENRADAS study.[12]

This inverse association between NSAID intake and radiographic progression can be explained by several mechanisms. First, NSAIDs act primarily by inhibiting COX enzymes and exerting their anti-inflammatory effects, leading to a reduction in the production of inflammatory mediators, mainly prostaglandins.[7] As inflammation plays a key role in the development and progression of spinal damage in axSpA, NSAIDs may help to slow the process of damage by reducing inflammation. Second, NSAIDs effectively relieve pain and are therefore used for pain management, which is the main symptom of axSpA. Improved pain control may lead to increased physical activity and mobility, potentially reducing the risk of spinal stiffness and deformity, which may contribute to the slowing of damage. Third, although the specific mechanisms remain to be investigated, some studies suggest that COX2 selectivity may have greater anti-inflammatory and potentially disease-modifying effects compared to ns-NSAIDs—beyond their pain-relieving effects—and this may result in a relatively superior effect on structural damage.[8,26] Therefore,

COX2i potentially offer a more targeted anti-inflammatory effect that may effectively block pathways leading to bone fusion. Finally, we observed that the effect of both NSAIDs on structural damage appeared numerically stronger in the r-axSpA group than in the nr-axSpA group. We suggest that this effect may be attributable to already present structural damage (ie, destructive/erosive changes), which is more common in advanced disease, being more likely to lead to more radiographic progression, and that there may be an interference between the inhibitory effect of NSAIDs on new bone formation and radiographic progression in advanced disease.

These findings have implications for clinical practice in potentially slowing disease progression, particularly suggesting that the beneficial effect of COX2i and ns-NSAIDs in reducing radiographic progression is more pronounced in patients with established radiographic disease (r-axSpA). However, further investigations are required to determine the differences regarding the effect of NSAIDs based on disease classification. If confirmed, these findings could be considered in treatment strategies by tailoring the use of NSAIDs to specific patient subgroups based on disease characteristics and radiographic features. Although NSAIDs are currently recommended as the first-line therapy to manage axSpA symptoms, their potential effect on slowing structural damage may add another dimension to their therapeutic value. Although the absolute effect sizes observed in our study were small, even modest reductions in radiographic progression may accumulate over time and translate into clinically meaningful differences, especially considering the irreversible nature of structural spinal damage in axSpA. At the population level, such reductions may also help decrease overall disease burden. However, it is important to determine the most appropriate type of NSAID, dosage, and duration of treatment to minimize potential side effects, such as gastrointestinal and cardiovascular risks, and maximize benefits.[27] Beyond safety considerations, studies have shown that controlling inflammation with NSAIDs may improve survival, as higher disease activity in axSpA is associated with increased mortality.[28,29] This highlights the importance of carefully balancing the risks and potential long-term benefits of NSAID use.

This study possesses several strengths, the foremost of which is the incorporation of the SCMMs. These models account for longitudinal repeated measures of exposure and outcome, as well as time-dependent confounders. The design of our study takes advantage of this methodology to mitigate biases that often occur in traditional GEE regression analyses when time-varying variables may be affected by past exposures or covariates. Evidence from previous research, including simulation studies, highlights the ability of SCMMs to accurately estimate beta coefficients for associations between time-varying exposures and outcomes. Their effectiveness is comparable to that of marginal structural models, which use inverse probability treatment weighting of PSs to handle this complex study design.[22] Second, our longitudinal statistical modeling isolates the within-patient effect, providing clinicians with the best possible evidence when conducting an RCT is not always feasible. Moreover, incorporating PS adjustment enhances the reliability of our findings. The further strength of this study is its inclusion of entire axSpA population, including both r-axSpA and nr-axSpA, and its assessment of different types of NSAIDs together, unlike other studies that focus either on specific subgroups or on a specific type of NSAID. Our findings underscore the importance of investigating potential differences in treatment response across the whole axSpA spectrum and different types of NSAIDs. In addition, based on our current knowledge, this study highlights for the first time that NSAIDs may have a modest but consistent effect on structural damage even in nr-axSpA population. Finally, usage of systematically collected data from a large, well-defined cohort with extended follow-up allowed a comprehensive and robust analysis.

However, it is important to acknowledge some limitations. First, the cohort's case report form did not provide data on certain factors that could influence the outcome, such as inflammation on MRI, or affect exposure, such as specific comorbidities or alternative causes of pain. Nonetheless, we considered all these factors in our DAG to describe causal relationships when building our models, and we included ASDAS (with CRP) in the multivariable analyses. Second, we excluded patients who did not have consecutive spinal radiographs required for radiographic progression calculation. Although the characteristics of included and excluded patients were similar, this exclusion may have introduced attrition bias. The third limitation concerns the collection of NSAID usage information, which relied on patient-reported intake and may therefore be subject to recall bias. Finally, radiographic progression in nr-axSpA was generally low, possibly independent of therapy. This likely reflects the lower level of baseline structural damage in nr-axSpA, which is associated with less subsequent progression, and may explain the limited NSAID effect observed in this subgroup.

In conclusion, our study adds evidence to support the beneficial effect of higher NSAID intake on slowing radiographic spinal progression in patients with axSpA, and particularly in patients with r-axSpA. More specifically, COX-2i showed a slightly superior effect compared to ns-NSAIDs across the whole axSpA spectrum.

ACKNOWLEDGMENTS

We thank Prof. M. Leirisalo-Repo (Finland), Prof. D. van der Heijde (the Netherlands) and Prof. M. Dougados (France) for scientific advice on the design of the cohort. We thank Beate Buss, Petra Tietz, and Annegret Langdon for monitoring the cohort; Johanna Callhoff, Anja Weiss, Joachim Listing, Maryna Verba, and Martina Niewerth for the data management support and statistical advice; Ani Dilbaryan for image scoring; Torsten Karge for developing the image scoring interface; and to all patients who volunteered to participate in this cohort. Furthermore, we would like to thank the following rheumatologists for inclusion of their patients: J. Brandt, H. Brandt, G.-R. Burmester, H. Deister, E. Edelmann, J. Emmerich, M. Enderlein, A. Gauliard, E. Gromnica-Ihle,

F. Heldmann, S. Hermann, U. von Hinüber, Ü. Hübner, K. Karberg, H. Nüßlein, R. Pelle-Lohfink, D. Pick, G. Reichmuth, E. Riechers, M. Rihl, R. Schmidt, S. Schnarr, U. Schneider, I.-H. Song, I. Spiller, U. Syrbe, V. Walz, S. Wassenberg, H. M. Wisseler, S. Zinke. Open Access funding enabled and organized by Projekt DEAL.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr Torgutalp confirms that all authors have provided the final approval of the version to be published, and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

ROLE OF THE STUDY SPONSOR

Abbott/AbbVie, Amgen, Centocor, Schering-Plough, and Wyeth had no role in the study design or in the collection, analysis, or interpretation of the data, the writing of the manuscript, or the decision to submit the manuscript for publication. Publication of this article was not contingent upon approval by Abbott/AbbVie, Amgen, Centocor, Schering-Plough, and Wyeth.

REFERENCES

1. Sieper J, Poddubnyy D. Axial spondyloarthritis. Lancet 2017; 390(10089):73–84.
2. Machado P, Landewé R, Braun J, et al. Both structural damage and inflammation of the spine contribute to impairment of spinal mobility in patients with ankylosing spondylitis. Ann Rheum Dis 2010;69(8): 1465–1470.
3. Poddubnyy D, Listing J, Haibel H, et al. Functional relevance of radiographic spinal progression in axial spondyloarthritis: results from the GErman SPondyloarthritis Inception Cohort. Rheumatology (Oxford) 2018;57(4):703–711.
4. Torgutalp M, Poddubnyy D. Emerging treatment options for spondyloarthritis. Best Pract Res Clin Rheumatol 2018;32(3): 472–484.
5. Ward MM, Deodhar A, Gensler LS, et al. 2019 update of the American College of Rheumatology/Spondylitis Association of America/Spondyloarthritis Research and Treatment Network recommendations for the treatment of ankylosing spondylitis and nonradiographic axial spondyloarthritis. Arthritis Rheumatol 2019;71(10):1599–1613.
6. Ramiro S, Nikiphorou E, Sepriano A, et al. ASAS-EULAR recommendations for the management of axial spondyloarthritis: 2022 update. Ann Rheum Dis 2023;82(1):19–34.
7. Blackwell KA, Raisz LG, Pilbeam CC. Prostaglandins in bone: bad cop, good cop? Trends Endocrinol Metab 2010;21(5): 294–301.
8. Nagano A, Arioka M, Takahashi-Yanaga F, et al. Celecoxib inhibits osteoblast maturation by suppressing the expression of Wnt target genes. J Pharmacol Sci 2017;133(1):18–24.
9. Wanders A, Heijde D, Landewé R, et al. Nonsteroidal antiinflammatory drugs reduce radiographic progression in patients with ankylosing spondylitis: a randomized clinical trial. Arthritis Rheum 2005;52(6): 1756–1765.
10. Kroon F, Landewé R, Dougados M, et al. Continuous NSAID use reverts the effects of inflammation on radiographic progression in patients with ankylosing spondylitis. Ann Rheum Dis 2012;71(10): 1623–1629.
11. Poddubnyy D, Rudwaleit M, Haibel H, et al. Effect of non-steroidal anti-inflammatory drugs on radiographic spinal progression in patients with axial spondyloarthritis: results from the German Spondyloarthritis Inception Cohort. Ann Rheum Dis 2012;71(10): 1616–1622.
12. Sieper J, Listing J, Poddubnyy D, et al. Effect of continuous versus on-demand treatment of ankylosing spondylitis with diclofenac over 2 years on radiographic progression of the spine: results from a randomised multicentre trial (ENRADAS). Ann Rheum Dis 2016;75(8): 1438–1443.
13. Proft F, Torgutalp M, Muche B, et al. Comparison of the effect of treatment with NSAIDs added to anti-TNF therapy versus anti-TNF therapy alone on the progression of structural damage in the spine over 2 years in patients with radiographic axial spondyloarthritis from the randomised-controlled CONSUL trial. Ann Rheum Dis 2024;83(5): 599–607.
14. Rudwaleit M, Haibel H, Baraliakos X, et al. The early disease stage in axial spondylarthritis: results from the German Spondyloarthritis Inception Cohort. Arthritis Rheum 2009;60(3):717–727.
15. Torgutalp M, Rios Rodriguez V, Dilbaryan A, et al. Treatment with tumour necrosis factor inhibitors is associated with a time-shifted retardation of radiographic spinal progression in patients with axial spondyloarthritis. Ann Rheum Dis 2022;81(9): 1252–1259.
16. Torgutalp M, Rios Rodriguez V, Proft F, et al. Treatment with tumor necrosis factor inhibitors is associated with a time-shifted retardation of radiographic sacroiliitis progression in patients with axial spondyloarthritis: 10-year results from the German Spondyloarthritis Inception Cohort. Arthritis Rheumatol 2022;74(9):1515–1523.
17. Machado P, Landewé R, Lie E, et al; Assessment of SpondyloArthritis international Society. Ankylosing Spondylitis Disease Activity Score (ASDAS): defining cut-off values for disease activity states and improvement scores. Ann Rheum Dis 2011;70(1):47–53.
18. van der Heijde D, Molto A, Ramiro S, et al. Goodbye to the term 'ankylosing spondylitis', hello 'axial spondyloarthritis': time to embrace the ASAS-defined nomenclature. Ann Rheum Dis 2024; 83(5):547–549.
19. Dougados M, Simon P, Braun J, et al. ASAS recommendations for collecting, analysing and reporting NSAID intake in clinical trials/epidemiological studies in axial spondyloarthritis. Ann Rheum Dis 2011;70(2):249–251.
20. van der Heijde D, Landewé R. Selection of a method for scoring radiographs for ankylosing spondylitis clinical trials, by the Assessment in Ankylosing Spondylitis Working Group and OMERACT. J Rheumatol 2005;32(10):2048–2049.
21. Textor J, van der Zander B, Gilthorpe MS, et al. Robust causal inference using directed acyclic graphs: the R package 'dagitty'. Int J Epidemiol 2016;45(6):1887–1894.
22. Keogh RH, Daniel RM, VanderWeele TJ, et al. Analysis of longitudinal studies with repeated outcome measures: adjusting for time-dependent confounding using conventional methods. Am J Epidemiol 2018;187(5):1085–1092.
23. Tomova GD, Arnold KF, Gilthorpe MS, et al. Adjustment for energy intake in nutritional research: a causal inference perspective. Am J Clin Nutr 2022;115(1):189–198.
24. Højsgaard S, Halekoh U, Yan J. The R package geepack for generalized estimating equations. J Stat Softw 2006;15:1–11.
25. *R*. The R Foundation; 2013.

26. Matsuyama A, Higashi S, Tanizaki S, et al. Celecoxib inhibits osteoblast differentiation independent of cyclooxygenase activity. Clin Exp Pharmacol Physiol 2018;45(1):75–83.
27. Braun J, Baraliakos X, Westhoff T. Nonsteroidal anti-inflammatory drugs and cardiovascular risk - a matter of indication. Semin Arthritis Rheum 2020;50(2):285–288.
28. Bakland G, Gran JT, Nossent JC. Increased mortality in ankylosing spondylitis is related to disease activity. Ann Rheum Dis 2011; 70(11):1921–1925.
29. Haroon NN, Paterson JM, Li P, et al. Patients with ankylosing spondylitis have increased cardiovascular and cerebrovascular mortality: a population-based study. Ann Intern Med 2015;163(6):409–416.

Arthritis & Rheumatology
Vol. 78, No. 3, March 2026, pp 592–602
DOI 10.1002/art.43332

American College of Rheumatology
Empowering Rheumatology Professionals

Association of Abatacept With Lower Mortality Risk Compared to Rituximab in Rheumatoid Arthritis–Associated Interstitial Lung Disease: An Emulated Target Trial

Po-Cheng Shih,[1,2] Shiow-Ing Wang,[3,4] and James Cheng Chung Wei[1,5,6,7]

Objective. The optimal treatment strategy for rheumatoid arthritis–associated interstitial lung disease (RA-ILD) remains uncertain, and direct comparative data between biologics are limited. This study aimed to evaluate the effectiveness and safety of abatacept compared with rituximab in patients with RA-ILD.

Methods. An emulated target trial was designed using the TriNetX US Collaborative Network database, including patients with RA-ILD, diagnosed between 2007 and 2024. Propensity score matching was used to balance baseline characteristics between the two treatment groups (abatacept and rituximab). The primary outcome was all-cause mortality, and secondary outcomes included respiratory events, medical utilization, and infection-related adverse events. Hazard ratios (HRs) with 95% confidence intervals (CIs) were estimated to use Cox proportional hazards models.

Results. In total, 1,615 patients per group were identified after matching for analysis. Abatacept was associated with a significantly lower risk of all-cause mortality compared with rituximab (HR, 0.689; 95% CI, 0.581–0.818) and a reduced risk of mechanical ventilation (HR, 0.698; 95% CI, 0.521–0.934). The subgroup analyses yielded consistent findings. Sensitivity analyses excluding patients with concomitant connective tissue diseases also demonstrated consistent results (mortality: HR, 0.679; 95% CI, 0.570–0.810), reinforcing the robustness of the findings.

Conclusion. Abatacept was associated with a lower risk of mortality compared with rituximab in patients with RA-ILD. Because clinicians may preferentially reserve abatacept for less aggressive RA-ILD, residual confounding by indication cannot be excluded; thus, the association should not be interpreted as proof of causality. Prospective randomized trials are needed to confirm whether abatacept confers a true survival advantage.

INTRODUCTION

Current American College of Rheumatology (ACR) guidelines recommend rituximab as the first-line biologic therapy for rheumatoid arthritis–associated interstitial lung disease (RA-ILD), with tocilizumab as an alternative.[1] However, emerging evidence suggests that abatacept may be a promising option.

The preliminary data of post hoc analysis of a placebo-controlled trial reported a lower incidence of ILD among patients with RA treated with abatacept.[2] Although an analysis of Medicare and MarketScan data found no significant difference in ILD exacerbation risk between abatacept and tumor necrosis factor inhibitors (TNFi), a national multicenter study of 263 patients with RA-ILD demonstrated higher drug persistence and improved or stable outcomes for both joint and lung function in those receiving abatacept.[3,4] Other studies have further demonstrated abatacept's effectiveness across diverse ILD patterns, including nonspecific interstitial pneumonia and usual interstitial pneumonia, irrespective of administration route (subcutaneous or intravenous).[5–7] These findings highlight the potential role of abatacept in the management of RA-ILD.

[1]Po-Cheng Shih, MD, James Cheng Chung Wei, PhD: Institute of Medicine, Chung Shan Medical University, Taichung, Taiwan; [2]Po-Cheng Shih, MD: Division of Allergy, Immunology and Rheumatology, Department of Internal Medicine, Changhua Christian Hospital, Changhua, Taiwan; [3]Shiow-Ing Wang, PhD: Center for Health Data Science, Department of Medical Research, Chung Shan Medical University Hospital, Taichung, Taiwan; [4]Shiow-Ing Wang, PhD: Department of Health Policy and Management, College of Health Care and Management, Chung Shan Medical University, Taichung, Taiwan; [5]James Cheng Chung Wei, PhD: Department of Allergy, Immunology and Rheumatology, Chung Shan Medical University Hospital, Taichung, Taiwan; [6]James Cheng Chung Wei, PhD: Office of Research and Development, Asia University, Taichung, Taiwan; [7]James Cheng Chung Wei, PhD: Graduate Institute of Integrated Medicine, China Medical University, Taichung, Taiwan.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/art.43332).

Author disclosures and graphical abstract are available online at https://onlinelibrary.wiley.com/doi/10.1002/art.43332.

[Correction added on 9 March 2026, after first online publication: Qing-Wen Wang was removed as author from this version and affiliation footnotes have been renumbered accordingly.]

Address correspondence via email to James Cheng Chung Wei, MD, PhD, at jccwei@gmail.com.

Submitted for publication March 5, 2025; accepted in revised form July 7, 2025.

Given the absence of direct head-to-head comparative evidence between abatacept and rituximab, we conducted this large-scale observational study using an emulated target trial design to evaluate the comparative effectiveness and safety of abatacept versus rituximab in patients with RA-ILD.

METHODS

Study design and data source. This retrospective cohort study employed a target trial emulation framework to compare the effectiveness and safety of abatacept versus rituximab in patients with RA-ILD, using deidentified electronic health record (EHR) data collected for clinical practice and administrative purposes, with optional linkage to administrative claims when available.[8,9] Data were obtained from the TriNetX US Collaborative Network, a federated EHR platform comprising data from 71 health care organizations across the United States.[9] TriNetX is a global, federated real-world data and analytics platform that aggregates deidentified EHR and, where available, linked administrative claims data from multiple collaborative networks. The major TriNetX dataset is sourced from participating US health care organizations. Each site exports pseudonymized EHR records from its clinical data management system into a secure, behind-the-firewall TriNetX node. When an external researcher issues a query, only preaggregated, deidentified results, compliant with HIPAA Safe Harbor, are returned; no patient-level data ever leave individual firewalls. Importantly, data are ingested in near-real time via periodic synchronizations, so the network continuously reflects up-to-date clinical activity, although the exact refresh frequency varies by site.[9]

For accurate diagnosis of RA, key diagnostic codes (International Classification of Diseases, Ninth Revision [ICD-9] and ICD-10, Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT), Logical Observation Identifiers Names and Codes (LOINC) used in TriNetX database have demonstrated good validity in external validations, positive predictive values of 82% to 84% in the Rheumatology Informatics System for Effectiveness (RISE) registry and MarketScan studies, 77.1% to 79.2% for seropositive RA, and 61.6% to 89.5% for seronegative RA in the Optum database.[10–12] In a prior Medicare-based validation study, the optimal claims-based algorithm, requiring two or more International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) codes of ILD, achieved a positive predictive value of 77% (95% confidence interval [CI], 67%–84%) for prevalent ILD and 96% (95% CI, 85%–100%) for incident ILD among patients with RA.[13]

The detailed study design is illustrated in Supplementary Figure 1. The study protocol was reviewed and approved by the institutional review board of Chung Shan Medical University Hospital (approval number CS2-21176). As this study was conducted using deidentified EHRs, informed consent from patients was waived.

Participants and eligibility criteria. Eligible patients had been diagnosed with RA, followed by newly diagnosed ILD between January 1, 2007, and June 30, 2024. Patients younger than 18 years were excluded. RA was identified using ICD-10-CM codes M05 and M06, and ILD was defined using codes J84.1, J84.8, and J84.9. Patients without clearly documented diagnoses or incomplete diagnostic information were excluded. Complete eligibility criteria are outlined in Supplementary Table 1. Diagnoses and treatments were identified using the ICD-10-CM codes (see Supplementary Table 2).

Treatment strategies. The study period was from January 1, 2007, to June 30, 2024. Patients were categorized into two cohorts based on the first prescription of either abatacept or rituximab during the study period. Treatment exposure was defined as the first prescription of abatacept or rituximab, identified via RxNorm codes, within the study period. The date of this initial prescription was defined as the index date (Figure 1).

Outcomes. Patients were observed from the index date until the occurrence of the outcome event, last medical visit, or for up to five years, whichever occurred first. The primary outcome was all-cause mortality. Secondary outcomes included hospitalization, mechanical ventilation, subsequent use of cyclophosphamide, plasma exchange, or intravenous immunoglobulin, identified through Current Procedural Terminology, RxNorm, and Healthcare Common Procedure Coding System codes. To assess infection-related complications, we also analyzed bacteremia and pneumonia as key secondary outcomes. Additionally, external causes of morbidity were used as a negative control outcome to assess potential residual confounding unrelated to RA-ILD treatments (Supplementary Table 2).

Emulation of the target trials. For the target population, RA and ILD diagnoses were identified using ICD-10-CM codes, supported by related clinical data, including pulmonary function tests and oxygen saturation levels. Patients with inconclusive or incomplete diagnostic data were excluded. This multilayered approach minimized potential diagnostic misclassification.

Baseline covariates were carefully selected to mitigate bias, incorporating more than 70 variables, including (1) demographic factors (age at index date, sex, race, and body mass index); (2) socioeconomic status (ICD-10-CM codes Z55–Z65); (3) medical utilizations; (4) severity of disease (RA and ILD): erythrocyte sedimentation rate (ESR), baseline oxygen saturation, forced vital capacity (FVC) ≤80%, and DL_{CO} ≤60%[14]; (5) comorbidities (cardiovascular diseases, diabetes mellitus, chronic lung and kidney diseases, cor pulmonale, secondary pulmonary hypertension, and autoimmune conditions—eg, idiopathic inflammatory

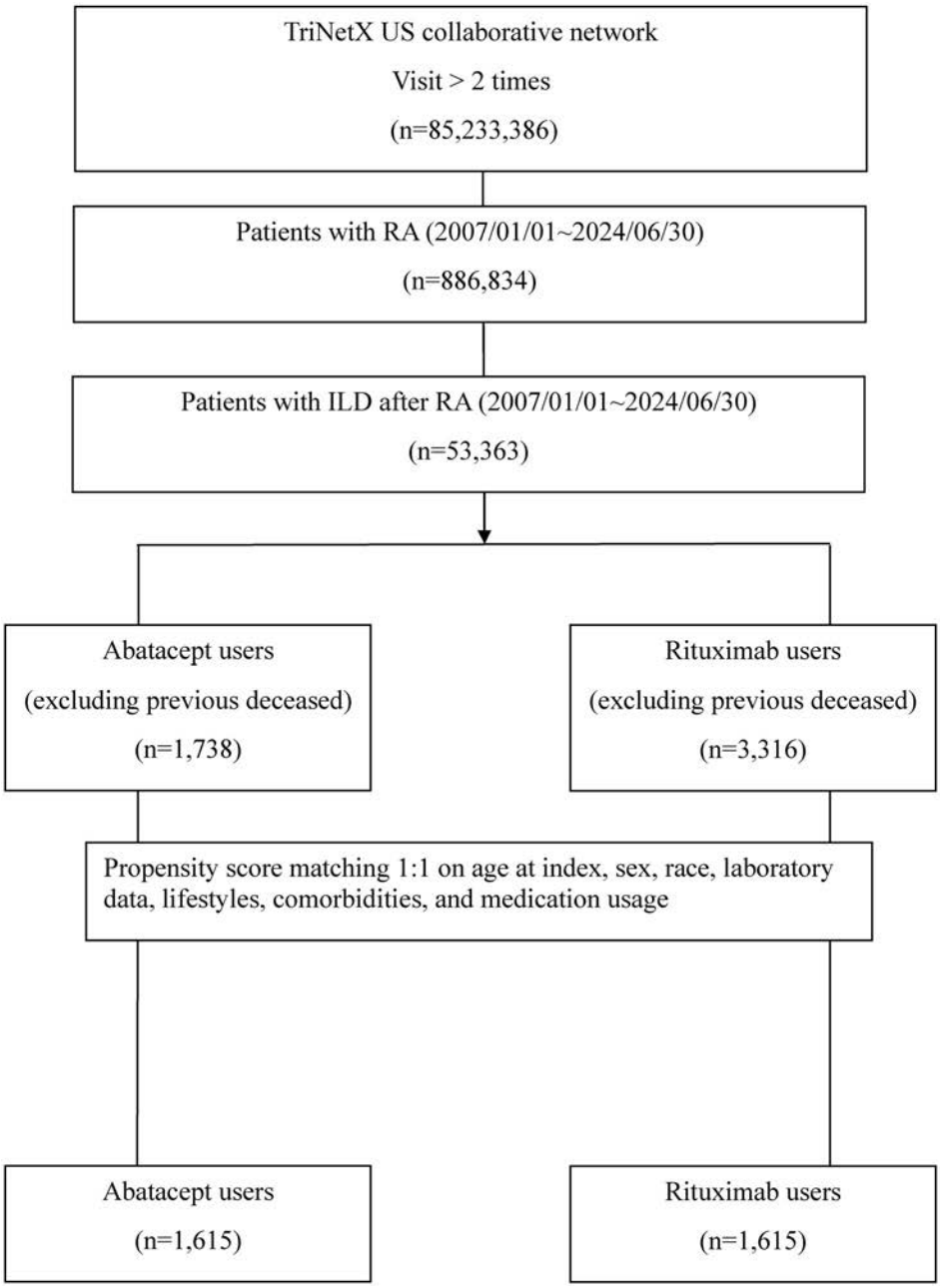


Figure 1. Flow chart of cohort construction. ILD, interstitial lung disease; RA, rheumatoid arthritis.

myositis [IIM], systemic sclerosis [SSc], Sjögren disease); and (6) medication history (use of steroids, methotrexate, mycophenolate, cyclosporine, and treatments for cardiovascular disease, pulmonary hypertension, and chronic lung disease—eg, inhaled steroids, bronchodilators). Variable definitions are detailed in Supplementary Table 2.

Statistical analyses. Propensity score matching (PSM) was conducted using demographic, lifestyle, health care utilization, comorbidity, laboratory, and medication data. Balance between groups was assessed using the standardized mean difference (SMD), with SMD <0.1 indicating adequate balance. Hazard ratios (HRs) with 95% CIs were estimated using Cox proportional hazards models. Kaplan–Meier analyses were used to estimate time-to-event for all-cause mortality and secondary outcomes. Statistical significance was defined as a two-sided P value <0.05.

Subgroup analyses. To explore potential treatment heterogeneity, subgroup analyses were performed based on the following: (1) sex (male vs female), to assess potential hormonal or immunologic differences in treatment response; (2) age (18–64 vs ≥65 years), to evaluate the influence of age-related factors on treatment outcomes; (3) prior exposure to abatacept or rituximab before first ILD code (yes vs no), to

evaluate whether previous exposure to either drug influenced treatment outcomes; (4) biologic-naïve vs biologic-experienced status, to assess whether prior exposure to biologics (other than abatacept or rituximab) modified treatment effectiveness; (5) oxygen dependence on baseline (yes vs no), to account for baseline lung disease severity; (6) steroid use within one month of index date (yes vs no), to evaluate the impact of concurrent glucocorticoid therapy on outcomes[15]; and (7) interval from ILD diagnosis to abatacept/rituximab initiation (≤2 years vs >2 years): to assess whether earlier versus later treatment initiation influences effectiveness and safety outcomes.

Sensitivity analyses. To ensure robustness, multiple sensitivity analyses were conducted: (1) analyzing outcomes restricted to a short-term follow-up period of up to one year after the index date; (2) washout period, which excluded patients who had received other biologics or JAK inhibitor (JAKi) within six months before the index date or had inpatient or critical care admissions within one month before treatment initiation; (3) treatment stability, which excluded patients who switched to another biologic or JAKi after the index date; (4) restricting the analysis to patients without IIM or SSc to reduce confounding and ensure a more homogeneous RA population; (5) excluding patients with malignancy, IIM, or SSc at baseline to mitigate potential bias due to physician treatment preferences influenced by malignancy history; (6) applying a more stringent RA-ILD definition by requiring at least two separate RA diagnoses in the EHR before the first ILD diagnosis, compared with the requirement of at least one RA diagnosis in the primary analysis, to enhance diagnostic specificity and reduce misclassification; (7) smoking history, which is an established risk factor for both the development and progression of RA-ILD (We identified patients with smoking history based on ICD-10-CM codes Z87.891, F17, and Z72.0 to assess the impact of smoking; due to inherent limitations and incomplete smoking status documentation, we restricted analysis to patients with documented smoking history, recognizing potential residual selection bias); (8) identifying patients with diagnosis of secondary pulmonary hypertension based on ICD-10-CM codes (I27.2); (9) identifying patients with a record of higher inflammatory biomarker within six months before index date; (10) identifying patients without any record of methotrexate use within six months before index date; (11) identifying patients with potential adverse events, which are potentially associated with steroid based on ICD-10-CM codes (T38.0X, E24.2, E09, M80, M81) before the index date[16]; and (12) reconstructing the cohort from 2007 to 2019 to avoid confounding by indication due to accumulating evidence or the implementation of clinical guidelines regarding RA-ILD treatment in recent years. These sensitivity analyses aimed to analyze the different situations to test robustness of our findings (Supplemental Statistical Analysis Plan).

RESULTS

Patient characteristics. In this emulated target trial, the abatacept group comprised 1,738 patients, and the rituximab group included 3,316 patients in the cohort construction. After PSM, each group comprised 1,615 well-matched patients, and baseline characteristics (demographics, comorbidities, medication use) were evenly balanced (SMD < 0.1; Table 1; Supplementary Figure 2).

All-cause mortality. During the five-year follow-up, the abatacept group had significantly fewer deaths compared with the rituximab group (219 vs 328 events; HR, 0.689; 95%, CI 0.581–0.818; Figure 2; Table 2).

Secondary outcomes. Compared with rituximab, abatacept was also associated with significantly lower risks of secondary outcomes, including hospitalization (HR, 0.882; 95% CI, 0.776–0.977), mechanical ventilation (HR, 0.698; 95% CI, 0.521–0.934), cyclophosphamide use (HR, 0.162; 95% CI, 0.090–0.291), and immunoglobulin use (HR, 0.232; 95% CI, 0.157–0.343). The incidence of pneumonia was significantly lower with abatacept (HR, 0.858; 95% CI, 0.754–0.977), whereas bacteremia risk was comparable between groups (Table 2).

Subgroup analysis. Subgroup analyses consistently supported the main findings, showing lower mortality associated with abatacept across most clinical scenarios, including sex, age groups, prior biologic exposure, concurrent steroid use, and interval from ILD diagnosis to abatacept/rituximab initiation (Supplementary Table 3). The forest plot for all-cause mortality stratified by subgroup analysis is shown in Figure 3.

Among oxygen-independent patients, abatacept significantly reduced the risk of all-cause mortality (HR, 0.702; 95% CI, 0.582–0.846) and hospitalization (HR, 0.888; 95% CI, 0.797–0.989). In contrast, outcomes were comparable between treatments among oxygen-dependent patients, possibly reflecting limited statistical power due to smaller subgroup size (n = 101). The results demonstrated a lower risk of all-cause mortality with abatacept among biologic-naïve patients; however, this benefit was not observed in biologic-experienced patients. Patients concurrently receiving glucocorticoids appeared to experience greater benefit, whereas those without glucocorticoid use showed less pronounced differences (Figure 3; Supplementary Tables 3–10).

Sensitivity analysis for patients with washout period (stricter inclusion criteria). To minimize confounding from severe baseline comorbidities, sensitivity analysis excluding recent biologic or JAKi use or hospitalization before treatment initiation still showed that abatacept was associated with a lower mortality risk (HR, 0.688; 95% CI, 0.524–0.902) and mechanical ventilation (HR, 0.6563; 95% CI, 0.322–0.984). However,

Table 1. Baseline Characteristics of Study Participants (Before and After PSM Matching)*

Variables	Before PSM			After PSM[a]		
	ABA cohort (n = 1,738)	RTX cohort (n = 3,316)	SMD	ABA cohort (n = 1,615)	RTX cohort (n = 1,615)	SMD
Age at index, mean ± SD, y						
Mean ± SD	63.9 ± 12.1	59.2 ± 13.9	0.365	63.4 ± 12.2	63.5 ± 11.8	0.013
Sex, n (%)						
Female	1,281 (73.7)	2,385 (71.9)	0.04	1,186 (73.4)	1,189 (73.6)	0.004
Male	394 (22.7)	870 (26.2)	0.083	386 (23.9)	382 (23.7)	0.006
Race, n (%)						
White	1,212 (69.7)	2,094 (63.1)	0.14	1,126 (69.7)	1,133 (70.2)	0.009
Black or African American	228 (13.1)	655 (19.8)	0.18	219 (13.6)	216 (13.4)	0.005
Asian	48 (2.8)	118 (3.6)	0.046	46 (2.8)	50 (3.1)	0.015
Lifestyles, n (%)						
Body weight index (≥25)	1,005 (57.8)	2,181 (65.8)	0.164	950 (58.8)	964 (59.7)	0.018
Laboratory data, n (%)						
Baseline oxygen saturation, at most 90%	96 (5.5)	305 (9.2)	0.141	92 (5.7)	101 (6.3)	0.024
Erythrocyte sedimentation rate (≥20 mm/h)	652 (37.5)	1,426 (43.0)	0.112	608 (37.6)	615 (38.1)	0.009
DLco predicted, at most 60%	11 (0.6)	52 (1.6)	0.09	11 (0.7)	11 (0.7)	0
FVC predicted, at most 80%	15 (0.9)	52 (1.6)	0.064	15 (0.9)	16 (1.0)	0.006
Blood gas Pao_2 <60%	27 (1.6)	88 (2.7)	0.077	26 (1.6)	28 (1.7)	0.01
Medical utilizations, n (%)						
Office or other outpatient services	1,078 (62.0)	2,175 (65.6)	0.074	993 (61.5)	1,005 (62.2)	0.015
Preventive medicine services	116 (6.7)	181 (5.5)	0.051	99 (6.1)	100 (6.2)	0.003
Emergency department services	368 (21.2)	795 (24.0)	0.067	321 (19.9)	321 (19.9)	0
Hospital inpatient services	310 (17.8)	894 (27.0)	0.22	295 (18.3)	291 (18.0)	0.006
Noninvasive ear or pulse oximetry for oxygen saturation	20 (1.2)	75 (2.3)	0.086	20 (1.2)	14 (0.9)	0.036
Pulmonary stress testing	95 (5.5)	329 (9.9)	0.168	93 (5.8)	97 (6.0)	0.011
Socioeconomic status, n (%)						
Persons with potential health hazards related to socioeconomic and psychosocial circumstances	38 (2.2)	72 (2.2)	0.001	31 (1.9)	33 (2.0)	0.009
Comorbidities, n (%)						
Hypertensive diseases	783 (45.1)	1,539 (46.4)	0.027	712 (44.1)	720 (44.6)	0.01
Diabetes mellitus	332 (19.1)	670 (20.2)	0.028	298 (18.5)	312 (19.3)	0.022
Ischemic heart diseases	320 (18.4)	606 (18.3)	0.004	282 (17.5)	277 (17.2)	0.008
Other chronic obstructive pulmonary disease	282 (16.2)	481 (14.5)	0.048	254 (15.7)	247 (15.3)	0.012
Asthma	254 (14.6)	477 (14.4)	0.007	218 (13.5)	217 (13.4)	0.002
Other and unspecified asthma	214 (12.3)	426 (12.8)	0.016	186 (11.5)	188 (11.6)	0.004
Mild intermittent asthma	55 (3.2)	67 (2.0)	0.072	38 (2.4)	40 (2.5)	0.008
Mild persistent asthma	28 (1.6)	29 (0.9)	0.067	14 (0.9)	12 (0.7)	0.014
Moderate persistent asthma	42 (2.4)	58 (1.7)	0.047	33 (2.0)	26 (1.6)	0.032
Severe persistent asthma	12 (0.7)	18 (0.5)	0.019	10 (0.6)	10 (0.6)	0
Chronic kidney disease	185 (10.6)	407 (12.3)	0.051	174 (10.8)	159 (9.8)	0.031
End stage renal disease	10 (0.6)	39 (1.2)	0.065	10 (0.6)	10 (0.6)	0
Cerebrovascular diseases	101 (5.8)	209 (6.3)	0.021	91 (5.6)	95 (5.9)	0.011
Dermatomyositis/polymyositis	18 (1.0)	411 (12.4)	0.466	18 (1.1)	17 (1.1)	0.006
Systemic sclerosis (scleroderma)	44 (2.5)	358 (10.8)	0.336	44 (2.7)	42 (2.6)	0.008
Sjögren disease	147 (8.5)	379 (11.4)	0.099	141 (8.7)	137 (8.5)	0.009
Other secondary pulmonary hypertension	153 (8.8)	597 (18.0)	0.273	150 (9.3)	147 (9.1)	0.006
Other specified pulmonary heart diseases	50 (2.9)	164 (4.9)	0.107	47 (2.9)	45 (2.8)	0.007
Medications, n (%)						
Antirheumatic agents						
Glucocorticoids for systemic use	1,286 (74.0)	2,702 (81.5)	0.181	1,201 (74.4)	1,206 (74.7)	0.007
Methotrexate	467 (26.9)	660 (19.9)	0.165	411 (25.4)	409 (25.3)	0.003
Mycophenolic acid	22 (1.3)	186 (5.6)	0.24	22 (1.4)	22 (1.4)	0
Mycophenolate mofetil	92 (5.3)	785 (23.7)	0.541	92 (5.7)	89 (5.5)	0.008
Cyclosporine	54 (3.1)	107 (3.2)	0.007	51 (3.2)	44 (2.7)	0.026
Inhalation medications						
Fluticasone	387 (22.3)	775 (23.4)	0.026	347 (21.5)	366 (22.7)	0.028
Budesonide	158 (9.1)	342 (10.3)	0.041	144 (8.9)	149 (9.2)	0.011
Mometasone	62 (3.6)	111 (3.3)	0.012	55 (3.4)	53 (3.3)	0.007

(Continued)

Table 1. (*Cont'd*)

Variables	Before PSM			After PSM[a]		
	ABA cohort (n = 1,738)	RTX cohort (n = 3,316)	SMD	ABA cohort (n = 1,615)	RTX cohort (n = 1,615)	SMD
Beclomethasone	10 (0.6)	26 (0.8)	0.025	10 (0.6)	10 (0.6)	0
Ciclesonide	10 (0.6)	10 (0.3)	0.041	10 (0.6)	10 (0.6)	0
Flunisolide	10 (0.6)	10 (0.3)	0.041	10 (0.6)	0 (0.0)	0.112
Albuterol	593 (34.1)	1,378 (41.6)	0.154	549 (34.0)	556 (34.4)	0.009
Levalbuterol	34 (2.0)	91 (2.7)	0.052	32 (2.0)	37 (2.3)	0.021
Formoterol	129 (7.4)	257 (7.8)	0.012	116 (7.2)	123 (7.6)	0.017
Salmeterol	111 (6.4)	188 (5.7)	0.03	98 (6.1)	105 (6.5)	0.018
Vilanterol	111 (6.4)	197 (5.9)	0.019	96 (5.9)	94 (5.8)	0.005
Tiotropium	98 (5.6)	160 (4.8)	0.037	87 (5.4)	88 (5.4)	0.003
Umeclidinium	68 (3.9)	95 (2.9)	0.058	61 (3.8)	56 (3.5)	0.017
Medications for asthma						
Montelukast	148 (8.5)	211 (6.4)	0.082	118 (7.3)	118 (7.3)	0
Zafirlukast	0 (0.0)	10 (0.3)	0.078	0 (0.0)	0 (0.0)	N/A
Zileuton	0 (0.0)	0 (0.0)	N/A	0 (0.0)	0 (0.0)	N/A
Medications for cardiovascular disease						
Alpha blocker	17 (1.0)	31 (0.9)	0.004	16 (1.0)	13 (0.8)	0.02
ACEi/ARB	439 (25.3)	858 (25.9)	0.014	404 (25.0)	407 (25.2)	0.004
Beta blocker	497 (28.6)	918 (27.7)	0.02	443 (27.4)	432 (26.7)	0.015
CCB	339 (19.5)	759 (22.9)	0.083	305 (18.9)	292 (18.1)	0.021
Diuretics	510 (29.3)	1,082 (32.6)	0.071	468 (29.0)	466 (28.9)	0.003
Lipid lower agents	544 (31.3)	905 (27.3)	0.088	495 (30.7)	496 (30.7)	0.001
Antiarrhythmic agent	568 (32.7)	1,357 (40.9)	0.172	531 (32.9)	523 (32.4)	0.011
Antithrombotic agents	624 (35.9)	1,490 (44.9)	0.185	584 (36.2)	586 (36.3)	0.003
Medications for pulmonary hypertension						
Sildenafil	33 (1.9)	143 (4.3)	0.139	33 (2.0)	29 (1.8)	0.018
Tadalafil	22 (1.3)	58 (1.7)	0.04	19 (1.2)	19 (1.2)	0
Epoprostenol	10 (0.6)	23 (0.7)	0.015	10 (0.6)	10 (0.6)	0
Treprostinil	10 (0.6)	21 (0.6)	0.007	10 (0.6)	10 (0.6)	0
Iloprost	0 (0.0)	0 (0.0)	N/A	0 (0.0)	0 (0.0)	N/A
Selexipag	10 (0.6)	10 (0.3)	0.041	10 (0.6)	10 (0.6)	0
Bosentan	10 (0.6)	10 (0.3)	0.041	10 (0.6)	10 (0.6)	0
Antifibrotic agents						
Nintedanib	38 (2.2)	103 (3.1)	0.057	37 (2.3)	40 (2.5)	0.012
Pirfenidone	10 (0.6)	15 (0.5)	0.017	10 (0.6)	10 (0.6)	0

* If the patient count is ≤10, results show the count as 10. ABA, abatacept; ACEi, angiotensin converting enzyme inhibitor; ARB, angiotensin receptor blocker; CCB, calcium channel blocker; FVC, forced vital capacity; N/A, not applicable; PSM, propensity score matching; RTX, rituximab; SMD, standardized mean difference.
[a] Propensity score matching was performed on age at index, sex, race, social economic status, lifestyles, comorbidities, laboratory data, and medication usage.

differences in hospitalization and pneumonia incidence lost statistical significance, potentially due to reduced patient numbers (Table 3).

Sensitivity analysis for exclusion of patients who switched to other biologics or JAKi. To evaluate treatment efficacy without confounding from subsequent medication changes, we excluded patients who switched to another biologic or JAKi after the index date. Abatacept remained associated with lower all-cause mortality risk (HR, 0.751; 95% CI, 0.602–0.937), lower hospitalization rate (HR, 0.790; 95% CI, 0.688–0.907), and reduced risk of mechanical ventilation (HR, 0.569; 95% CI, 0.377–0.857) (Supplementary Table 11).

Sensitivity analysis for one-year period outcomes. To account for potential confounding from medication switching and patient attrition, we performed a one-year outcome analysis. The findings remained consistent, showing a significantly lower risk of all-cause mortality (HR, 0.511; 95% CI, 0.372–0.702), hospitalization (HR, 0.755; 95% CI, 0.660–0.863), and mechanical ventilation (HR, 0.565; 95% CI, 0.364–0.879) (Supplementary Table 12).

Sensitivity analysis for patients without concomitant IIM or SSc. Recognizing the potential influence of comorbidities, we conducted a sensitivity analysis limited to patients without IIM or SSc. The findings remained consistent, demonstrating a lower risk of all-cause mortality in the abatacept group compared with rituximab (HR, 0.679; 95% CI, 0.570–0.810; Supplementary Table 13).

Sensitivity analysis for patients without IIM, SSc, or malignancy. Malignancy could be an important factor

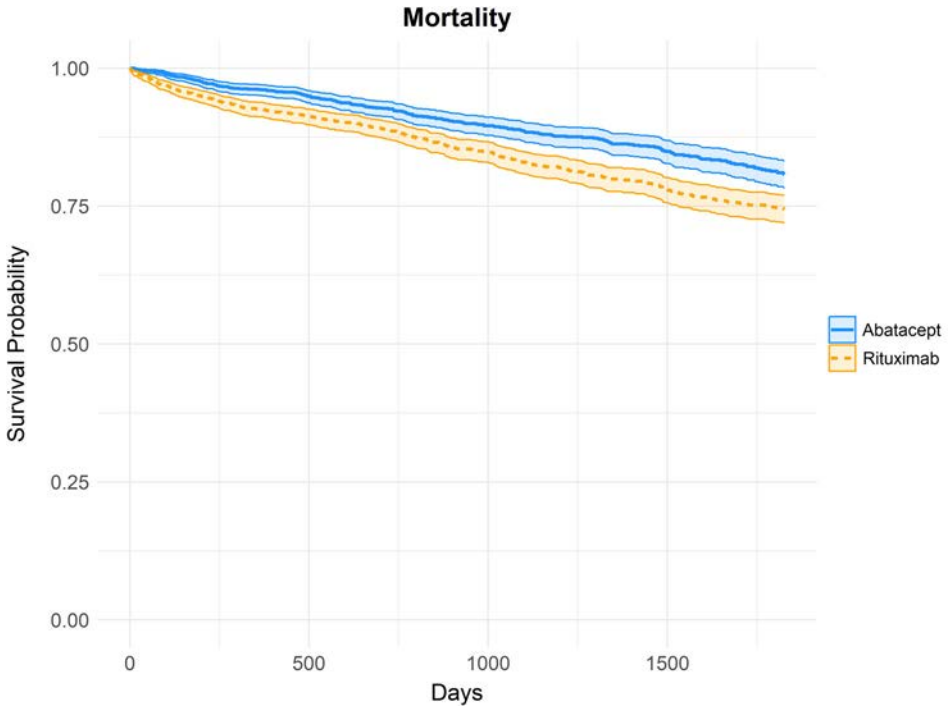


Figure 2. Kaplan–Meier Plots for all-cause mortality.

influencing selection bias, as physicians may preferentially prescribe rituximab rather than abatacept for patients with a history of malignancy. To address this potential bias, we conducted a sensitivity analysis excluding patients with baseline malignancy (ICD-10-CM codes: C00–C96). Results for all-cause mortality remained consistent with the primary analysis (HR, 0.757; 95% CI, 0.618–0.927; Supplementary Table 14), strengthening the robustness across confounding factors.

Sensitivity analysis for stringent definition of RA-ILD. Applying stricter RA-ILD diagnostic criteria (at least two incidents of RA diagnosis) continued to show consistent associations, with abatacept significantly associated with lower all-cause mortality risk (HR, 0.689; 95% CI, 0.578–0.822; Supplementary Table 15).

Sensitivity analysis for smoking history. When restricting analysis to patients with documented smoking history, abatacept maintained a significantly lower mortality risk (HR, 0.670; 95% CI, 0.514–0.874), supporting robustness across confounding factors (Supplementary Table 16).

Sensitivity analysis for concomitant secondary pulmonary hypertension. No significant differences were

Table 2. Risk of Outcomes*

Outcomes	Patients with outcome		Hazard ratio (95% CI)[a]
	ABA cohort (n = 1,615)	RTX cohort (n = 1,615)	
All-cause mortality	219	328	**0.689 (0.581–0.818)**
Medical utilization			
Hospital inpatient services	701	778	**0.882 (0.796–0.977)**
Mechanical ventilator	76	111	**0.698 (0.521–0.934)**
Cyclophosphamide use	13	79	**0.162 (0.090–0.291)**
Plasma exchange	≤10	14	0.428 (0.164–1.114)
Immunoglobulin use	31	132	**0.232 (0.157–0.343)**
Serious infection (occurrence)			
Bacteremia	57	59	0.985 (0.684–1.418)
Pneumonia	424	494	**0.858 (0.754–0.977)**

* If the patient count is ≤10, results show the count as 10. Bold font represents a standardized difference of >0.1.
ABA, abatacept; CI, confidence interval; RTX, rituximab.
[a] Propensity score matching was performed on age at index, sex, race, social economic status, lifestyles, comorbidities, laboratory data, and medication usage.

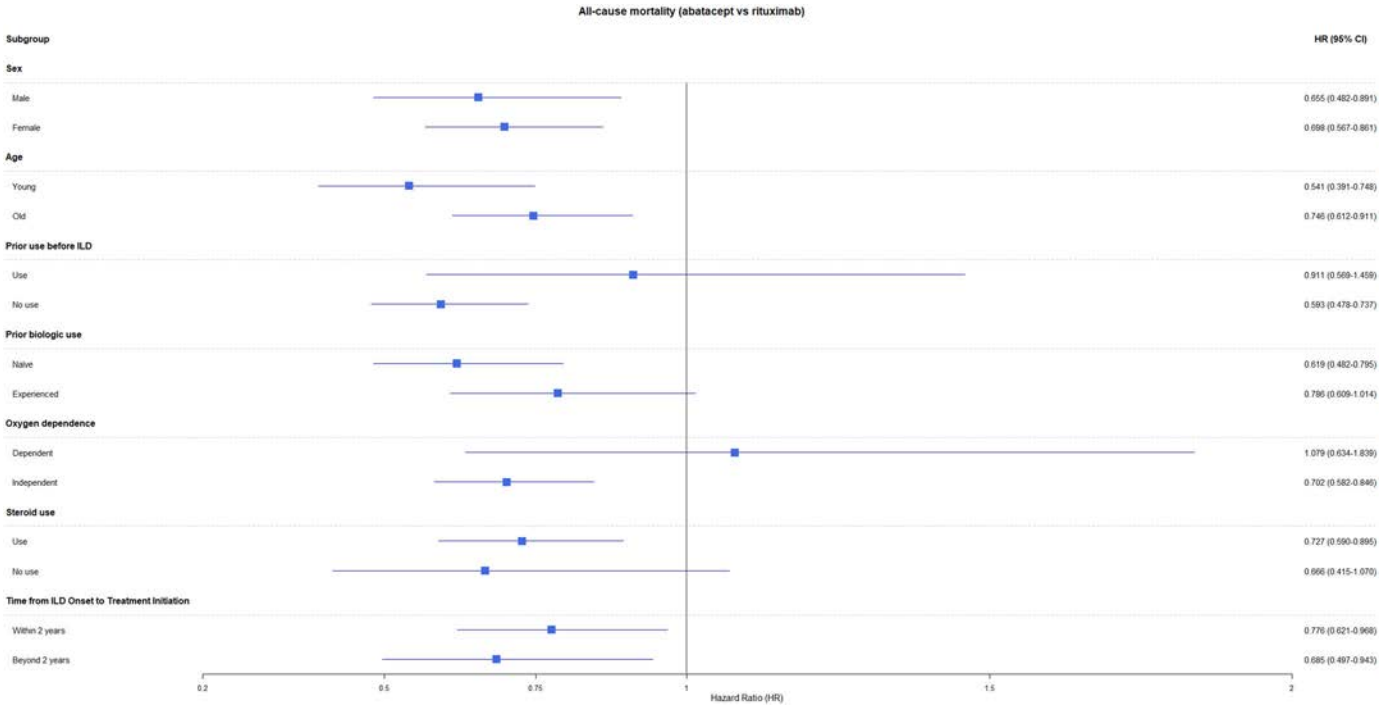


Figure 3. Forest plots for subgroup analysis (all-cause mortality). CI, confidence interval; HR, hazard ratio; ILD, interstitial lung disease. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.43332/abstract.

observed between the two cohorts, possibly due to limited statistical power arising from the small number of patients in the pulmonary arterial hypertension subgroup (Supplementary Table 17).

Sensitivity analysis for high marker of inflammation. For the population with a high marker of inflammation in baseline, the results demonstrated that abatacept was associated with comparable risk of mortality (HR, 0.828; 95%, CI 0.632–1.085) and lower risk of hospitalization (HR, 0.844; 95%, CI 0.717–0.993) and mechanical ventilation (HR, 0.618; 95%, CI 0.383–0.998). Notably, unlike in the main analysis in which abatacept was associated with reduced mortality, this effect was not observed in patients with high baseline markers of inflammation, although benefits for hospitalization and mechanical ventilation remained (Supplementary Table 18).

Sensitivity analysis for without methotrexate at baseline. In these populations, abatacept was associated with lower risk of all-cause mortality (HR, 0.807; 95% CI, 0.582–0.858), which is consistent with the main analysis (Supplementary Table 19).

Table 3. Risk of Outcomes (Add-on 6-Month Washout Period and 1-Month Stabilization Period)*

Outcomes	Patients with outcome		Hazard ratio(95% CI)[a]
	ABA cohort (n = 682)	RTX cohort (n = 682)	
All-cause mortality	86	132	**0.688 (0.524–0.902)**
Medical utilization			
Hospital inpatient services	275	294	0.951 (0.807–1.121)
Mechanical ventilator	19	35	**0.563 (0.322–0.984)**
Cyclophosphamide use	≤10	23	**0.218 (0.083–0.574)**
Plasma exchange	≤10	≤10	0.678 (0.113–4.059)
Immunoglobulin use	11	52	**0.214 (0.112–0.410)**
Serious infection (occurrence)			
Bacteremia	17	16	1.128 (0.570–2.233)
Pneumonia	155	195	0.810 (0.656–1.000)

* If the patient count is ≤10, results show the count as 10. Bold font represents a standardized difference of >0.1. Every subgroup analysis will be matched again beyond the main analysis. ABA, abatacept; CI, confidence interval; RTX, rituximab.

[a] Propensity score matching was performed on age at index, sex, race, social economic status, lifestyles, comorbidities, laboratory data, and medication usage. Excluding patients with admission and critical care visit record within 1 month before index date.

Sensitivity analysis for concomitant steroid-associated adverse events at baseline. Among patients with a history of common steroid-associated adverse events at baseline, which may reflect relatively high cumulative steroid exposure, abatacept was associated with a lower risk of all-cause mortality (HR, 0.625; 95% CI, 0.482–0.810) compared to rituximab (Supplementary Table 20).

Sensitivity analysis for study period (2007–2019). To address potential confounding by indication, the study period was restricted to 2007–2019. During this interval, abatacept remained associated with a lower risk of all-cause mortality (HR, 0.716; 95% CI, 0.569–0.899; Supplementary Table 21).

DISCUSSION

In this emulated target trial, abatacept was associated with more favorable outcomes compared to rituximab in patients with RA-ILD across most clinical scenarios evaluated. Overall, abatacept was associated with a lower risk of all-cause mortality, regardless of sex, age, interval between diagnosis and medication initiation, or study period. Subgroup and sensitivity analyses yielded results consistent with the primary analysis.

Although multiple observational studies suggest a potential benefit of abatacept in RA-ILD, reported endpoints are generally confined to stabilization or modest improvement in pulmonary function or radiographic progression. In a multicenter study of 263 patients with RA-ILD who received abatacept with a 12-month follow-up, 91.9% did not experience worsening dyspnea, 87.7% maintained stable FVC, and 90.6% preserved stable DLco.[4] However, 23.6% discontinued therapy, primarily due to adverse events (11.4%) and serious respiratory infections (n = 25).[4]

In our overall cohort of patients with RA-ILD who initiated abatacept or rituximab, the five-year all-cause mortality was 13% for abatacept and 20% for rituximab. These figures are congruent with a prior retrospective study that reported a five-year mortality of approximately 18% among rituximab-treated patients.[17] When we restricted the analysis to biologic-naïve patients, the five-year mortality remained comparable to the literature (17.5%).[17] A recently published US target trial emulation observed 73 deaths over a three-year follow-up in 237 patients receiving non-TNFi biologics (tocilizumab, sarilumab, abatacept, rituximab, tofacitinib, baricitinib, and upadacitinib). Compared with that study, our cohort had a lower proportion of methotrexate use in the year preceding the index date and excluded tocilizumab and JAKi, factors that likely contributed to the lower absolute mortality we observed. When we repeated the analysis in patients without baseline methotrexate use, mortality remained similar (abatacept: 169 of 1,261 [13.4 %] vs rituximab: 257 of 1,261 [20.4 %]; HR, 0.707; 95% CI, 0.582–0.858), supporting the consistency of our main findings. The differing treatment mixes across studies remain a potential concern when comparing mortality rates.

In the analysis restricted to abatacept-treated RA-ILD cohorts, previously reported mortality rates are of the same order of magnitude as ours. Fernández-Díaz et al documented six deaths among 273 patients (2.2%) during a mean follow-up of 22 months, with only two deaths attributed to ILD progression, both events occurring after abatacept had been withdrawn.[4] By comparison, our study recorded 58 deaths among 1,615 abatacept users within the first year of follow-up (3.6%). Although our cohort is larger and the observation window differs, the absolute early mortality percentages are broadly comparable, reinforcing the consistency of abatacept-associated outcomes across diverse settings.

Confounding by indication was the principal methodologic concern in retrospective design, because abatacept is considered for patients with comparatively milder RA-ILD. To mitigate this bias, we conducted a 1:1 propensity-score–matched cohort that simultaneously balanced socioeconomic status, systemic inflammatory activity reflected by the ESR, baseline oxygenation, pulmonary function (FVC and DLco) together with the presence of pulmonary hypertension, and baseline therapy, including systemic glucocorticoids, conventional and biologic disease modifying antirheumatic drugs (DMARDs), inhaled respiratory medications, pulmonary vasodilators, and antifibrotic agents as the proxy of disease severity. This comprehensive matching strategy struggled to ensure that disease severity, access to care, and relevant concomitant treatments were comparable between the abatacept and rituximab groups, thereby minimizing residual confounding.

To deal with residual confounding, we undertook a series of prespecified subgroup analyses. First, we stratified patients by biologic exposure history. Among bio-naïve individuals, abatacept was associated with lower all-cause mortality than rituximab, whereas in the bio-experienced stratum, reflecting potentially more refractory disease, the difference did not reach statistical significance. Second, we examined proxies of ILD severity at treatment initiation. Separate models were run for patients who (1) required supplemental oxygen around the index date and (2) did not receive systemic glucocorticoids within ±30 days of index. Owing to limited sample size, neither comparison demonstrated a significant treatment effect, yet point estimates were directionally consistent with the primary analysis. Third, we assessed the impact of timing of biologic initiation relative to ILD diagnosis. Patients who started abatacept or rituximab within two years of their first ILD diagnosis (early initiators) could potentially represent a cohort with more rapidly progressive or highly active pulmonary disease, whereas those treated beyond two years after diagnosis might have more indolent ILD. Mortality did not differ meaningfully between early and late initiators, suggesting that the observed treatment effect is robust to variation in ILD duration at the time of biologic initiation.

To minimize residual confounding, we conducted a suite of sensitivity analyses. We applied an "as-treated" restriction by censoring patients who switched to another biologic or targeted synthetic DMARD after the index date. Furthermore, we eliminated individuals with IIM, SSc, or malignancy, comorbidities associated with poorer prognosis and a clinical preference for rituximab. Diagnostic specificity was strengthened by requiring two or more RA codes before enrollment. We then repeated the analysis within a high-severity subset, characterized by markedly elevated markers of inflammation, secondary pulmonary hypertension, or prior steroid-related adverse events. Finally, recognizing that methotrexate is often withheld in advanced ILD, we excluded baseline patients treated with methotrexate.[18] Large-scale studies and formal guidelines on using abatacept and rituximab in RA-ILD have only recently emerged. The British Society for Rheumatology published its first recommendations in 2019, the joint Spanish Rheumatology-Pulmonology guideline appeared in 2022, and the ACR released an interdisciplinary systemic autoimmune rheumatic disease–ILD guideline in 2023.[1,19,20] Accordingly, we limited our study period to 2007–2019 to examine whether outcomes differed before versus after the earliest of these recommendations. Across all iterations, effect estimates remained directionally consistent and of similar magnitude to the primary analysis, reinforcing the robustness of the observed abatacept survival benefit.

This study has several methodologic strengths that enhance the validity and generalizability of its findings. First, we employed a target trial emulation using a large-scale, real-world dataset (TriNetX US Collaborative Network), which includes a diverse patient population from 71 health care organizations across the United States. This large sample size, currently one of the largest in RA-ILD treatment comparative studies, significantly enhances statistical power and clinical practice. Second, PSM was applied to more than 70 baseline covariates, achieving SMDs below 0.1, thereby minimizing potential confounding. Third, extensive sensitivity analyses, including washout periods, assessment of biologic switching, exclusion of confounding comorbidities, malignancy, and application of stringent RA-ILD definitions, consistently supported our main findings, reinforcing the reliability of the results. Fourth, comprehensive subgroup analyses (eg, sex, age, biologic-naïve vs experienced, and oxygen dependence) demonstrated consistent treatment associations across clinically relevant patient populations, further enhancing the generalizability and clinical relevance. Lastly, with external causes of morbidity as a negative control outcome, the results effectively addressed residual confounding concerns, reinforcing confidence that observed differences were attributable to treatment strategies rather than baseline imbalances or confounding biases.

However, several limitations warrant consideration. First, despite rigorous eligibility criteria and PSM across more than 70 covariates, residual confounding due to unmeasured factors may still exist, including detailed smoking history, RA disease activity measures, and ILD severity based on imaging and pulmonary function trajectories. Importantly, despite comprehensive PSM, residual baseline differences in comorbidity severity might persist, because rituximab is more frequently prescribed to patients with greater disease complexity or higher comorbidity burden. To mitigate this, we conducted comprehensive sensitivity analyses and subgroup analyses, consistently reinforcing our main findings. However, residual differences in baseline disease severity cannot be eliminated by observational methods alone. Second, despite emulating a target trial, balancing >70 baseline variables, and confirming robustness in multiple subgroups and sensitivity analyses, we cannot fully eliminate confounding by indication. In routine practice, clinicians may preferentially prescribe abatacept to patients perceived to have milder or less progressive RA-ILD, potentially biasing the observed survival advantage. Accordingly, our findings should be interpreted as associations rather than definitive evidence of causality. Third, diagnostic misclassification remains possible, as ILD diagnoses relied on ICD-10 codes without direct validation from radiologic or complete pulmonary function test results with pulmonologist. To address this limitation, we performed sensitivity analyses employing more stringent diagnostic criteria, maintaining consistent results. Fourth, adherence to medication regimens could not be confirmed directly through prescription data alone. However, sensitivity analyses excluding patients who switched to other biologics or JAKi after initiation provided robust findings, possibly indicating observed effects likely reflect true differences between abatacept and rituximab treatment. Fifth, our database lacked detailed information regarding continuous patient enrollment and health care coverage changes, potentially leading to incomplete follow-up data. Although we mitigated this by requiring patients to have at least two follow-up visits, this limit still exists. Sixth, detailed data on disease severity, such as clinical medical record and FVC percentage predicted, were unavailable, although we incorporated proxies including DL_{CO} predicted, ESR, and blood gas parameters to partially compensate. Missing data, inherent to real-world databases, limit comprehensive severity assessment. Seventh, the TriNetX platform does not provide person-years incidence rates or detailed at-risk counts required for optimal Kaplan–Meier plots. Although this limitation restricts certain analyses, HRs still reliably represent treatment effects. Eighth, due to the limitation in platform, stratified Cox regression analyses were not available in comparing analysis, limiting direct statistical control for baseline heterogeneity. Nevertheless, extensive subgroup analyses helped confirm consistent effectiveness and safety across clinically relevant patient subgroups. Ninth, retrospective observational designs inherently limit causality inferences compared with randomized controlled trials. However, our target trial emulation enhances causal inference beyond typical observational studies in our best, closely approximating randomized trial conditions. Lastly, our study population was derived solely from a US-based health care database, potentially limiting generalizability

internationally. The differences in genetic backgrounds, disease severity, and health care systems worldwide may interfere with outcomes. Future international multiorganization studies are needed to validate the generalizability of these findings globally.

This large, real-world target trial emulation strengthens the evidence that abatacept may offer a more favorable outcome than rituximab in RA-ILD, with consistently lower risks of all-cause mortality across the subgroup and sensitivity analyses. By providing the first head-to-head comparison of these two biologics, the study provides a critical evidence gap in RA-ILD management. Nevertheless, residual confounding by indication cannot be fully excluded, most notably the likelihood that clinicians reserve abatacept for patients with less aggressive diseases, so the observed association should not be interpreted as definitive causality. Prospective randomized controlled trials are still required to confirm whether abatacept leads to a true survival advantage and to guide future treatment recommendations.

AUTHOR CONTRIBUTIONS

Conceptualization: PCS, JCCW, QWW. Formal analysis: SIW. Resources & Data Curation: SIW, PCS. Interpretation of the data: SIW, PCS. Writing –original draft: PCS. Writing – review & editing: all authors reviewed and edited the manuscript. All authors read and approved the final manuscript. As corresponding author, Dr Wei confirms that all authors have provided the final approval of the version to be published and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. Johnson SR, Bernstein EJ, Bolster MB, et al. 2023 American College of Rheumatology (ACR)/American College of Chest Physicians (CHEST) guideline for the treatment of interstitial lung disease in people with systemic autoimmune rheumatic diseases. Arthritis Rheumatol 2024;76(8):1182–1200.
2. Dieudé P, Sparks J, Fischer A, et al. POS1050 incidence rates of interstitial lung disease among patients with rheumatoid arthritis treated with abatacept: a post hoc pooled analysis from a compendium of clinical trials. Ann Rheum Dis 2023;82:843.
3. Kang EH, Jin Y, Desai RJ, et al. Risk of exacerbation of pulmonary comorbidities in patients with rheumatoid arthritis after initiation of abatacept versus TNF inhibitors: a cohort study. Semin Arthritis Rheum 2020;50(3):401–408.
4. Fernández-Díaz C, Castañeda S, Melero-González RB, et al. Abatacept in interstitial lung disease associated with rheumatoid arthritis: national multicenter study of 263 patients. Rheumatology (Oxford) 2020;59(12):3906–3916.
5. Atienza-Mateo B, Fernández-Díaz C, Vicente-Rabaneda EF, et al; Spanish Collaborative Group of Abatacept in Interstitial Lung Disease Associated with Rheumatoid Arthritis. Abatacept in usual and in non-specific interstitial pneumonia associated with rheumatoid arthritis. Eur J Intern Med 2024;119:118–124.
6. Shoda T, Kotani T, Koyama M, et al. The therapeutic efficacy of abatacept for rheumatoid arthritis-associated interstitial lung disease: insights from a 12-month trial using semi-quantitative chest high-resolution computed tomography imaging. J Clin Med 2024;13(19):5871.
7. López-Maraver M, Serrano-Combarro A, Atienza-Mateo B, et al. Subcutaneous vs intravenous abatacept in rheumatoid arthritis-interstitial lung disease. National multicentre study of 397 patients. Semin Arthritis Rheum 2024;68:152517.
8. Malek S, Stein E, Swartzbaugh M, et al. RWD68 The TriNetX linked EHR and administrative claims network. Value Health 2024;27(6):S370.
9. Ludwig RJ, Anson M, Zirpel H, et al. A comprehensive review of methodologies and application to use the real-world data and analytics platform TriNetX. Front Pharmacol 2025;16:1516126.
10. Anand P, Zhang Y, Merola D, et al. Comparison of EHR data-completeness in patients with different types of medical insurance coverage in the United States. Clin Pharmacol Ther 2023;114(5):1116–1125.
11. Curtis JR, Xie F, Zhou H, et al. Use of ICD-10 diagnosis codes to identify seropositive and seronegative rheumatoid arthritis when lab results are not available. Arthritis Res Ther 2020;22(1):242.
12. Lee H, Sparks JA, Lee SB, et al. Validation of serostatus of rheumatoid arthritis using ICD-10 codes in administrative claims data. Pharmacoepidemiol Drug Saf 2023;32(5):586–591.
13. Meehan M, Shah A, Lobo J, et al. Validation of an algorithm to identify incident interstitial lung disease in patients with rheumatoid arthritis. Arthritis Res Ther 2022;24(1):2.
14. Brooks R, Baker JF, Yang Y, et al. The impact of disease severity measures on survival in U.S. veterans with rheumatoid arthritis-associated interstitial lung disease. Rheumatology (Oxford) 2022;61(12):4667–4677.
15. Nonaka F, Fukui S, Michitsuji T, et al. The impact of glucocorticoid use on the outcomes of rheumatoid arthritis in a multicenter ultrasound cohort study. Int J Rheum Dis 2024;27(3):e15118.
16. Su YJ, Lin CY, Hsu CY. Analyzing the risk of osteoporosis and fracture in rheumatoid arthritis patients who have been treated with various biologics. Int J Rheum Dis 2024;27(2):e15055.
17. Druce KL, Iqbal K, Watson KD, et al. Mortality in patients with interstitial lung disease treated with rituximab or TNFi as a first biologic. RMD Open 2017;3(1):e000473.
18. Park E, Iqbal R, Giles JT, et al. Use of methotrexate and TNF inhibitors in patients with rheumatoid arthritis-associated interstitial lung disease: a survey of rheumatologists. Clin Rheumatol 2024;43(9):3029–3032.
19. Holroyd CR, Seth R, Bukhari M, et al. The British Society for Rheumatology biologic DMARD safety guidelines in inflammatory arthritis. Rheumatology (Oxford) 2019;58(2):e3–e42.
20. Narváez J, Díaz Del Campo Fontecha P, Brito García N, et al. SER-SEPAR recommendations for the management of rheumatoid arthritis-related interstitial lung disease. Part 2: treatment. Reumatol Clin (Engl Ed) 2022;18(9):501–512.

Arthritis & Rheumatology
Vol. 78, No. 3, March 2026, pp 603–612
DOI 10.1002/art.43424

American College of Rheumatology
Empowering Rheumatology Professionals

Distinct HLA Associations for Antibody Multireactivity With Citrulline-Containing Type II Collagen Epitopes Versus More Limited Antibody Reactivity With Citrulline-Containing IgG Epitopes in Rheumatoid Arthritis

S. Janna Bashar,[1] Courtney B. Myhr,[1] Adam H. Titi,[1] Zihao Zheng,[1,2] and Miriam A. Shelef[1,3]

Objective. Anticitrullinated protein antibodies (ACPAs) in rheumatoid arthritis (RA) can be promiscuous, with cross-reactive binding to many antigens containing short motifs, or private with little cross-reactivity. Also, ACPA reactivity patterns differ among patients with RA, including for motif-containing epitopes in important self-antigens like collagen and IgG (bound by RA-associated rheumatoid factors [RFs]), with limited understanding of the underlying mechanism. The objective of this study was to determine if HLA alleles associate with ACPA reactivity patterns.

Methods. For 100 ACPA+RF+ participants with RA, serum IgG binding was quantified by enzyme-linked immunosorbent assay to 10 citrulline-containing peptides derived from Type II collagen and IgG1 (nine with motifs), and HLA loci were genotyped. Also, antibody and serum multireactivity were evaluated. HLA alleles present differentially in RA participants with high versus low IgG binding to specific peptides, as well as with multireactivity versus limited reactivity were identified by Fisher's exact test.

Results. Serum IgG multireactivity for citrulline-glycine motif-containing collagen peptides was high, at least partially due to promiscuous antibodies. HLA-DQA1*01:02 was present in more participants with anticitrullinated collagen antibodies and multireactive sera. In contrast, serum multireactivity was low for IgG1-derived peptides due at least in part to more private antibodies. Shared epitope–containing HLA-DRB1*04:01 was present more frequently in participants with RA-associated RFs irrespective of the citrulline-serine motif and less frequently in participants with anticitrullinated collagen antibodies. Several HLA alleles associated with specific antibody reactivities.

Conclusion. Different HLA alleles may contribute to the different reactivity patterns of promiscuous anticitrullinated collagen antibodies and more private RA-associated RFs.

INTRODUCTION

Anticitrullinated protein antibodies (ACPAs) are highly specific markers of rheumatoid arthritis (RA), a chronic autoimmune and inflammatory disease. Since their initial discovery over two decades ago,[1] much has been learned about how ACPAs bind an enormous repertoire of antigens. Many ACPAs are actually multireactive antimodified protein antibodies that can recognize thousands of citrullinated (arginines converted to citrullines), homocitrullinated (lysines converted to homocitrullines), and other antigens.[2] The multireactivity of at least some ACPAs appears to be due to the recognition of very short motifs like citrulline-glycine, irrespective of protein identity.[2] In contrast, other ACPAs are more selective and bind a limited number of antigens, leading to the labels "promiscuous" and "private."[3] These different ACPAs bind a variety of proteins, such as fibrin,[4] histones,[5] and vimentin,[6] some with important roles in RA pathogenesis, in either a specific or nonspecific manner. The factors that underlie the development of different ACPAs are still emerging yet are central to understanding RA pathophysiology.

One important autoantibody target in RA is Type II collagen, an articular cartilage protein. Injection of Type II collagen with

Supported by US Army Medical Research Acquisition Activity through the Peer Reviewed Medical Research Program (grant W81XWH-18-1-0717) to Dr Shelef.

[1]Department of Medicine, University of Wisconsin–Madison; [2]Department of Statistics, University of Wisconsin–Madison; [3]William S. Middleton Memorial Veterans Hospital, Madison, Wisconsin.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/art.43424).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/art.43424.

Address correspondence via email to Miriam A. Shelef, MD, PhD, at mshelef@medicine.wisc.edu.

Submitted for publication March 13, 2025; accepted in revised form September 30, 2025.

adjuvant drives inflammatory arthritis in mice,[7] and anticollagen antibodies induce arthritis in naive mice.[8] Antibodies against native Type II collagen are also found in human RA,[9] and at least some of those anticollagen antibodies bind the same conformational native epitope as a pathogenic murine anticollagen antibody.[10] People with RA also generate autoantibodies against citrullinated collagen, even more so than against native collagen.[11] Interestingly, IgG in RA can recognize cyclic citrulline–containing collagen-derived peptides, particularly when a citrulline-glycine multireactivity motif is present,[12] whereas murine anticollagen antibodies bind to conformational native epitopes.[13] Of note, RA antibodies that react with cyclic citrulline–containing collagen peptides also bind to arthritic joints.[14] However, not everyone with RA generates these antibodies, and there is a wide variety of IgG-binding patterns to different collagen peptides, even when citrulline-glycine is present.[12] The mechanisms underlying these differences are poorly understood.

IgG is another important autoantibody target in RA. Rheumatoid factors (RFs), which recognize the Fc region of IgG, bind ACPA immune complexes to enhance macrophage-mediated inflammation[15] and complement-mediated inflammation.[16] RFs are canonically known to bind two conformational epitopes of IgG,[17,18] but RFs also have unique disease-specific IgG reactivities and multireactivities.[19–22] For example, whereas COVID-19–associated RFs bind IgG and pathogen epitopes based upon an aspartic acid–dominated tripeptide motif, RA-associated RFs, whose reactivity overlaps with multireactive ACPAs, bind citrulline-containing and homocitrulline-containing linear IgG epitopes, potentially due in part to a citrulline-serine motif.[20–23] Like promiscuous and private ACPAs, RFs have varying degrees of multireactivity.[24] Also, similar to anticitrullinated collagen antibodies, RA sera have differing patterns of RA-associated RF binding to IgG epitopes despite the presence of citrulline-serine,[20–23] with no clear underlying mechanism.

One factor that may underlie the development of antibodies with different reactivities is the HLA. RA risk and ACPAs are most famously linked to HLA alleles with the shared epitope, a five amino acid consensus sequence found in HLA-DRB1.[25,26] Shared epitope–containing HLAs have been reported to have preferential affinity for citrulline-containing peptides, although variability exists among HLAs and peptides.[27–31] Also, HLA alleles without the shared epitope, including HLA-DQ and HLA-DP alleles, as well as amino acids outside of the shared epitope region have been associated with RA, ACPAs, and the presentation of citrulline-containing peptides.[32–36]

Some work has been done to evaluate HLA alleles in the context of specific antibodies as well. Because RA-associated RFs are newly identified, no HLA studies have been performed. However, antibodies against Type II collagen were found to be associated with HLA-DRB1*01 and HLA-DRB1*03, but the specific anticollagen antibodies and the specific HLA alleles are unknown.[37] Also, HLA-DRB1*04 was found to be associated with private antibodies against a native Type II collagen triple helical peptide, a citrulline-containing α-enolase peptide, and citrullinated fibrinogen.[38] However, neither specific HLA alleles nor multireactive antibodies were described. These findings provide evidence of a role for HLA in shaping the RA autoantibody repertoire, but mysteries remain regarding how HLA alleles relate to specific antibodies and multireactivity. Thus, the objective of this study was to determine if specific HLA alleles associate with specific anticitrullinated collagen antibodies, RA-associated RFs, and/or multireactivity to better understand the development of pathogenic antibodies in RA.

MATERIALS AND METHODS

Human subjects. Human subjects research was approved by the University of Wisconsin Institutional Review Board and was conducted in compliance with the Declaration of Helsinki. All human subjects provided written informed consent. Sera and DNA from subjects diagnosed with RA by a rheumatologist and with anti–cyclic citrullinated peptide and RF levels more than two times the upper limit of normal as well as sera from control subjects without known autoimmune or inflammatory disease as previously defined were obtained from the University of Wisconsin Rheumatology Biorepository.[39,40] Participants with RA are described in Supplementary Table 1.

Enzyme-linked immunosorbent assay. Enzyme-linked immunosorbent assay (ELISA) was performed as previously described[21] with minor modifications. ELISA plates were incubated with a 0.1 μM concentration of peptides (Supplementary Table 2), incubated with sera diluted at 1:500 for collagen Type II α1 (COL2A1) and 1:200 for IgG1 peptides, and IgG was detected with mouse anti-human IgG (clone JDC-10, SouthernBiotech). RA sera with IgG binding values >1 SD above mean control IgG binding to each citrulline-containing peptide (marked with a B to represent the amino acid citrulline) were defined as "high binding" or "positive" versus all other sera defined as "low binding" or "negative."

HLA genotyping. Genotyping of HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DQB1, HLA-DRB1, and HLA-DRB345 alleles was performed using next-generation sequencing by CD Genomics.

Antibody purification. Antibodies that bind to COL2A1-885B and IgG1-289B were purified and used in ELISA as previously,[20] with modifications. Briefly, streptavidin-sepharose beads (Abcam, Waltham, MA) were incubated with an equal volume of 0.02 mM biotinylated peptide, washed, and loaded onto a column. Sera diluted 1:4 were added and incubated overnight at 4°C; columns were then washed, and antibody was eluted from the beads and concentrated using Amicon Ultra 30 kDa

molecular weight cut-off filters (MilliporeSigma, Burlington, MA). Antibody concentration was determined by ELISA using standard curves generated from purified IgG (Novus Biologicals, Centennial, CO). Purified antibodies were tested at 20 ng/mL to determine multireactivity.

Statistics. IgG-binding levels were compared by Kruskal-Wallis test with Dunn's multiple comparisons test for more than two groups and a Mann-Whitney test for two groups. The frequency of the presence of at least one copy of each unique HLA allele was compared for high and low binding groups by Fisher's exact test. All analyses were performed using Prism (GraphPad Software, San Diego, CA), and $P < 0.05$ was considered significant.

RESULTS

We began by evaluating IgG binding in seropositive RA to all possible 12 amino acid linear peptides from COL2A1 in native, citrulline-containing, and homocitrulline-containing forms using data from two previously published high-density peptide arrays.[23,41] We observed very high binding to many citrulline-containing peptides and more modest binding to homocitrulline-containing peptides (Supplementary Figure 1), consistent with previous results for these subjects and linear peptides derived from other proteins.[21,23] Also, the highest bound peptides overwhelmingly contained citrulline-glycine, suggesting that this multireactivity motif is a major factor in anticitrullinated collagen reactivity.

To better understand the differential reactivities of potentially pathogenic anticollagen antibodies that target citrulline-glycine, we honed in on five linear COL2A1 peptides that contained citrulline-glycine, were highly bound by RA IgG in the array experiments, and had homologous epitopes in mice bound by arthritogenic autoantibodies when the peptides were in their triple helical form.[10,13,42] In ELISA, we saw, as expected, significantly higher IgG binding in RA to the citrulline-containing form than the native form of all five COL2A1 peptides, as well as higher binding in RA than in controls (Figure 1A). Also, similar to antibody binding to cyclic citrulline–containing collagen-derived peptides,[12] different RA sera demonstrated different degrees and patterns of IgG reactivity to the linear citrulline-containing peptides despite the presence of citrulline-glycine (Figure 1B). However, in general, serum multireactivity was high with >50% of sera with reactivity to any collagen peptide demonstrating reactivity against all five peptides (Supplementary Figure 2). Sera with anti–COL2A1-751B antibodies had the highest level of multireactivity. Among sera with these antibodies, 88% had IgG that bound four or five peptides.

We also evaluated IgG binding to five IgG1-derived citrulline-containing peptides that are known to be bound at high levels in seropositive RA, four of which have citrulline-serine.[20–22] As with COL2A1, RA sera had significantly more binding to citrulline-containing IgG1 peptides than control sera, with variability among sera in both the degree and pattern of reactivity (Figure 2). In contrast to COL2A1, serum multireactivity was low for the citrulline-containing IgG1-derived peptides (Supplementary Figure 2). Among sera with IgG that bound any IgG1 peptide, only approximately 10% had IgG that bound all five IgG1 peptides and approximately one-third had IgG that bound four to five peptides. Sera with any reactivity with IgG1-derived peptides most commonly bound two or three IgG1 peptides, as opposed to five peptides for collagen. However, it should be noted that the higher and more variable RA-associated RF levels in controls raises the cut-off for positivity in RA, which may underestimate RA-associated RF presence in our cohort.

Nonetheless, given the differences in multireactivity among sera for collagen- and IgG1-derived peptides, we evaluated antibody promiscuity. Antibodies that bind COL2A1-885B were elevated in all 40 of the RA sera with reactivity to at least four collagen peptides as well as sera with more limited reactivity. Also, COL2A1-885B has only one citrulline and thus likely only one epitope recognized by purified ACPAs. Therefore, we purified antibodies that bound to COL2A1-885B from three subjects with reactivity to all 5 peptides and three subjects with reactivity to 2 peptides and evaluated IgG binding of the anti-COL2A1-885B antibodies to all 10 COL2A1 and IgG1 peptides. As shown in Figure 3A, anti–COL2A1-885B antibodies from multireactive sera bound all five collagen peptides but not any IgG1 peptides despite serum multireactivity with IgG1-derived peptides. The binding of anti–COL2A1-885B antibodies from limited reactivity sera primarily mirrored serum reactivity but not quite as closely as for the multireactive sera.

We performed similar experiments evaluating multireactivity with IgG1 peptides (Figure 3B). As noted above, serum multireactivity was low for IgG1-derived peptides. However, anti–IgG1-289B antibodies were present in all but 1 of the 13 RA serum samples with multireactivity to four or five peptides. Moreover, anti–IgG1-289B antibodies were present at sufficiently high levels for antibody purification experiments, and IgG1-289B contains only a single citrulline, a combination of features not present for the other four peptides. Thus, we purified anti–IgG1-289B antibodies from sera that bound 4 to 5 IgG1 peptides or 2 IgG1 peptides and evaluated IgG binding to all 10 peptides. In contrast to the anti–COL2A1-885B results, anti–IgG1-289B antibodies purified from two of three multireactive sera had no cross-reactivity with the other peptides. Anti–IgG1-289B antibodies purified from limited reactivity sera more closely mirrored serum reactivity. Taken together, these data suggest that promiscuous antibodies contribute to high serum multireactivity with citrulline-glycine–containing collagen-derived peptides, whereas multiple, more private antibodies may underly the low serum multireactivity to IgG1-derived peptides.

We then evaluated if the presence of IgG that binds to specific peptides might correlate with specific HLA alleles. We

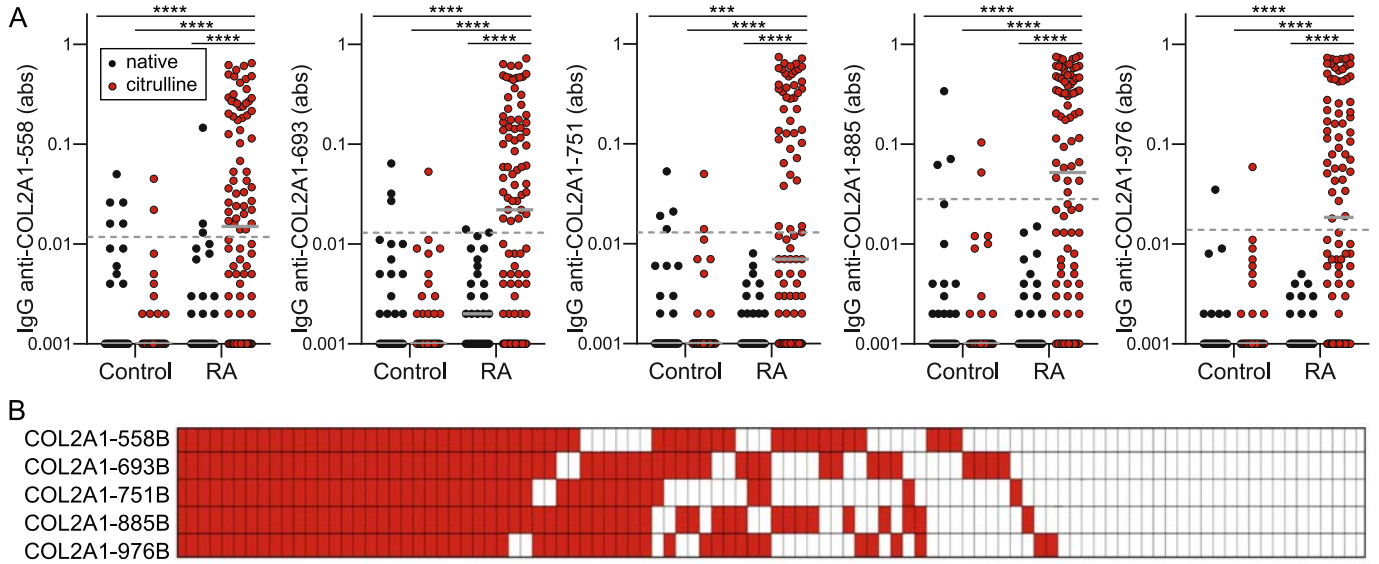


Figure 1. High, yet variable, IgG binding to citrulline-containing Type II collagen-derived linear peptides in RA. (A) IgG binding measured by enzyme-linked immunosorbent assay for RA (n = 32) and control sera (n = 32) to five native linear COL2A1 peptides and IgG binding of control sera (n = 32) to the citrulline-containing form of the peptides were compared to IgG binding for RA sera (n = 100) to the citrulline-containing peptides by Kruskal-Wallis with Dunn's multiple comparisons test (***$P < 0.001$; ****$P < 0.0001$; gray bars indicate medians; dashed lines indicate cut-offs for high IgG binding 1 SD above the mean for control IgG binding to citrulline-containing peptides). (B) Heatmap of positive binding (red) for each participant with RA (column) and peptide (row). abs, absorbance; B, citrulline; COL2A1, collagen Type II α1; RA, rheumatoid arthritis Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.43424/abstract.

genotyped the HLA loci of the participants with RA and then determined if specific alleles were more highly represented in participants with RA with high or low binding for each peptide. For collagen (Table 1), we identified DQA1*01:02 as present in a higher percentage of RA participants with high IgG binding to COL2A1-751B versus low IgG binding to that peptide. In contrast, DRB3*02:02, DQA1*05:05, and shared epitope–containing DRB1*04:01 were present in more RA participants with low IgG binding to COL2A1-885B, COL2A1-976B, and COL2A1-693B, respectively.

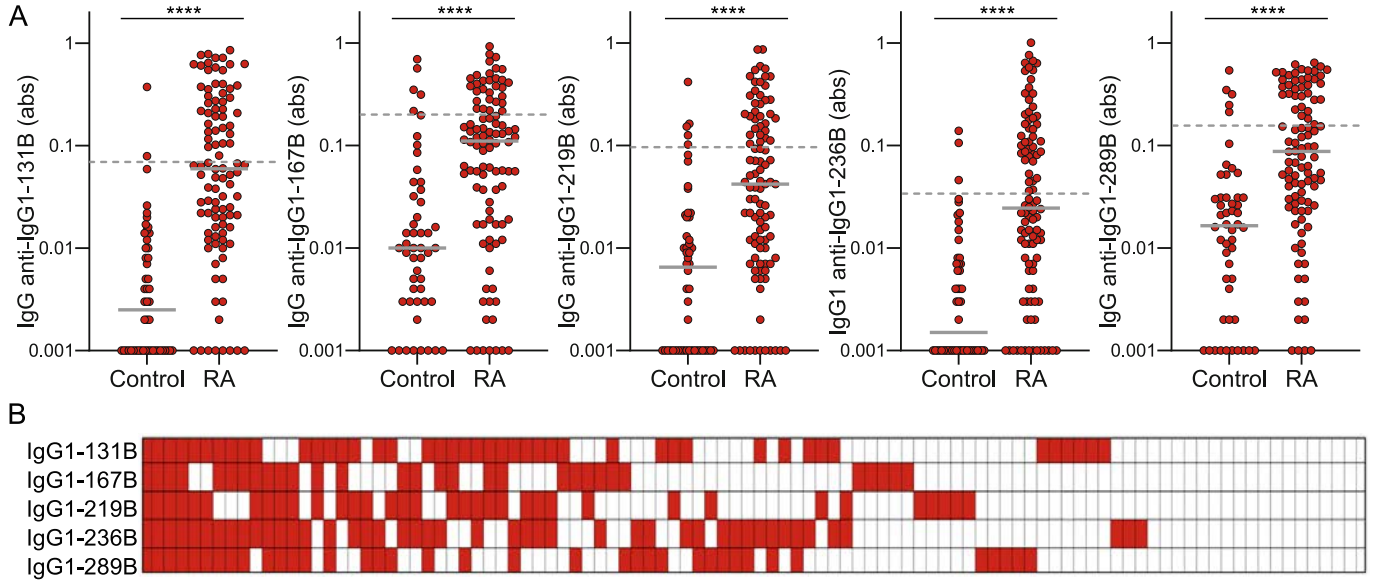


Figure 2. Increased, but variable, IgG binding to citrulline-containing IgG1 peptides in RA. (A) IgG binding measured by enzyme-linked immunosorbent assay for control (n = 50) versus RA sera (n = 100) to citrulline-containing IgG1-derived peptides was evaluated by Mann-Whitney test (****$P <$ 0.0001; gray bars indicate medians; dashed lines indicate cut-offs for high IgG binding 1 SD above the mean for control IgG binding). (B) Heatmap of positive binding (red) for each participant with RA (column) and peptide (row). abs, absorbance; B, citrulline; RA, rheumatoid arthritis. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.43424/abstract.

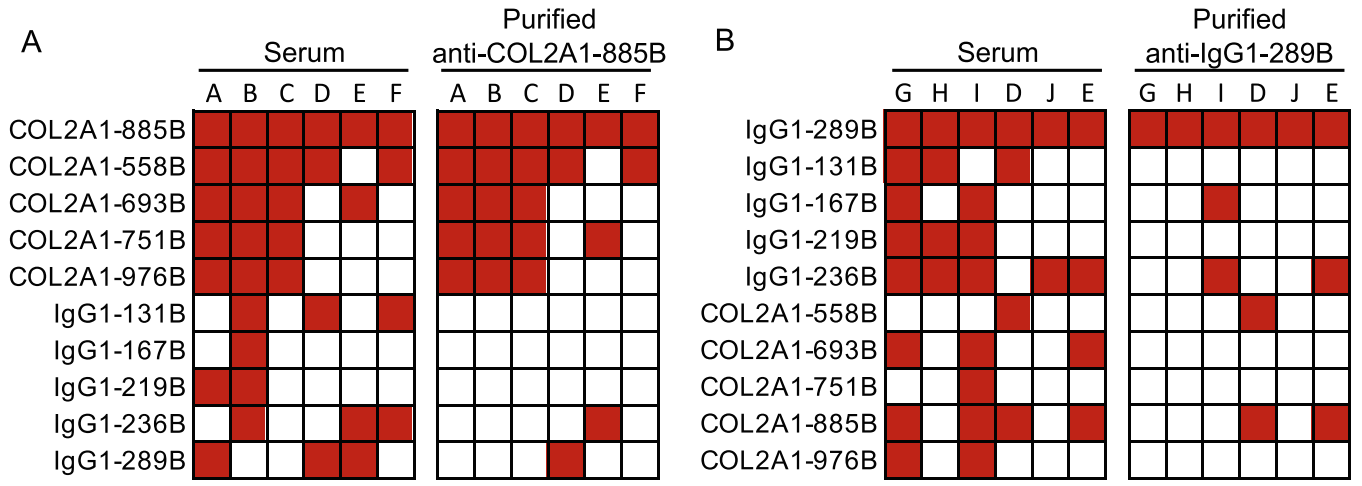


Figure 3. Rheumatoid arthritis antibody and serum multireactivity for Type II collagen- and IgG1-derived peptides. (A) Antibodies that bind the COL2A1-885B peptide were purified from three sera multireactive for all five collagen peptides and three sera with IgG reactivity to only two collagen peptides. IgG anti–COL2A1-885B binding to the five collagen and five IgG peptides was quantified by enzyme-linked immunosorbent assay. (B) Antibodies that bind to the IgG1-289B peptide were purified from sera with reactivity to four or five IgG1 peptides or reactivity to two IgG1 peptides and then tested in enzyme-linked immunosorbent assay for binding to all 10 peptides as in (A). Letters indicate specific participants. Positive results for serum reactivity according to previously noted cut-offs, and binding values >5% of the peptide used to purify the tested antibody are depicted in red. COL2A1, collagen Type II α1. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.43424/abstract.

For IgG binding to IgG1-derived peptides (Table 2), we found that shared epitope–containing DRB1*04:01 is present in more RA participants with high IgG binding to IgG1-167B, whereas DPB1*03:01 is present in more participants with low binding to that peptide. Also, DQA1*01:05 and shared epitope–containing DRB1*10:01 were present more frequently in high IgG-binding subjects, and DPB1*04:02 in more low IgG-binding subjects for IgG1-219B. However, all participants who had DRB1*10:01 also had DQA1*01:05 and vice versa, making differentiating between those two alleles for analysis impossible.

Given the multireactivity observed for collagen peptides, we were surprised that each HLA variant associated with antibodies for only a single collagen peptide, although the only positive association (DQA1*01:02; Table 1) was with the antibody that had the highest level of serum multireactivity (IgG anti–COL2A1-751B; Supplementary Figure 2). Because our sample size was only 100, we hypothesized that patterns of associations might be revealed if we looked at trends, not just significant *P* values. So, for each HLA variant significantly associated with antibodies that recognized any peptide, we created a heatmap of association trends for all 10 peptides. As shown in Figure 4, participants with IgG that binds any of the five collagen peptides show at least a trend toward more frequent presence of DQA1*01:02 and less frequent presence of DRB3*02:02. DQA1*05:05 had similar results as DRB3*02:02 but for only four of five peptides. Less uniformity was seen for the IgG1-derived peptides, with no HLA demonstrating even a trend toward association with antibodies against all five peptides. However, DRB1*04:01 had at least a trend toward increased presence in participants with IgG that binds IgG1-131B, IgG1-167B, or IgG1-236B. Combined with the positive association of DRB1*10:01 with IgG anti–IgG1-219B, at least a trend toward more frequent presence of shared epitope–containing HLA exists for participants with RA-associated RFs. Of note, the negative association of DRB1*04:01 with IgG that binds citrulline-containing collagen peptides was not uniform across citrulline-glycine–containing peptides.

Next, we evaluated the relationship between HLA and serum multireactivity. We compared HLA presence for participants with IgG binding to multiple peptides versus limited or no binding. For

Table 1. Association of HLA alleles with IgG that binds specific Type II collagen peptides*

Peptide	HLA allele	Allele frequency in high binders	Allele frequency in low binders	*P* value[a]
COL2A1-693B, n/N (%)	DRB1*04:01	19/55 (35)	25/45 (56)	0.044
COL2A1-751B, n/N (%)	DQA1*01:02	18/43 (42)	12/57 (21)	0.029
COL2A1-885B, n/N (%)	DRB3*02:02	1/54 (2)	6/46 (13)	0.046
COL2A1-976B, n/N (%)	DQA1*05:05	1/51 (2)	7/49 (14)	0.029

* COL2A1, collagen Type II α1.
[a] Fisher's exact test.

Table 2. Association of HLA alleles with IgG that binds specific IgG1 peptides

Peptide	HLA allele	Allele frequency in high binders	Allele frequency in low binders	*P* value[a]
IgG1-167B, n/N (%)	DRB1*04:01	19/30 (63)	25/70 (36)	0.015
	DPB1*03:01	1/30 (3)	14/70 (20)	0.035
IgG1-219B, n/N (%)	DPB1*04:02	5/35 (14)	23/65 (35)	0.035
	DQA1*01:05[b]	4/35 (11)	0/65 (0)	0.013
	DRB1*10:01[b]	4/35 (11)	0/65 (0)	0.013

[a] Fisher's exact test.
[b] The same four subjects possessed both of these alleles.

the collagen peptides, DQA1*01:02 was present in 43% of the 40 participants with RA with multireactivity across four to five peptides versus 22% of the 60 subjects with limited or no reactivity (0–3 peptides; $P = 0.044$). Moreover, all multireactive sera bound COL2A1-885B, and DQA1*01:02 was present in 43% of those 40 participants compared to only 14% of the 14 participants with anti–COL2A1-885B and limited reactivity (1–3 peptides, $P = 0.058$). For IgG1 peptides, only 13 participants had multireactivity across four to five peptides, which limited analysis. However, 34 participants had high IgG binding to three to five IgG1 peptides, and DRB1*04:01 was more frequent in these participants than in participants with binding to zero to two peptides (59% versus 36%; $P = 0.036$). These HLA associations with multireactivity are consistent with the peptide-specific associations depicted in Figure 4. Together, these data suggest that multireactivity to citrulline-glycine–containing collagen peptides is associated with HLA-DQA1*01:02, whereas reactivity with some IgG1 peptides, with or without citrulline-serine, is associated with HLA-DRB1*04:01.

One factor to consider regarding these results is the high level of linkage disequilibrium among HLA alleles. Therefore, for HLA alleles commonly in linkage disequilibrium[43] with DRB1*04:01 or DQA1*01:02 that were present in more than one participant in our cohort, we included the *P* values (all >0.05) for the above analyses (Supplementary Table 3). We also evaluated the association for each partial haplotype (ie, the combination of the allele of interest and the allele in linkage disequilibrium) for each antibody and multireactivity (Supplementary Table 4). The only statistically significant results were reduced presence of DRB1*04:01/DQB1*03:02 in participants with versus without IgG anti–COL2A1-693B (20% versus 40%; $P = 0.045$), IgG anti–COL2A1-976B (20% versus 39%; $P = 0.047$), and anticollagen multireactivity (18% versus 37%; $P = 0.045$). These findings are consistent with the negative association of DRB1*04:01 with IgG anti–COL2A1-693B (Table 1) and the trend toward a negative association of DRB1*04:01 with anticollagen multireactivity (Supplementary Table 3; DRB1*04:01 present in 33% of people with versus 52% of people without anticollagen multireactivity; $P = 0.067$).

Finally, intrigued by the opposite relationship of DRB1*04:01 with antibodies that recognize IgG1 versus collagen peptides, we divided participants into four groups: multireactivity for collagen

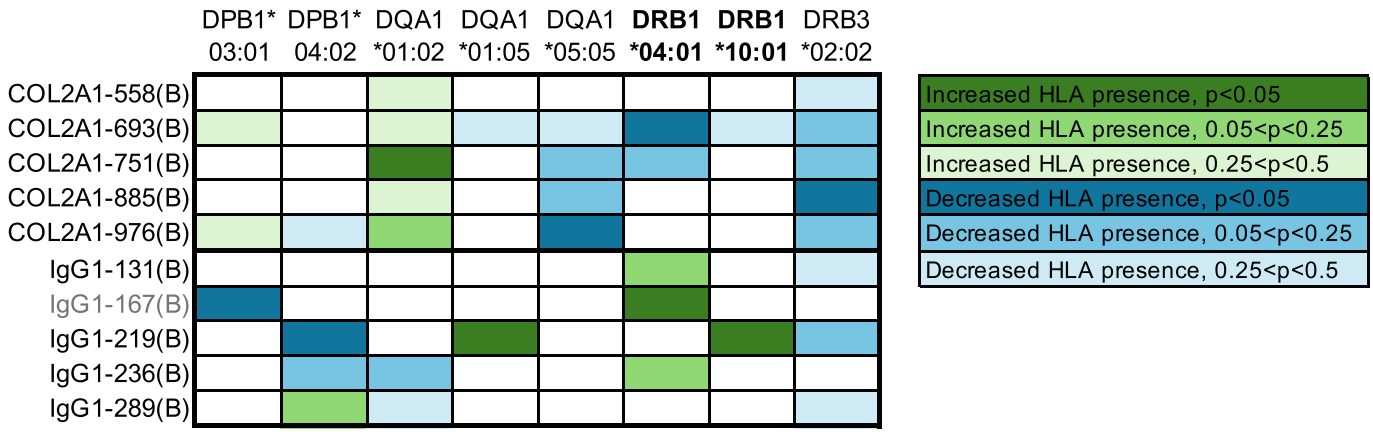


Figure 4. Heatmap of HLA presence in participants with rheumatoid arthritis with IgG binding to citrulline-containing collagen and IgG1 peptides. For each HLA allele significantly associated with an antipeptide antibody in participants with rheumatoid arthritis, we depicted in a heatmap the association of that HLA allele with all 10 antibodies. Cells are colored different shades of green (HLA allele present more frequently in subjects with the antibody versus without) or blue (HLA allele present less frequently in subjects with the antibody versus without) based on the *P* value in a Fisher's exact test. White cells represent $P \geq 0.5$. HLA alleles in bold contain the shared epitope, and peptide in gray has neither citrulline-serine or citrulline-glycine. COL2A1, collagen Type II α1.

and IgG1 peptides, multireactivity for collagen only, multireactivity for IgG1 only, and no multireactivity (Supplementary Figure 3A). We then compared the presence of each allele in a pairwise manner using a Fisher's exact test. The only significant associations were found for DRB1*04:01, potentially due to the smaller sample sizes that result from dividing participants into four groups, with the strongest trend seen for DQA1*01:02. As shown in Supplementary Figure 3B, DRB1*04:01 was present in 74% of participants with multireactivity for IgG1 and not collagen, the two conditions positively associated with DRB1*04:01, compared with 28% of participants with multireactivity for collagen and not IgG1, the two conditions negatively associated with DRB1*04:01 ($P = 0.006$). The difference was less dramatic when comparing the presence of DRB1*04:01 in participants with multireactivity for IgG1 and not collagen versus participants with only one condition positively associated with DRB1*04:01 (ie, no multireactivity; 74% versus 41%; $P = 0.027$) or multireactivity with both peptides sets (74% versus 40%; $P = 0.080$). In contrast, DQA1*01:02 strongly trended toward greater presence in participants with multireactivity to collagen only as compared with subjects without collagen multireactivity either with (48% versus 16%; $P = 0.052$) or without (48% versus 24%; $P = 0.062$) IgG1 multireactivity. Together, these data re-enforce the positive association of DRB1*04:01 with both the presence of IgG1 multireactivity and the absence of collagen multireactivity as compared with the positive association of DQA1*01:02 with collagen multireactivity independent from IgG1 multireactivity.

DISCUSSION

This study aimed to determine if specific HLA alleles associate with specific anticitrullinated collagen antibodies, RA-associated RFs, and/or multireactivity in RA. We discovered a novel positive association between DQA1*01:02 and multireactivity with citrulline-glycine–containing Type II collagen-derived peptides. Further, we found a positive association of shared epitope–containing DRB1*04:01 with antibodies that bind citrulline-containing IgG epitopes in a more limited manner and a negative association of DRB1*04:01 with antibodies that bind citrulline-glycine–containing collagen peptides. We also identified several new HLA associations with specific antibodies.

To assess HLA associations, we first evaluated reactivity patterns of RA sera with collagen- and IgG1-derived peptides. As expected, sera had a variety of IgG-binding patterns across participants, even when citrulline-glycine or citrulline-serine motifs were present. However, serum multireactivity was more prominent for collagen than IgG1 peptides. Similarly, purified anti–COL2A1-885B antibodies were promiscuous, and anti–IgG1-289B antibodies were more private. The multireactivity against collagen peptides is likely due to the presence of citrulline-glycine, not the collagen protein, because citrulline-glycine is a motif for ACPA antigens based on a wide variety of peptides derived from many different proteins.[2,23,44] Type II collagen has many citrulline-glycine pairs, which may underlie its frequent targeting in RA. In contrast, citrulline-serine was only identified in a single study, and it was a motif for RA serum IgG binding, not monoclonal ACPA binding.[23] Thus, citrulline-serine could be a motif for more private antibodies that, perhaps, have unique reactivities based on binding to amino acids that neighbor the motif. Additional studies are needed to evaluate citrulline-serine, citrulline-glycine, IgG1, and collagen more independently to draw definitive conclusions.

The differential multireactivity for antibodies that bind citrulline-containing collagen and IgG epitopes is intriguing, especially in the context of RA development. Because antibodies that bind citrullinated collagen have a high likelihood of being promiscuous, the original tolerance-breaking antigen is unknown. Because of molecular mimicry with Epstein-Barr virus proteins, citrullinated Type II collagen could be an original antigen driving RA.[45] Alternatively, tolerance could be lost at a mucosal surface,[46] with promiscuous ACPAs developing during the preclinical ACPA expansion period,[47] perhaps due to epitope spreading. The promiscuous ACPAs could then bind citrullinated collagen in a nonspecific, although potentially pathogenic, manner. In contrast, RA-associated RFs appear to be more private, although not completely private given some cross-reactivity (Figure 3) and the binding of some promiscuous ACPAs to citrulline-containing and homocitrulline-containing IgG epitopes.[21] Given this level of privacy, the presence of IgG at sites of infection, IgG citrullination,[48] and RF cross-reactivity with pathogen antigens,[20,49] tolerance could be specifically lost for citrullinated IgG early on the path to RA.[50] However, future studies evaluating these antibodies in the preclinical RA period would be needed to more definitely determine antigen chronology.

In addition to, or perhaps due to, the differences in multireactivity, we identified different HLA associations for anticitrullinated collagen antibodies and RA-associated RFs. DQA1*01:02 was the only HLA allele positively associated with citrullinated collagen reactivity in our cohort. It was present in more RA participants with IgG binding to COL2A1-751B (Table 1), the antibody with the most associated serum multireactivity (Supplementary Figure 2), with similar trends for the other collagen peptides (Figure 4). DQA1*01:02 was also present in more RA participants with serum multireactivity to collagen peptides. Finally, among RA participants with anti–COL2A1-885B (whose anticollagen multireactivity mirrored that of serum; Figure 3), those with multireactive sera had a strong trend of more frequently having DQA1*01:02 than those with limited reactivity sera. Together, these data suggest, albeit to some extent indirectly, that DQA1*01:02 is associated with promiscuous IgG that binds to citrulline-glycine–containing collagen peptides. To the best of our knowledge, these findings provide the first evidence for a role for DQA1*01:02 in RA as well as the first positive HLA association with multireactivity. Interestingly, DQA1*01:02 is associated with

increased lupus risk in multiple populations,[51] and lupus autoantibodies can be promiscuous,[52] raising the possibility that DQA1*01:02 could contribute to the development of promiscuous antibodies across diseases.

As opposed to the increased presence of DQA1*01:02 in participants with RA with anticitrullinated collagen antibodies, there was a decreased presence of DRB3*02:02 and DQA1*05:05, most prominently in RA participants with anti–COL2A1-885B and anti–COL2A1-976B, respectively (Table 1), as well as anticollagen antibodies in general (Figure 4). Given these broad associations across citrulline-glycine–containing collagen peptides, these HLAs may be negatively associated with multireactivity. Little is known about these two alleles in the context of RA, leaving many unanswered questions about how they could protect against the formation of promiscuous RA autoantibodies. Also, there was a decreased presence of shared epitope–containing DRB1*04:01 in RA participants with anti–COL2A1-693B with a similar trend for anti–COL2A1-751B. Consistent with these findings, there was a trend toward a negative association of DRB1*04:01 with anticollagen multireactivity and a significant negative association of the DRB1*04:01/DQB1*03:02 partial haplotype with IgG anti–COL2A1-693B, IgG anti–COL2A1-976B, and anticollagen multireactivity. Interestingly, DRB1*04 positively associates with the native form of a peptide almost identical to COL2A1-558 (C^{III}),[38] suggesting different HLA associations for citrullinated versus native epitopes of Type II collagen.

As opposed to the negative association with anticitrullinated collagen antibodies, DRB1*04:01 was present more frequently in RA participants with IgG anti–IgG1-167B (Table 2) with a similar trend for IgG1-131B and IgG1-236B (Figure 4). DRB1*04:01 was also present more frequently in RA participants with serum multireactivity for IgG1 peptides. However, multireactivity was low for RA-associated RFs, and IgG1-167B does not have citrulline-serine. Thus, DQA1*01:02 may be associated with promiscuous citrulline-glycine–reactive anticollagen antibodies, whereas DRB1*04:01 may be associated with a subset of more private RA-associated RFs. Interestingly, another shared epitope–containing HLA, DRB1*10:01, was more frequent in RA participants with IgG anti–IgG1-219B (although, as noted above, it is in complete linkage disequilibrium with DQA1*01:05, limiting the strength of this association). Nonetheless, shared epitope–containing HLA could be positively associated with private ACPAs, including RA-associated RFs, in contrast to a negative association with promiscuous citrulline-glycine–reactive anticollagen antibodies. In support of this conclusion, shared epitope–containing HLAs recently have been shown to associate with noncitrulline-glycine motif ACPAs more than citrulline-glycine motif ACPAs.[53] Also consistent with this finding is the positive association of DRB1*04 with three private antibodies[38] and the positive association of DRB1*04:01 with ACPAs that bind a citrulline-containing fibrin peptide without citrulline-glycine (SGIGTLDGF**B**H**B**HPD).[54] These studies did not evaluate the HLA-DQ locus, but future studies should.

Finally, DPB1*04:02 and DPB1*03:01 were present more commonly in participants with RA with low binding to IgG1-219B and IgG1-167B, respectively. How DPB1 alleles impact RA development is only emerging,[32] and the relationship between DPB1*04:02 and DPB1*03:01 and RA-associated RFs will require more exploration.

The mechanism underlying the associations of HLAs with autoantibodies is not completely clear but may be due to HLA-binding affinities for antigens, as previously explored for shared epitope–containing alleles. The more frequent presence of DRB1*04:01 in participants with RA-associated RFs, two that bind a peptide with citrulline-serine and one that binds a peptide without a motif, is consistent with DRB1*04:01 binding a variety of peptides with different amino acids adjacent to citrulline.[29,36] DRB1*10:01, which was positively associated with anti–IgG1-219B, has similar variability.[55,56] Moreover, the negative association of DRB1*04:01 with anti–COL2A1-693B in our study is consistent with poor binding of DRB1*04:01 to a citrulline-containing collagen peptide with five identical amino acids (**B**GFPG) as COL2A1-693B.[56] These different binding patterns may be explained by the shared epitope–defined P4 pocket contacting citrulline[28] and the nonshared epitope residues of other pockets differentially driving affinity for the rest of the peptide antigen.[29] Thus, differential HLA affinity may be driving the differential HLA associations with specific antibodies. Alternatively, the hapten carrier model, in which PAD4 (carrier) is targeted by T cells that help B cells reacting to citrullinated peptides (haptens),[57] may be contributing to the development of private and/or promiscuous ACPAs.

There are several limitations to our study. Most importantly, our sample size is small, which prevents statistical significance when correcting for multiple comparisons. Also, our participants are all North American and 85% White (Supplementary Table 1). These limitations reduce the statistical power, robustness, and generalizability of our findings. To partially address the issue of statistical power, we provided information about potentially biologically relevant trends, although these findings must be interpreted with caution. Further, although we performed a partial analysis of haplotypes, the small sample size may have led us to underestimate the contributions from alleles in linkage disequilibrium. Also, we were unable to purify antibodies that bind all 10 peptides from all sera to comprehensively evaluate multireactivity, limiting the strength of our conclusions about HLA associations with multireactivity. Additionally, our studies did not include peptides with citrulline-glycine or citrulline-serine that were not collagen derived or IgG1 derived, respectively, or collagen peptides that did not contain citrulline-glycine, preventing us from differentiating between findings due to motifs versus specific proteins. Peptides also represented only linear, and not conformational, epitopes. Finally, we only demonstrated HLA associations with antibodies; thus, mechanisms remain unknown.

Nonetheless, this study identified a novel positive association for DQA1*01:02 and a negative association for DRB1*04:01 with promiscuous antibodies that bind citrulline-glycine–containing Type II collagen-derived peptides in RA as well as a positive association between shared epitope–containing HLA alleles, most prominently DRB1*04:01, and RA-associated RFs, which appear to be more private. Future studies evaluating additional peptides from other important antigens as well as larger, more diverse cohorts are needed to further refine these associations as well as to define the underlying mechanisms.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr Shelef confirms that all authors have provided the final approval of the version to be published, and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. Schellekens GA, de Jong BA, van den Hoogen FH, et al. Citrulline is an essential constituent of antigenic determinants recognized by rheumatoid arthritis-specific autoantibodies. J Clin Invest 1998;101(1):273–281.
2. Steen J, Forsström B, Sahlström P, et al. Recognition of amino acid motifs, rather than specific proteins, by human plasma cell-derived monoclonal antibodies to posttranslationally modified proteins in rheumatoid arthritis. Arthritis Rheumatol 2019;71(2):196–209.
3. Ge C, Holmdahl R. The structure, specificity and function of anti-citrullinated protein antibodies. Nat Rev Rheumatol 2019;15(8):503–508.
4. Masson-Bessière C, Sebbag M, Girbal-Neuhauser E, et al. The major synovial targets of the rheumatoid arthritis-specific antifilaggrin autoantibodies are deiminated forms of the alpha- and beta-chains of fibrin. J Immunol 2001;166(6):4177–4184.
5. Pratesi F, Dioni I, Tommasi C, et al. Antibodies from patients with rheumatoid arthritis target citrullinated histone 4 contained in neutrophils extracellular traps. Ann Rheum Dis 2014;73(7):1414–1422.
6. Harre U, Georgess D, Bang H, et al. Induction of osteoclastogenesis and bone loss by human autoantibodies against citrullinated vimentin. J Clin Invest 2012;122(5):1791–1802.
7. Courtenay JS, Dallman MJ, Dayan AD, et al. Immunisation against heterologous type II collagen induces arthritis in mice. Nature 1980;283(5748):666–668.
8. Stuart JM, Dixon FJ. Serum transfer of collagen-induced arthritis in mice. J Exp Med 1983;158(2):378–392.
9. Morgan K, Clague RB, Collins I, et al. Incidence of antibodies to native and denatured cartilage collagens (types II, IX, and XI) and to type I collagen in rheumatoid arthritis. Ann Rheum Dis 1987;46(12):902–907.
10. Burkhardt H, Koller T, Engström A, et al. Epitope-specific recognition of type II collagen by rheumatoid arthritis antibodies is shared with recognition by antibodies that are arthritogenic in collagen-induced arthritis in the mouse. Arthritis Rheum 2002;46(9):2339–2348.
11. Burkhardt H, Sehnert B, Bockermann R, et al. Humoral immune response to citrullinated collagen type II determinants in early rheumatoid arthritis. Eur J Immunol 2005;35(5):1643–1652.
12. Liang B, Ge C, Lonnblom E, et al. The autoantibody response to cyclic citrullinated collagen type II peptides in rheumatoid arthritis. Rheumatology (Oxford) 2019;58(9):1623–1633.
13. Schulte S, Unger C, Mo JA, et al. Arthritis-related B cell epitopes in collagen II are conformation-dependent and sterically privileged in accessible sites of cartilage collagen fibrils. J Biol Chem 1998;273(3):1551–1561.
14. Li Q, Li Y, Liang B, et al. Rheumatoid arthritis sera antibodies to citrullinated collagen type II bind to joint cartilage. Arthritis Res Ther 2022;24(1):257.
15. Sokolove J, Johnson DS, Lahey LJ, et al. Rheumatoid factor as a potentiator of anti-citrullinated protein antibody-mediated inflammation in rheumatoid arthritis. Arthritis Rheumatol 2014;66(4):813–821.
16. Anquetil F, Clavel C, Offer G, et al. IgM and IgA rheumatoid factors purified from rheumatoid arthritis sera boost the Fc receptor- and complement-dependent effector functions of the disease-specific anti-citrullinated protein autoantibodies. J Immunol 2015;194(8):3664–3674.
17. Artandi SE, Calame KL, Morrison SL, et al. Monoclonal IgM rheumatoid factors bind IgG at a discontinuous epitope comprised of amino acid loops from heavy-chain constant-region domains 2 and 3. Proc Natl Acad Sci USA 1992;89(1):94–98.
18. Terness P, Kohl I, Hübener G, et al. The natural human IgG anti-F(ab')2 antibody recognizes a conformational IgG1 hinge epitope. J Immunol 1995;154(12):6446–6452.
19. Bonagura VR, Artandi SE, Davidson A, et al. Mapping studies reveal unique epitopes on IgG recognized by rheumatoid arthritis-derived monoclonal rheumatoid factors. J Immunol 1993;151(7):3840–3852.
20. Amjadi MF, Parker MH, Adyniec RR, et al. Novel and unique rheumatoid factors cross-react with viral epitopes in COVID-19. J Autoimmun 2024;142:103132.
21. Mergaert AM, Zheng Z, Denny MF, et al. Rheumatoid factor and anti-modified protein antibody reactivities converge on IgG epitopes. Arthritis Rheumatol 2022;74(6):984–991.
22. Titi AH, Krisko BT, Bashar SJ, et al. Rheumatoid arthritis-associated rheumatoid factors post-COVID-19. Front Immunol 2025;16:1553540.
23. Zheng Z, Mergaert AM, Fahmy LM, et al. Disordered antigens and epitope overlap between anti-citrullinated protein antibodies and rheumatoid factor in rheumatoid arthritis. Arthritis Rheumatol 2020;72(2):262–272.
24. Burastero SE, Casali P, Wilder RL, et al. Monoreactive high affinity and polyreactive low affinity rheumatoid factors are produced by CD5+ B cells from patients with rheumatoid arthritis. J Exp Med 1988;168(6):1979–1992.
25. Gregersen PK, Silver J, Winchester RJ. The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. Arthritis Rheum 1987;30(11):1205–1213.
26. van der Helm-van Mil AH, Verpoort KN, Breedveld FC, et al. The HLA-DRB1 shared epitope alleles are primarily a risk factor for anti-cyclic citrullinated peptide antibodies and are not an independent risk factor for development of rheumatoid arthritis. Arthritis Rheum 2006;54(4):1117–1121.
27. Hill JA, Southwood S, Sette A, et al. Cutting edge: the conversion of arginine to citrulline allows for a high-affinity peptide interaction with the rheumatoid arthritis-associated HLA-DRB1*0401 MHC class II molecule. J Immunol 2003;171(2):538–541.

28. Scally SW, Petersen J, Law SC, et al. A molecular basis for the association of the HLA-DRB1 locus, citrullination, and rheumatoid arthritis. J Exp Med 2013;210(12):2569–2582.
29. Ting YT, Petersen J, Ramarathinam SH, et al. The interplay between citrullination and HLA-DRB1 polymorphism in shaping peptide binding hierarchies in rheumatoid arthritis. J Biol Chem 2018;293(9): 3236–3251.
30. Sidney J, Becart S, Zhou M, et al. Citrullination only infrequently impacts peptide binding to HLA class II MHC. PLoS One 2017;12(5): e0177140.
31. Auger I, Sebbag M, Vincent C, et al. Influence of HLA-DR genes on the production of rheumatoid arthritis-specific autoantibodies to citrullinated fibrinogen. Arthritis Rheum 2005;52(11):3424–3432.
32. Raychaudhuri S, Sandor C, Stahl EA, et al. Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis. Nat Genet 2012;44(3):291–296.
33. Guo J, Zhang T, Cao H, et al. Sequencing of the MHC region defines HLA-DQA1 as the major genetic risk for seropositive rheumatoid arthritis in Han Chinese population. Ann Rheum Dis 2019;78(6): 773–780.
34. Lagutkin D, Panaifo L, Nurul-Aain AF, et al. Revisiting the link between HLA-DRB1 alleles and autoantibodies in rheumatoid arthritis: association of non-shared epitope alleles *09 and *15 with high levels of anti-citrullinated peptide/protein antibodies. ACR Open Rheumatol 2025; 7(1):e11767.
35. Kampstra AS, van Heemst J, Moustakas AK, et al. The increased ability to present citrullinated peptides is not unique to HLA-SE molecules: arginine-to-citrulline conversion also enhances peptide affinity for HLA-DQ molecules. Arthritis Res Ther 2016;18(1):254.
36. Anderson KM, Roark CL, Portas M, et al. A molecular analysis of the shared epitope hypothesis: binding of arthritogenic peptides to DRB1*04 alleles. Arthritis Rheumatol 2016;68(7):1627–1636.
37. Manivel VA, Mullazehi M, Padyukov L, et al. Anticollagen type II antibodies are associated with an acute onset rheumatoid arthritis phenotype and prognosticate lower degree of inflammation during 5 years follow-up. Ann Rheum Dis 2017;76(9):1529–1536.
38. Snir O, Widhe M, von Spee C, et al. Multiple antibody reactivities to citrullinated antigens in sera from patients with rheumatoid arthritis: association with HLA-DRB1 alleles. Ann Rheum Dis 2009;68(5): 736–743.
39. Rebernick R, Fahmy L, Glover C, et al. DNA area and NETosis analysis (DANA): a high-throughput method to quantify neutrophil extracellular traps in fluorescent microscope images. Biol Proced Online 2018; 20(1):7.
40. Mergaert AM, Bawadekar M, Nguyen TQ, et al. Reduced anti-histone antibodies and increased risk of rheumatoid arthritis associated with a single nucleotide polymorphism in PADI4 in North Americans. Int J Mol Sci 2019;20(12):3093.
41. Bashar SJ, Zheng Z, Mergaert AM, et al. Limited biomarker potential for IgG autoantibodies reactive to linear epitopes in systemic lupus erythematosus or spondyloarthropathy. Antibodies (Basel) 2024; 13(4):87.
42. Bajtner E, Nandakumar KS, Engström A, et al. Chronic development of collagen-induced arthritis is associated with arthritogenic antibodies against specific epitopes on type II collagen. Arthritis Res Ther 2005;7(5):R1148–R1157.
43. Creary LE, Sacchi N, Mazzocco M, et al. High-resolution HLA allele and haplotype frequencies in several unrelated populations determined by next generation sequencing: 17th International HLA and Immunogenetics Workshop joint report. Hum Immunol 2021;82(7): 505–522.
44. Dam CE, Houen G, Trier NH. The dependency on neighboring amino acids for reactivity of anti-citrullinated protein antibodies to citrullinated proteins. Scand J Clin Lab Invest 2016;76(5):417–425.
45. Robinson WH, Younis S, Love ZZ, et al. Epstein-Barr virus as a potentiator of autoimmune diseases. Nat Rev Rheumatol 2024;20(11): 729–740.
46. Holers VM, Demoruelle MK, Kuhn KA, et al. Rheumatoid arthritis and the mucosal origins hypothesis: protection turns to destruction. Nat Rev Rheumatol 2018;14(9):542–557.
47. Sokolove J, Bromberg R, Deane KD, et al. Autoantibody epitope spreading in the pre-clinical phase predicts progression to rheumatoid arthritis. PLoS One 2012;7(5):e35296.
48. Tutturen AE, Fleckenstein B, de Souza GA. Assessing the citrullinome in rheumatoid arthritis synovial fluid with and without enrichment of citrullinated peptides. J Proteome Res 2014;13(6):2867–2873.
49. Nakamura M, Burastero SE, Notkins AL, et al. Human monoclonal rheumatoid factor-like antibodies from CD5 (Leu-1)+ B cells are polyreactive. J Immunol 1988;140(12):4180–4186.
50. Shelef MA. New relationships for old autoantibodies in rheumatoid arthritis. Arthritis Rheumatol 2019;71(9):1396–1399.
51. Selvaraja M, Too CL, Tan LK, et al. Human leucocyte antigens profiling in Malay female patients with systemic lupus erythematosus: are we the same or different? Lupus Sci Med 2022;9(1):e000554.
52. DeGiorgio LA, Konstantinov KN, Lee SC, et al. A subset of lupus anti-DNA antibodies cross-reacts with the NR2 glutamate receptor in systemic lupus erythematosus. Nat Med 2001;7(11):1189–1193.
53. Mathsson Alm L, Westerlind H, Gehring I, et al; Swedish Rheumatology Quality Register Biobank Study Group (SRQb); The Danish Rheumatologic Biobank Study Group. Recognition of glycine versus nonglycine citrulline motifs dictating the HLA class II association of anticitrullinated protein antibodies: insights from autoantibody profiling of 6,900 Scandinavian patients with rheumatoid arthritis. Arthritis Rheumatol 2025;77(9):1179–1193.
54. Larid G, Pancarte M, Offer G, et al. In rheumatoid arthritis patients, HLA-DRB1*04:01 and rheumatoid nodules are associated with ACPA to a particular fibrin epitope. Front Immunol 2021;12:692041.
55. James EA, Moustakas AK, Bui J, et al. HLA-DR1001 presents "altered-self" peptides derived from joint-associated proteins by accepting citrulline in three of its binding pockets. Arthritis Rheum 2010;62(10):2909–2918.
56. Chemin K, Pollastro S, James E, et al. A novel HLA-DRB1*10: 01-restricted T cell epitope from citrullinated type II collagen relevant to rheumatoid arthritis. Arthritis Rheumatol 2016;68(5):1124–1135.
57. Roudier J, Balandraud N, Auger I. How RA associated HLA-DR molecules contribute to the development of antibodies to citrullinated proteins: the hapten carrier model. Front Immunol 2022;13:930112.

Arthritis & Rheumatology
Vol. 78, No. 3, March 2026, pp 613–629
DOI 10.1002/art.43430

American College of Rheumatology
Empowering Rheumatology Professionals

Gut Microbiota–Derived Tryptophan Indole Metabolites Ameliorate Collagen-Induced Arthritis in Mice Via Aryl Hydrocarbon Receptor Activation

Liang Wu,[1] Qiuping Zhang,[1] Zhigang Tang,[1] Yanhong Li,[1] Tong Wu,[1] Lvyi Chen,[2] Chunyu Tan,[1] Lingshu Zhang,[1] Xing Ji,[1] Sijun Zhang,[1,3] Yinlan Wu,[1] Aline Bozec,[4] Mario M. Zaiss,[4] and Yubin Luo[1,3]

Objective. Our study's objective is to investigate the specific role of tryptophan metabolism, especially that of microbiome-derived metabolites, in the development of rheumatoid arthritis (RA).

Methods. We employed metabolomics to profile metabolites in 53 individuals at high risk for RA (PreRA), 30 patients with established RA, and 38 healthy individuals. Fecal microbiota transplantation (FMT) and collagen-induced arthritis (CIA) mouse models were used to investigate the impact of gut microbiome on arthritis severity, gut barrier function, and metabolic change. Treg cell differentiation and epithelial cells' barrier function were assessed by flow cytometry, immunofluorescence staining, and Western blotting. Co-immunoprecipitation and luciferase were applied for molecular mechanism studies.

Results. Dysregulated tryptophan metabolism exists in individuals with RA and PreRA, as well as in FMT mice, characterized by a shift toward the kynurenine pathway and reduced activity of serotonin and indole pathways. Indole-3-lactic acid (ILA) and indole-3-acetic acid (IAA) significantly alleviated arthritis in CIA mice by expanding Treg cells via the classical aryl hydrocarbon receptor (AhR)–aryl hydrocarbon receptor nuclear translocator–xenobiotic response element signaling pathway. Moreover, ILA repaired the leaking gut by increasing Zo-1 and occludin expression in Caco-2 cells, which was blocked by AhR antagonist CH223191. Moreover, CH223191 treatment could significantly reverse the improving effects of ILA and IAA on arthritis in mice.

Conclusion. These findings indicate that tryptophan indole metabolites may play a negative regulatory role in the progression of RA by affecting Treg cell development and intestinal gut barrier function.

INTRODUCTION

Numerous studies have demonstrated a critical role of gut microbiota dysbiosis in the pathogenesis of both preclinical and established rheumatoid arthritis (RA).[1–3] One of the key processes begins at the intestinal epithelial barrier, which is characterized by reduced levels of tight junction proteins such as zonulin-1 (ZO-1).[1,2] This disruption leads to the breakdown of immunologic self-tolerance in the mucosal immune system, accompanied by expansion of T helper 17 (Th17) cells. These Th17 cells or bacteria may disseminate to distal organs, such as synovial joints, thereby eliciting inflammatory responses.[2] Another consequence of gut microbiota dysbiosis is the aberrant metabolic profile observed in individuals at the preclinical and established RA stages.[1,4–6] Metabolites, especially those regulated by the gut microbiota, may also participate in the etiology of RA initiation or development. However, whether other types of metabolites in addition to short-chain fatty acids regulate immune function in RA has not been fully elucidated.

Supported by the National Natural Science Foundation of China (82271828,82302049), the Key Research and Development Project of Guangxi Province (AB25069068), the Sichuan University From 0 to 1 Innovation Project (2022SCUH0020), the Scientific Application and Foundation Project of the Science and Technology Department of Sichuan Province (2023NSFSC1702), and the Sichuan Province International Cooperative Foundation (2023YFH0081).

[1]Department of Rheumatology and Immunology, Laboratory of Rheumatology and Immunology, West China Hospital, Sichuan University, Chengdu, Sichuan, China; [2]School of Pharmaceutical Sciences, South-Central Minzu University, Wuhan, China; [3]Frontiers Science Center for Disease-related Molecular Network, West China Hospital, Sichuan University, Chengdu, Sichuan, China; [4]Department of Internal Medicine 3-Rheumatology and Immunology, Friedrich-Alexander University of Erlangen-Nürnberg and Universitätsklinikum Erlangen, Erlangen, Germany.

Authors L. Wu and Q. Zhang contributed equally to this work.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/art.43430).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/art.43430.

Address correspondence via email to Yubin Luo, PhD, at luoyubin2016@163.com.

Submitted for publication July 11, 2024; accepted in revised form August 27, 2025.

Tryptophan (Trp) is an indispensable amino acid that humans cannot synthesize and is supplied by dietary protein. Trp can be metabolized through three major pathways: the kynurenine (Kyn), serotonin, and indole pathways. Multiple studies have shown a skewed distribution of Trp metabolites in autoimmune diseases such as RA, systemic lupus erythematosus, and spondyloarthritis, characterized by an elevated Kyn/Trp ratio in the serum and urine.[7–12] Gut microbiota dysbiosis induces lupus-like autoimmunity in mice through an increase in Kyn. Low dietary tryptophan prevents autoimmune pathology, whereas high dietary tryptophan exacerbates the disease.[13] Seymour et al demonstrated that increases in Trp-derived indoles are required for the development of collagen-induced arthritis (CIA).[14] However, Moulin et al reported that an overactivated Kyn pathway but a relative deficiency of kynurenic acid (KYNA) and xanthurenic acid (XA) metabolites affects RA development. Moreover, the administration of the recombinant aminoadipate aminotransferase, a recombinant form of Kyn aminotransferase 2, has been shown to have a protective effect on CIA mice by restoring KYNA and XA production,[15] indicating that counteracting Trp catabolism alterations may be a new therapeutic strategy for RA therapy.

Unlike the serotonin and Kyn pathways, which rely on endogenous host Trp metabolism, enzymes in the intestinal microbiota catabolize Trp into various bioactive metabolites. These metabolites serve as key signaling molecules in host–microbial cross-talk and contribute to intestinal and systemic homeostasis.[16,17] Numerous gram-negative bacteria, such as *Escherichia coli*, as well as gram-positive species, including *Clostridium spp* and *Bacteroides spp*, synthesize indole from Trp via the tryptophanase enzyme.[16,18] Other bacteria, such as *Clostridium sporogenes*, convert tryptophan into indole-3-lactic acid (ILA), indole 3-propionic acid (IPA), and indole-3-aldehyde via various other metabolic pathways.[19–21] Other groups of bacteria capable of converting tryptophan include Lactobacillus, Bacteroides, Clostridium, Lactobacillus reuteri, and Lactobacillus spp. has been reported to convert tryptophan to ILA.[22,23] Several *Bacteroides* species and *Clostridium bartlettii* can also catabolize tryptophan into ILA and indole-3-acetic acid (IAA),[24] whereas *Bifidobacterium spp* can produce ILA.[24,25] To date, whether changes in the levels of indole metabolites associated with gut microbiota dysbiosis contribute to arthritic diseases in humans remains unknown, especially in the early stage of the disease.

The aryl hydrocarbon receptor (AhR) is a transcription factor activated by environmental pollutants in addition to a numerous Trp metabolites derived from either the microbiota or the endogenous Kyn pathway.[26] Importantly, AhR activation by some of these ligands has been linked to the differentiation and function of both innate and adaptive immune cells,[27] highlighting the importance of the Trp-AhR axis in modulating immunity. Many indole derivatives are known AhR ligands and thus function as immune modulatory compounds.[26] For example, *Lactobacillus reuteri*–derived ILA activates AhR to generate immunoregulatory $CD4^+CD8^+$ intraepithelial lymphocytes.[22] Tryptamine and indole-3-acetate reduce fatty acid– and lipopolysaccharide (LPS)-stimulated production of proinflammatory cytokines in macrophages and suppress inflammatory responses in hepatocytes.[28]

The present study tested the hypothesis that gut microbiota dysbiosis in RA disrupts the balance of immunomodulatory microbiota metabolites derived from the microbiota, thereby increasing susceptibility to arthritis in mice. We used a metabolomics approach to identify the commonly altered metabolite types between patients with PreRA (individuals at high risk for RA) and RA as well as gut microbiota–derived tryptophan indole catabolites. Moreover, the effects of altered indole metabolites on immune reactions and gut barrier function, as well as the potential mechanism, were assessed both in vitro and in vivo.

METHODS

Study cohort, patient characteristics, and sample collection. The study procedure was approved by the Biomedical Research Ethics Committee, West China Hospital of Sichuan University (no. 2021[790]), and informed written consent was obtained from all participants according to the Declaration of Helsinki. Patients with PreRA were consecutively recruited from West China Hospital, Sichuan University, China, between April 2017 and May 2021. These participants had positive serum anti-citrulline protein antibody (ACPA), with or without arthralgia at the time of enrollment. The absence of arthritis was confirmed by physical examination of 44 joints. Patients with RA were diagnosed according to the 2010 American College of Rheumatology classification criteria for RA.[29] Most patients with RA were under treatment (Supplemental Table 1). All patients were naive to biologic agents. Age-, sex-, and ethnicity-matched healthy controls (HCs) with no personal history of inflammatory arthritis were recruited. ACPA-negative profiles for the HCs were obtained from the health management center. Participants from all three groups were aged ≥18 years. The exclusion criteria were as follows: having a history of antibiotic treatment or surgery in the previous three months; current extreme diet; major organ dysfunction; cancer; other rheumatic or autoimmune diseases, including osteoarthritis, systemic lupus erythematosus, Sjögren disease or diabetes; and no cardiovascular diseases. A total of 121 participants were enrolled, comprising 53 individuals with PreRA, 30 patients with RA, and 38 HCs. Detailed demographic and clinical parameters are reported in Supplemental Table 1. Serum samples were freshly collected from each participant within one hour after excretion and were immediately frozen and stored at –80°C for further analyses.

16sRNA sequencing, untargeted and targeted metabolomics performance. Detailed methods are provided in the Supplementary Materials. 16S rRNA sequencing was

performed by Majorbio Biopharm Technology Co., Ltd. The liquid chromatography mass spectrometry/mass spectrometry assays were conducted by Novogene Co., Ltd. Trp-targeted metabolomics and lipid metabolomics analyses were performed by MetWare.

Mice and treatment. Six- to seven-week-old male DBA/1 mice were purchased from GemPharmatech laboratory animal company. All mice were raised in a specific pathogen-free animal facility at the Department of Laboratory Animal Center of West China Hospital, Sichuan University, and housed under a 12-hour light-dark cycle. All animal experiments were approved by the Sichuan University Animal Care and Use Committee and followed the recommendations from the Guide for the Care and Use of Laboratory Animals of Chinese Association for Laboratory Animal Science (No. 20220222024). Mice were gavaged with ILA or IAA (5 mM, 20 mM) (MedChemExpress) dissolved in sterile phosphate buffered saline (PBS) (pH 7.4). Each group had six mice. Mice were orally gavaged with ILA or IAA (100 μL/mouse) every day starting on day 21 before the second immunization. Control mice received pH and sodium-matched PBS.

Induction of CIA. Mice were intradermally immunized at the base of the tail with 100 μL of immunization-grade chicken type II collagen (Sigma, catalog number [cat]: C9301) emulsified in complete Freund's adjuvant containing 5 mg/mL killed *Mycobacterium tuberculosis* (Biological Product Institute) on day 0. Twenty-one days after the first immunization, the mice received a booster in the base of the tail and were monitored for clinical signs of arthritis.

Assessment of arthritis and histologic analysis. Arthritis severity was scored blindly from day 21 to day 42. The clinical scores for each paw were evaluated every other day and scored individually on a scale of 0 to 4, resulting in a maximum score of 16 in each mouse. For histology, whole paw joints were isolated and fixed in 4% paraformaldehyde, decalcified in EDTA, and then embedded in paraffin. Tissue sections (4 μm) were stained using hematoxylin and eosin.

Fecal microbiota transplantation. Fecal microbiota transplantation (FMT) was performed as previously described.[1] In brief, six samples from each population group (PreRA, RA, HCs) were selected at random and mixed at equal weight. One gram of the mixed stool was suspended in 5 mL of sterile PBS and centrifuged at 3,000 revolutions per minute for five minutes to obtain the supernatant for FMT. Each mouse used for FMT was orally treated with an aliquot of 200 μL suspension. For FMT experiments, adult male mice were cohoused for two weeks and pretreated with sterile drinking water containing 1 mg/mL ampicillin, neomycin, metronidazole, and 0.5 mg/mL vancomycin for 14 days. Mice were divided into several groups and were fed every other day for two weeks with feces supernatant of individuals from the HC, PreRA, and RA groups. Each group had six mice. Before CIA induction, seven-week-old male DBA/1 mice were cohoused and treated with antibiotics as described and then inoculated with feces from HC, PreRA or RA samples every other day for 14 days. All mice were then immunized with chicken type II collagen as mentioned above.

Gut permeability test. Intestinal permeability was assessed by measuring fluorescein isothiocyanate (FITC)-dextran fluorescence in serum. Mice were deprived of food and water for four hours and then administered nondigestible FITC-conjugated dextran, 4 kDa (Sigma) at a dose of 0.6 mg/g body weight. Fifty microliters of serum were collected from the tail vein and diluted with PBS (1:1). The fluorescence intensity of FITC in serum was measured using a synergy plate reader (Biotek).

Measurement of cytokines in sera. A total of 200 μL of blood was collected by retro-orbital puncture. Sera were then collected and immediately frozen and stored at –80°C for cytokines measurement. IL-6, tumor necrosis factor (TNF)-α, and IL-1b levels in sera were measured using Mouse Th1/Th2/Th17 Cytometric Bead Array kit according to the manufacturer's instructions (BD Biosciences, cat: C560485).

Flow cytometry analysis. Flow cytometry was performed on immune cells from various organs, including the small intestinal lamina propria, Peyer's patches (PPs), mesenteric lymph nodes (mLNs), and spleen. For the small intestine, mesenteric fat was removed, and PPs were collected. The remaining tissues were dissected and incubated in Hank's balanced salt solution with 5 mM EDTA to remove epithelial cells and subsequently digested with 2 mg/mL collagenase type IV and 2.5 U/mL nuclease. Lamina propria lymphocytes were harvested at the interphase between a 40% and an 80% Percoll gradient. Cells were then resuspended in staining solution (PBS with 2% bovine serum albumin) for flow cytometry. For spleen tissues, red blood cells were removed with lysis buffer. The single-cell suspension was filtered through a 70-μm cell strainer, washed with PBS, and resuspended in staining solution for flow cytometry analysis. The fluorochrome-conjugated monoclonal antibodies used include Fixable Viability Stain 700, APC-Cy7 anti-mouse CD3e, PercpCy5.5 anti-mouse CD4, FITC anti-mouse IFN-γ, APC anti-mouse IL-4, PE anti-mouse IL-17A, PE anti-mouse CD25, APC antimouse Foxp3, and BV510 anti-mouse CD19. All of the antibodies and relative isotype controls were purchased from BD Biosciences. The Foxp3/Transcription Factor Staining Buffer Set (eBioscience, CN) was used for Foxp3 staining. Samples were analyzed by BD FACSCelesta, and subsequent analysis was performed with FlowJo software.

Cell culture and treatment. The human colon cancer cell lines Caco-2 was obtained from the ATCC and cultured with complete Dulbecco's modified Eagle's medium (DMEM) medium (10% fetal bovine serum [FBS], penicillin 100U/mL, streptomycin 100 μg/mL). To detect the direct role of ILA or IAA on intestinal permeability, Caco-2 cells were treated with ILA or IAA (0.5, 1, 2, 5 mM) for 48 hours. Forty-eight hours later, cells were collected for real-time polymerase chain reaction (RT-PCR) and Western blotting detection. To test the protective effect of ILA on RA gut microbiome–induced intestinal permeability reduction, Caco-2 cells were first incubated in DMEM (10% FBS, without penicillin and streptomycin) with supernatant isolated from the stool of individuals with RA and HC for three hours. Then, the medium was removed, and the cells were washed three times with sterile PBS. The cells were then cultured with fresh DMEM with different doses of ILA at 37°C and 5% CO_2 for another 48 hours. Finally, the cells were collected for RNA isolation or fixed with formaldehyde for fluorescent staining.

Immunofluorescence staining. Caco-2 cells were fixed with 4% paraformaldehyde for 10 minutes and blocked with 5% goat serum. The cells were then incubated with rabbit anti-human/mouse ZO-1 or occludin antibody at 1:200 at 4°C overnight. Anti-rabbit Alexa Fluor 647 (1:500, Cell Signaling Technology [CST], cat: 8940S) was used as the secondary antibody. All the mages were collected under the fluorescence microscope and analyzed by the pathologic analysis software (QuPath).

Western blot assay. Proteins were obtained from cells or intestine tissue by lysis buffer (Beyotime Biotechnology). After obtaining the whole-cell extracts, a BCA Protein Assay Kit (Beyotime Biotechnology) was applied to determine the protein concentrations. The total protein at equal amount were separated by SDS-PAGE. They were transferred onto membranes and blocked with 5% (weight/volume) nonfat milk for our hour at room temperature. Subsequently, membranes were incubated with specific primary antibodies at 4°C for overnight. The antibodies used in the experiments were as follows: rabbit anti-human/mouse AhR antibody (1:1,000, CST, cat: 83200S); rabbit anti-human/mouse ZO-1 antibody (1:1,000, Affinity Biosciences, cat: AF5145); mouse anti-human/mouse GAPDH antibody (1:10,000, Immunoway); rabbit anti-human/mouse occludin antibody (1:1,000, Affinity Biosciences, cat: DF7504); rabbit anti-human Claudin 1 antibody (1:1,000, Affinity Biosciences, cat: AF0127); rabbit anti-mouse ARNT antibody (1:1,000, MedChemExpress, cat: HY-P80705); and rabbit anti-mouse CYP1A1 antibody (1:1,000, Santa Cruz Biotechnology). They were incubated with anti-rabbit/mouse secondary antibody; then the bands were visualized by using ECL Plus reagent.

Quantitative RT-PCR. Total RNA was isolated from the small intestine or cells using the TRIzol (Invitrogen). RNA was digested with DNaseI and reverse transcribed into complementary DNA using an oligo d(T) primer. Quantitative RT-PCR was performed using SYBR Green Master Mix (Applied Biological Materials). Samples were analyzed in triplicate, and the relative RNA expression was normalized to that of beta-actin. Primer sequences were shown in Supplemental Table 2.

Immunochemical staining. Paraffin-embedded small intestine sections were stained with antibodies against ZO-1 or occludin. The paraffin sections of small intestine were dewaxed and cleaned three times by PBST (0.1% tween-20 in PBS). We incubated these sections with EDTA buffer (pH = 8.0) at 95°C for 25 minutes. Then, sections were sinked in 3% H_2O_2 solution for 20 minutes and blocked with 5% goat serum for 30 minutes, followed with incubation with rabbit anti-human/mouse ZO-1 antibody (1:200, Affinity Biosciences, cat: AF5145) or rabbit anti-human/mouse occludin antibody (1:200, Affinity Biosciences, cat: DF7504) at 4°C overnight. The paraffin sections were stained by DAB reagent. All the images were collected under the microscope (Leica) and analyzed by the pathologic analysis software (QuPath).

Treg differentiation in vitro. Twenty-four–well plates were precoated with anti-mouse CD3 antibody and incubated overnight at 4°C. Plates were washed with PBS before usage. Mouse spleens were collected and ground. The cells suspension was passed through a 70-um strainer and centrifuged at 1,500 rpm × 5 min. After splitting red blood cells, the cells were resuspended in PBS and incubated with Fc receptor blocker. The naive $CD4^+$ T cells were then isolated by EasySep Mouse Naive CD4+ T Cell Isolation Kit following the manufacturer's instruction (STEMCELL Technologies, cat: 19765;). Naive $CD4^+$ T cells were suspended in 1640 complete medium and incubated with anti-mouse CD28 antibody (0.5 μg/mL). Mouse Treg Differentiation Supplement was added following the manufacturer's instruction (ImmunoCult Mouse Treg Differentiation Supplement; STEMCELL Technologies, cat:10957). The proportion of Treg cells was measured by flow cytometry after three days, and cells were collected for RT-PCR analysis.

Co-immunoprecipitation assay. Cells were isolated from mice spleen and incubated in the incubator for six hours to remove adherent cells. Cells were then divided into three groups and treated with 50 μM ILA, 50 μM IAA, or PBS as control for two hours. Subsequently, they were lysed with RIPA lysis buffer for 30 minutes and centrifuged at 12,000 g for five minutes. Cell lysates were incubated with anti-Hsp90 (3 μg) antibody (Proteintech, cat: 1317-1-AP) or IgG overnight at 4°C, followed by incubation with 20 μL Protein G Agarose (CST) for another two hours at 4°C. After washing with PBST, proteins complex was denatured in 1 × SDS loading buffer at 95°C thermal incubation. Proteins were immune-blotted with relative antibodies.

Luciferase assay. 293T cells were seeded into 24-well plates and incubated in the incubator until the confluence reached approximately 70%. Constructed xenobiotic response element (XRE)-luciferase promoter luciferase reporter into pGL3 basic plasmid was transfected into the cells with Lipofectamine 3000 (Thermo Fisher Scientific). The next day, the cells were exposed to 10 μM ILA and IAA. After treatment for 24 hours, a dual-luciferase reporter assay system (Beyotime) was used to detect the firefly and *Renilla* luciferase activities according to the manufacturers' instruction.

AhR translocation detection. Splenic cells were isolated and treated with ILA and IAA for one hour, then fixed with 4% paraformaldehyde for 30 minutes, and permeabilized with 0.2% Triton-100 for 20 minutes. Subsequently, they were blocked with 5% bovine serum albumin for one hour. Cells were then incubated with anti-AhR antibody overnight at 4°C, followed with incubation with Alexa Fluor 488 conjugated anti-rabbit IgG (H^+L), F(ab')2 Fragment (Conjugate) antibody (CST, cat:4412s) for one hour. DAPI staining was used for nucleus detection. The images were obtained under a fluorescence microscope.

Statistical analysis. The experiment data were presented as the mean ± SEM. To determine significantly different bacterial taxa among the three groups, we applied the Kruskal-Wallis test, with multiple tests corrected by the Benjamini-Hochberg false discovery rate test. For cross-sectional analyses of differences in baseline characteristics, diversity indices and other numerical variables among the three groups, one-way analysis of variance (ANOVA) was applied, followed by Bonferroni's correction. For animal studies, statistical significance was determined routinely using one-way ANOVA followed by Bonferroni's post hoc test for multiple comparisons. In vitro experiments were repeated at least three times. The *t*-test was applied to determine statistical significance between two groups. *P* values <0.05 were considered significant. Correlation was determined by Pearson correlation. Statistical tests were performed with GraphPad Prism software, version 8.0 (GraphPad).

RESULTS

Cohort demographics. A total of 121 participants were enrolled in this study, including 38 HCs, 53 patients with PreRA, and 30 patients with RA. Age, sex, body mass index, and smoking status were comparable across all groups (Supplemental Table 1). Among the 53 patients with PreRA, 18 reported a history of arthralgia; however, no arthritis was detected based on physical examination of 44 joints at the time of enrollment. During a five-year follow-up period, 12 of the 53 patients with PreRA developed RA. In the PreRA cohort, the positivity rates for ACPA and IgM rheumatoid factor (IgM-RF) were 100% and 25%, respectively, whereas they were 97% and 87%, respectively, in patients with RA (ACPA, $P = 0.36$; IgM-RF, $P < 0.0001$). Serum antibody titers were significantly higher in patients with RA than in those with PreRA, with ACPA levels being 1.88-fold higher (median: 330.3 U/mL vs 175.3 U/mL) and IgM-RF levels being 5.52-fold higher (median: 413.2 IU/mL vs 74.8 IU/mL). The mean disease duration in patients with RA was 17.7 months (median 13.0 months), and the mean disease activity score (DAS28) was 4.91, indicating the presence of moderate disease in these patients with early RA. Most of the patients with RA were receiving disease-modifying antirheumatic drugs (DMARDs), prednisone, or nonsteroidal anti-inflammatory drug (NSAID) treatment, but none were receiving any biologic agents.

Alteration in tryptophan catabolism in the early stages of RA and associations with inflammation and disease activity. To identify whether metabolic alteration contributes to arthritis initiation and development, we employed untargeted metabolomics to compare metabolite profile changes in the serum between patients with RA and HCs. Here, principal component analysis (PCA) revealed and detected significant differences in serum metabolites between RA patients and HCs (Supplemental Figure 1A). Compared with HCs, there were 620 different metabolites detected in the serum of patients with RA, of which 291 metabolites were decreased and 329 metabolites were increased (Supplemental Figure 1B). The heatmap illustrated the top 50 differential serum metabolites between patients with RA and HCs (Supplemental Figure 1C). Database analysis using Kyoto Encyclopedia of Genes and Genomes (KEGG), Human Metabolome Database (HMDB), PubChem, and the Small Molecule Pathway Database (SMPDB) indicated that differential metabolites in the RA group were primarily enriched in pathways related to D-glutamine and glutamate metabolism, arginine biosynthesis, arachidonic acid metabolism, and tryptophan metabolism (Supplemental Figure 1D). Further examination of these differential metabolites revealed that amino acid metabolites were the most prominent category in patients with RA (Figure 1A; Supplemental Table 3). Meanwhile, we noticed that the levels of indole derivatives, which constitute one of the metabolic pathways of tryptophan, were significantly decreased in patients with RA. The abundances of lipid metabolites in the RA group displayed as following: fatty acid metabolites were increased, including eicosapentaenoic acid, itaconic acid, heptanoic acid, and 20-Hete. Prenol lipids were also elevated, whereas steroid metabolites, including testosterone sulfate, cortisone, 11-aminoundecanoic acid, and estradiol sulfate, were mostly decreased. Additionally, levels of purine nucleotides, keto acids, and their derivatives were elevated in patients with RA, whereas bilirubin metabolism was reduced compared to HCs.

Our previous study revealed a disturbed serum metabolite profile existing in the PreRA stage.[1] Using the same PreRA cohort metabolism data, we conducted comparative analyses for 620 altered metabolites across KEGG, HMDB, PubChem, and

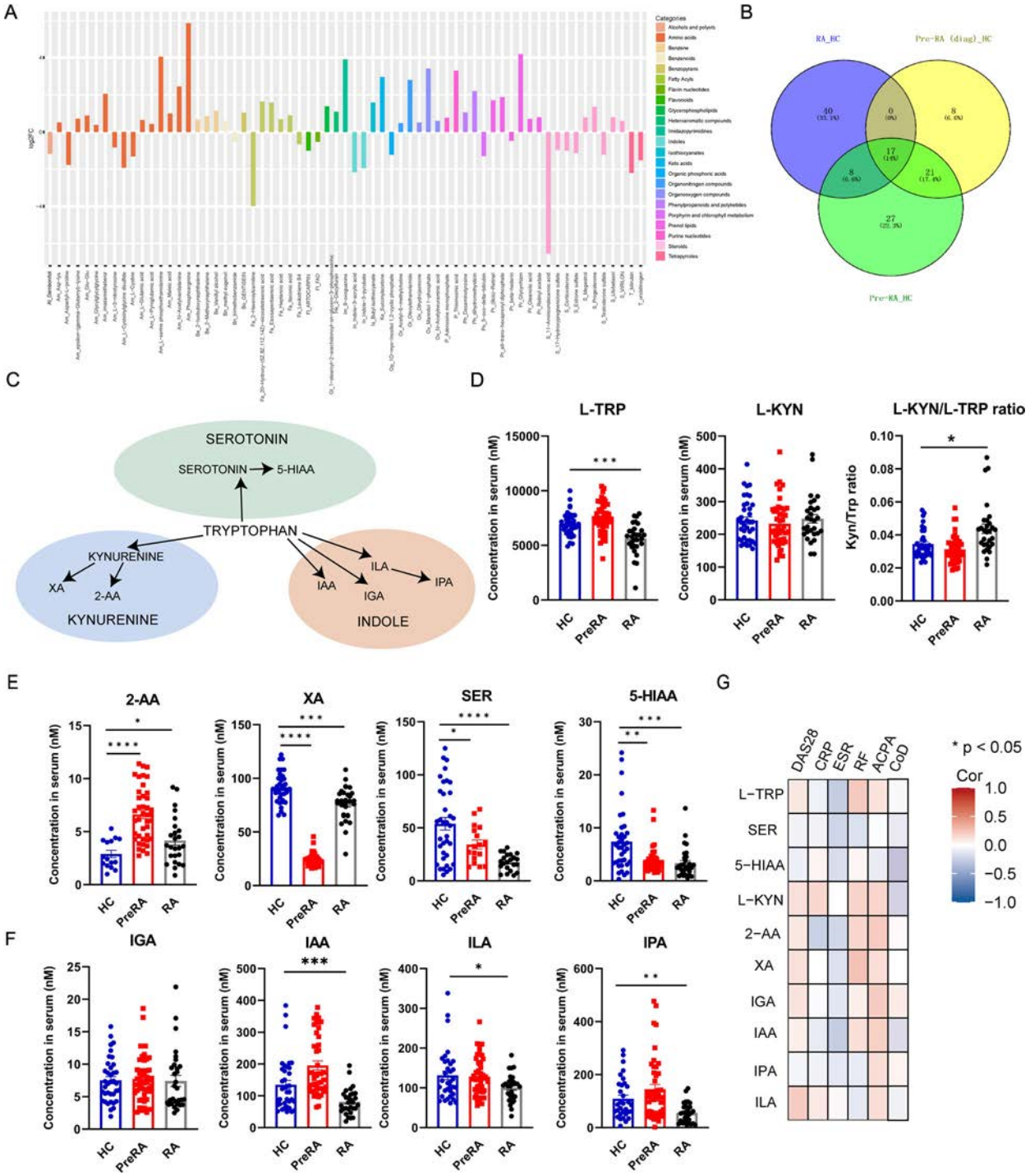


Figure 1. Lowered serum tryptophan metabolites are associated with PreRA and RA stages. (A) Differential metabolite categories between RA and HC groups. Differential metabolites with $\log_2$ fold change ≥0.5 or ≤−0.5 are displayed on the y-axis (RA: n = 30; HC: n = 38). (B) Venn diagram illustrating the overlapping metabolites among RA vs HC, PreRA vs HC, and PreRA (diagnosis) vs HC. (C) Schematic diagram of three major tryptophan metabolic pathways. (D–F) Levels of tryptophan-related metabolites in serum of RA, PreRA, and HC groups. (G) Correlation analysis of CoD, ACPA, RF, ESR, CRP, and DAS28 indexes and tryptophan metabolites in patients with RA by Spearman analysis. Data are expressed as mean ± SEM. Statistical significance was assessed by one-way analysis of variance followed by Bonferroni post hoc tests. $^{*}P < 0.05$, $^{**}P < 0.01$; $^{***}P < 0.001$. $^{****}P < 0.0001$. 2-AA, 2-aminobenzoic acid; 5-HIAA, 5-hydroxyindoleacetic acid; ACPA, anti-citrulline protein antibody; CoD, course of disease; CRP, C-reactive protein; DAS28, Disease Activity Score in 28 joints; ESR, erythrocyte sedimentation rate; HC, healthy control; IAA, indole-3-acetic acid; IGA, 3-indoleglyoxylic acid; ILA, indole-3-lactic acid; IPA, indolepropionic acid; L-KYN, L-kynurenine; L-TRP, L-tryptophan; PreRA, high risk for RA; RA, rheumatoid arthritis; RF, rheumatoid factor; SER, serotonin; XA, xanthurenic acid.

SMPDB databases to identify commonly altered metabolites among three groups: patients with PreRA versus HCs, patients with PreRA who progressed to RA (diagnosed) versus HCs, and patients with RA versus HCs. Compared with the HC group, 65 differentially abundant metabolites were identified in the RA group, 46 differentially abundant metabolites were identified in the PreRA (diagnosed) group, and 73 differentially abundant metabolites were identified in the PreRA group. Venn diagram analysis identified 17 different metabolites commonly shared by the three groups (Figure 1B). The 17 different metabolites were then assigned to different categories (Supplemental Figure 1E). 20-Hete, progesterone, megestrol, and retinal levels were greater in RA and PreRA serum samples than in the those of HCs, but the levels of indole-3-pyruvate, indole-3-acrylic acid, and bilirubin were lower in RA and PreRA serum samples (Supplemental Figure 1E).

We further focused on Trp metabolism, given its established link to gut microbiome alterations and PreRA disease status.[1] Using Trp-targeted metabolomics, we analyzed changes in Trp metabolites. Trp can be catabolized by three basic metabolic pathways, namely, the Kyn, serotonin, and indole pathways (Figure 1C). The L-tryptophan (L-TRP) level was comparable between patients with PreRA and HCs but significantly decreased in the sera of patients with RA. The level of L-kynurenine (L-KYN) was not affected in individuals with PreRA or RA. However, the ratio of tryptophan to Kyn in the RA group was greater than in HC group (Figure 1D). Levels of 2-aminobenzoic acid, a Kyn pathway metabolite, were significantly greater in PreRA and RA groups, whereas xanthurenic acid (XA) levels were lower than in controls (Figure 1E). The levels of serotonin and its downstream metabolite 5-hydroxyindoleacetic acid (5-HIAA) gradually decreased in PreRA and RA groups compared with the HC group (Figure 1E). The 3-indoleglyoxylic acid (IGA) level was unaffected. However, the contents of IAA, ILA, and IPA, the indole derivatives of tryptophan, were significantly decreased in the serum of the RA group (Figure 1F), suggesting the decreased indole metabolism of tryptophan.

To assess whether tryptophan metabolism is associated with disease development, we then performed correlation analyses between the altered tryptophan metabolites and the course of disease (CoD), DAS28, markers of inflammation including C-reactive protein (CRP) and erythrocyte sedimentation rate (ESR), and antibody titers (RF and ACPA). L-Kyn was positively correlated with ACPA, RF, CRP, and DAS28. In contrast, we observed a negative correlation between 5-HIAA, IAA, ILA, and IPA and CoD, ESR, RF, and DAS28 (Figure 1G). Taken together, these data suggest positive regulation of the Kyn metabolic pathway but an overall negative regulatory role of the serotonin or indole metabolic pathway in RA.

Gut microbiome alterations in patients with PreRA and RA induce changes in murine tryptophan metabolism. The gut microbiota–derived metabolites play crucial roles in maintaining intestinal homeostasis and immunologic balance. Because previous studies have indicated that the gut microbiota from PreRA or RA contributes to arthritis development,[1,30] we investigated which metabolites may contribute to FMT-induced arthritis exacerbation in mice. To do so, we inoculated the mice with human fecal microbiota from HCs, patients with PreRA, and patients with RA into germ-free mice and subsequently induced an arthritis model (Figure 2A). The ace and Chao indices were reduced in the PreRA and RA stool-transplanted mice compared to the HC stool-transferred mice, indicating lower intestinal bacterial diversity, which mirrored the change in HC, PreRA, and RA groups (Supplemental Figure 2A and D). At the genus level, those with PreRA and RA exhibited decreased abundances of *Bacteroides*, *Romboutsia*, and *Lactobacillus* genera bacteria compared to the HC population, whereas patients with RA showed significant increased abundances of *Lachnoclostridium* and *Coprococcus 2* (Supplemental Figure 2B and C). Consistently, the gut microbiome of the receiving mice also displayed different microbiome compositional changes after HC, PreRA, and RA fecal transfer (Supplemental Figure 2E). At the genus level, the proportions of the gut microbiota, especially genera *Lactobacillus*, *Bacteroides*, *Blautia*, and *Parabacteroides*, significantly differ among three groups (Figure 2B). Notably, there were significantly reduced abundances of *Lactobacillus*, *L reuteri*, *Bacteroides fragilis*, and *Clostridium spp* bacteria in PreRA and RA stool-transplanted mice (Supplemental Figure 2B). Moreover, compared with those of the HCs, more severe arthritis and clinical scores were observed in the knees of the PreRA and RA groups of mice after the second boost immunization (Figure 2C and D). Moreover, inflammatory cells, synovial hyperplasia, and bone destruction were also detected in the PreRA and RA stool-transferred mice compared with in the HC and PBS groups (Figure 2C). The PreRA and RA stool also caused an intestinal barrier disruption, as evidenced by the injured microvilli in the small intestine of FMT-transferred mice and significantly increased dextran uptake (Figure 2E and F). Consistently, increased frequencies of Th17 cells were detected in the mLNs of the PreRA and RA groups (Figure 2G). The proportions of Treg cells gradually decreased in PreRA and RA fecal-transferred mice (Figure 2H), Meanwhile, the percentages of Th1 and Th2 cells were comparable in the mLNs and PP among HC, PreRA, and RA-FMT mice (Supplemental Figure 3A and B). No significant differences in T helper cell subsets were observed in the lamina propria among HC, PreRA, and RA-FMT mice (Supplemental Figure 3C).

Untargeted metabolomics analysis indicated dysregulated oxidized lipid and tryptophan metabolism in PreRA and RA populations (Supplemental Figure 1C).[1] To determine which metabolites are involved in the aggravated arthritis caused by gut microbiota of those with PreRA or RA, we detected lipid and tryptophan metabolism using lipidomics and Trp-targeted metabolomics, respectively. However, the levels of oxidized lipid metabolites were comparable among HC, PreRA, and RA-FMT

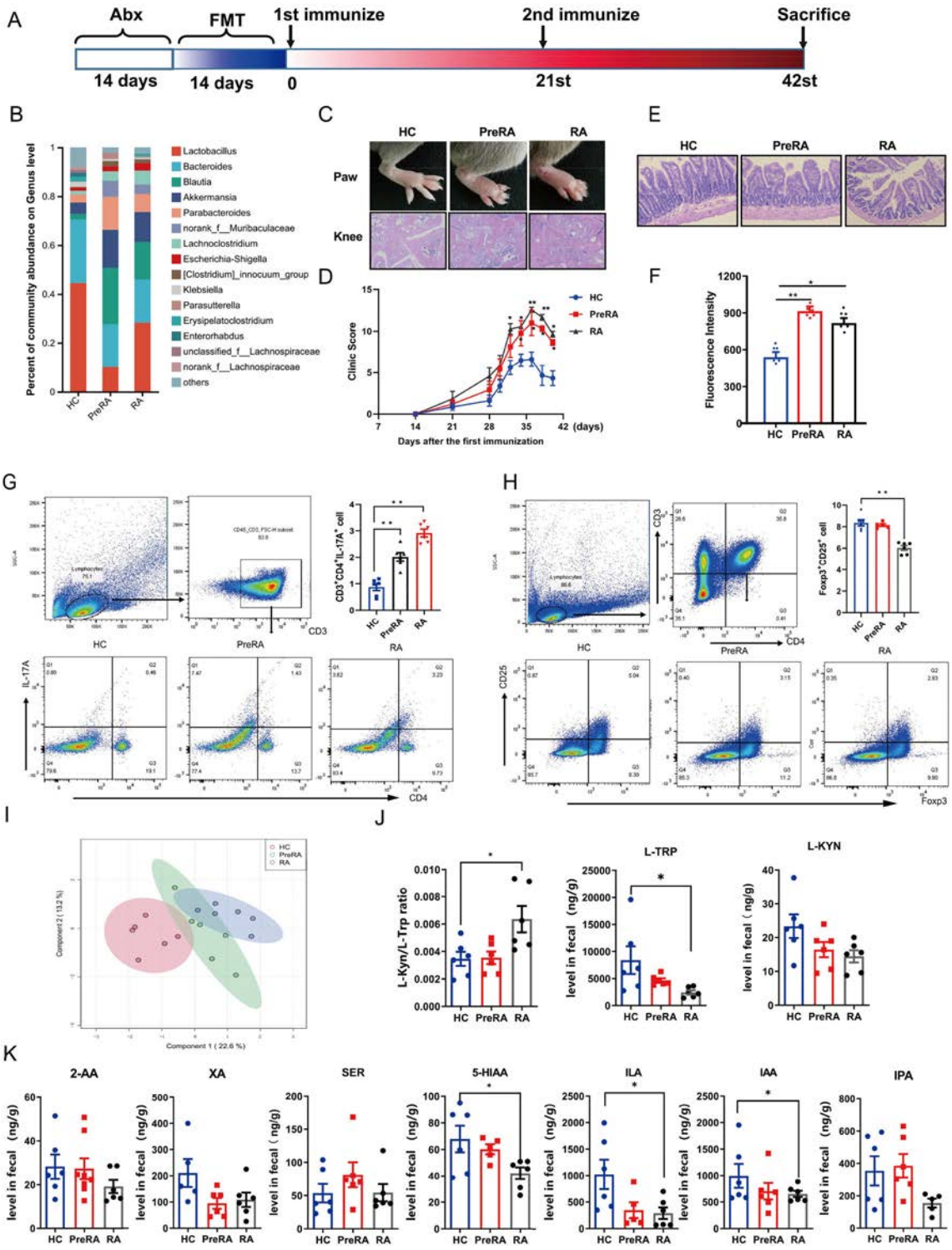


Figure 2. Fecal microbiota from patients with RA promote arthritis development and alter tryptophan metabolism. (A) Schematic representation of FMT and induction of arthritis (n = 6 per group). (B) Differences in bacterial genus distribution and abundance among HC, PreRA, and RA-FMT mice. (C) Representative image of mice hind paw and H&E staining images of the knees from HC, PreRA, and RA-FMT mice. (D) Arthritis score indicating disease severity in the three groups of mice. (E) Histologic assessment of microvilli integrity in the small intestine of FMT mice. (F) Fluorescence intensity of fluorescein isothiocyanate–conjugated dextran in serum from each group of mice. (G, H) Representative flow cytometry plots of Treg (G) and Th17 cells (H) and quantification of these two cells in mesenchymal lymphoid node. (I) Partial least-squares discriminant analysis comparison of serum metabolomes among HC, PreRA, and RA stool-transplanted mice. (J–K) Target detection of tryptophan-related metabolites in each group of mice. Data are expressed as mean ± SEM. Statistical significance was assessed by one-way analysis of variance followed by Bonferroni post hoc tests. $^{*}P < 0.05$, $^{**}P < 0.01$. Abx, antibiotics; FMT, fecal microbiota transplantation. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.43430/abstract.

mice (Supplemental Figure 4). PreRA, RA, and HC stool-transferred mice presented altered tryptophan metabolism profiles, as shown in Figure 2I. L-TRP levels gradually decreased in the HC, PreRA, and RA groups (Figure 2J). Although the level of L-KYN was not significantly affected, the L-TRP/L-KYN ratio was significantly increased in RA-FMT mice, which was similar to the change in the human cohort (Figure 2J). Notably, 2-ketoadipic acid (2-KA), a terminal metabolite of Kyn, was significantly increased in PreRA and RA stool-transferred mice (Supplemental Figure 5). In contrast, the level of the serotine metabolite 5-HIAA was reduced in RA-FMT mice (Figure 2K). Moreover, the levels of N-formylkynurenine, an intermediate of tryptophan metabolism, and ILA and IAA, indole metabolites, were significantly lower in the PreRA and RA groups than in the HC group (Figure 2K; Supplemental Figure 5). Meanwhile, the levels of other tryptophan metabolites including 2-KA, 2-AF, 5-HTOL, NAS, IEt and TRM, ICAld, IGA, IA, and IAM were comparable among three groups. Collectively, these data suggest a metabolic shift from serotonin and indole pathways to the Kyn pathway in arthritic mice.

ILA//IAA supplementation negatively regulates arthritis in mice by enhancing gut barrier function and promoting Treg cells. Indole metabolism is entirely dependent on the gut microbiota. Given the reduced levels of indole pathway metabolites observed in FMT-treated mice, we hypothesized that these indole metabolites, particularly the derivatives ILA and IAA, may inversely correlate with arthritis incidence. To test this possibility, we induced arthritis in RA stool-transferred mice (FMT-CIA) and orally administered mice with different doses of ILA and IAA (5 and 20 mM)[31,32] on day 21 before the second immunization. As shown in Figure 3A and B, ILA and IAA supplement significantly reduced the joint and paw swelling of FMT-CIA mice, accompanied by the lower clinical arthritis scores and inflammatory cells infiltration in the joints. Meanwhile, the levels of TNF-α and IL-1β in serum was also significantly reduced in arthritis mice. In addition, we detected a significant increase in Treg cells in the spleen of ILA- and IAA-treated CIA mice (Figure 3D and E). However, no significant difference was detected in the proportions of Treg in the mLNs or PP (Supplemental Figure 6). Th1, Th2, and Th17 cells were comparable in mLNs or PP among different groups (Supplemental Figure 6). Because an intestinal barrier dysfunction also contributes to arthritis occurrence and development,[1,2] we performed immunochemical staining to measure the expression of ZO-1 and occludin in the small intestine of mice. ILA treatment, but not IAA, was associated with increased ZO-1 and occludin staining in the small intestine, as confirmed by enhanced dextran uptake (Figure 3F and G).

CIA mice also displayed altered Typ metabolism similar to human patients with RA, with reduced serum levels of indole derivatives.[15] To evaluate the therapeutic potential of ILA and IAA in this model, we supplemented CIA mice with ILA or IAA as well. As shown in Figure 3H and I, CIA mice treated with ILA or IAA displayed less paw swelling and the lower clinical scores from days 21 to 49. Moreover, the levels of the proinflammatory cytokines TNF-α, IL-1β, and IL-6 were also significantly reduced in the CIA mice after ILA or IAA supplement (Figure 3J). Similar to the RA-FMT arthritis mice, we also found a significant expansion of Treg cells in the spleens of ILA- and IAA-treated CIA mice (Figure 3K and L, Supplemental Figure 7) and impaired gut barrier function (Figure 3M and N). Collectively, these data suggest that ILA and IAA exert negative regulation on arthritis development by expanding Treg cells and enhancing gut barrier function, also suggesting their therapeutic potential for arthritis treatment.

ILA/IAA promotes Treg cell expansion via AhR activation. Indole derivatives are established endogenous AhR agonists.[33] To determine whether AhR activation mediates ILA- or IAA-induced Treg expansion and gut barrier repair in arthritic mice, we isolated naive $CD4^+$ T cells from control mice and differentiated them into Treg cells in vitro under ILA or IAA treatment. As shown in Figure 4A–C, ILA (10, 100 μM) and IAA (10 μM) significantly enhanced Treg cell differentiation, accompanied by increased Foxp3 gene expression. Notably, the expression of CYP1A, an AhR-regulated target gene, was dramatically increased by different concentrations of ILA or IAA at the transcriptional and protein levels (Figure 4D–F), suggesting that AhR activation occurred when the cells were exposed to ILA or IAA. However, the co-treatment with CH223191 (10 μM), a known AhR antagonist,[34] significantly reversed ILA- or IAA-induced Treg expansion and reduced gene expression of Foxp3 (Figure 4G–I). Moreover, ILA or IAA facilitated the nuclear translocation of AhR, as shown by immunofluorescence staining (Figure 4J), and the formation of AhR/ARNT complexes (Figure 4K). Finally, the activity of the XRE-luciferase reporter gene was also elevated in cells after ILA or IAA treatment (Figure 4L). These results indicate that ILA or IAA induces Treg cell expansion in an AhR-dependent manner.

ILA enhances intestinal tight junctions via AhR activation. ILA treatment increased the expression of intestinal tight junction proteins ZO-1 and occludin in CIA mice. We next tested whether ILA regulates the expression of these proteins in vitro. To do so, we first treated Caco-2 cells with different concentrations of ILA (0.5, 1, 2, and 5 mM). As shown in Figure 5A–C, ILA dose-dependently increased Zo-1 and occludin expression at the transcriptional and protein levels. We further assessed ILA's protective effect against gut barrier disruption by treating Caco-2 cells with fecal supernatants from patients with RA (Figure 5D). Stimulation with the RA stool fluid caused a reduction in Zo-1 and occludin expression, as detected by RT-PCR and immunofluorescence staining (Figure 5E and K), but HC stool fluid did not cause these changes of Zo-1 and occluding

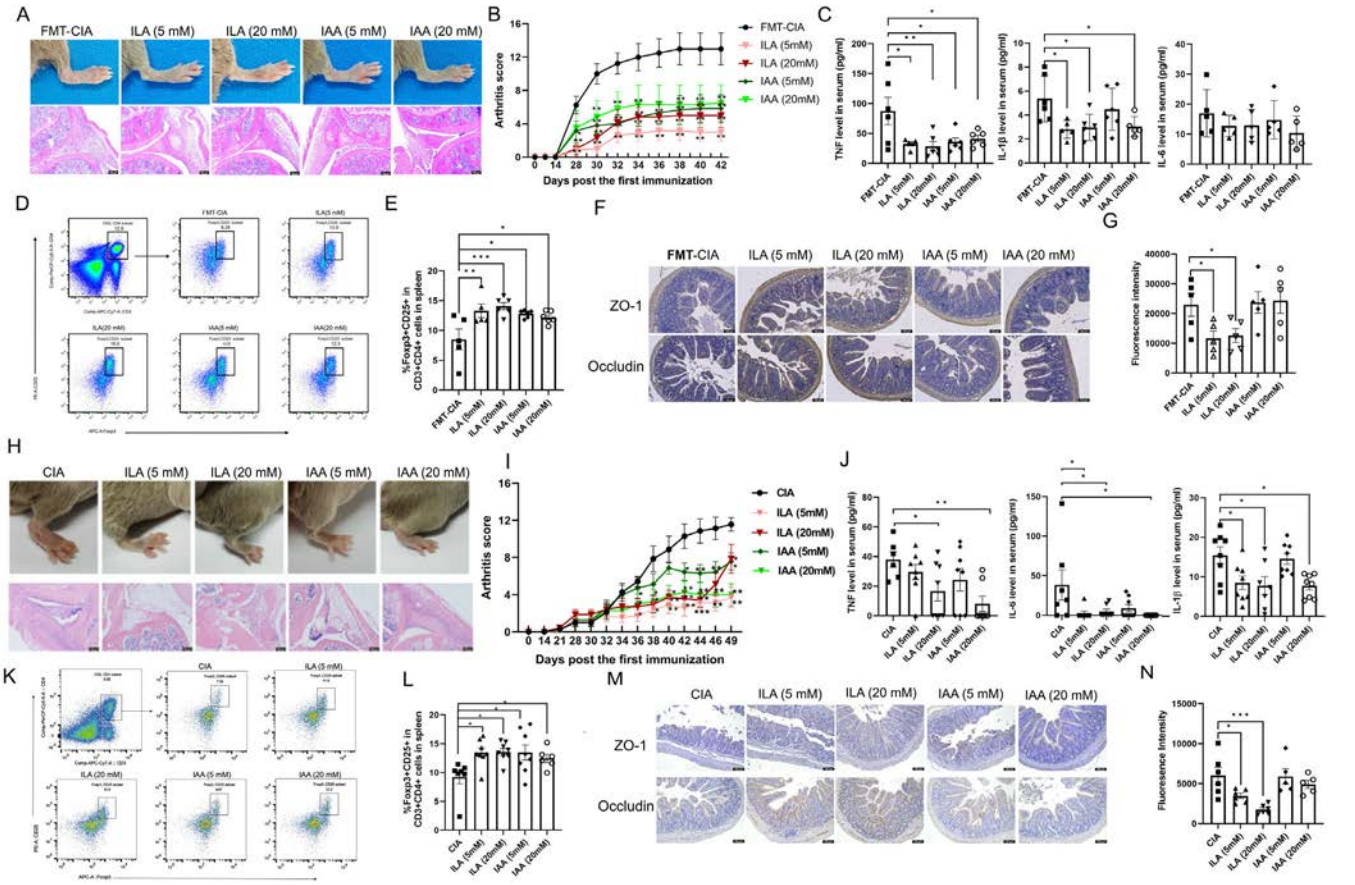


Figure 3. ILA/IAA supplement effectively reduces arthritis in mice by enhancing gut barrier function and Treg cell promotion. (A) Representative photographs of paw swelling and H&E staining images of the knees in RA fecal-transferred mice FMT with collagen-induced arthritis. (B) Clinical scores in arthritis mice receiving different doses of ILA (5, 20 mM) and IAA (5, 20 mM). (C) The levels of TNF-α, IL-1β, and IL-6 in sera of different groups. (D, E) Representative flow cytometry plots (D) and quantification of Treg cells (E) in spleen. (F) The expressions of Zo-1 and occludin in the small intestine of mice using histochemical staining. (G) Serum fluorescence intensity of FITC-conjugated dextran. (H) Paw image and H&E staining of knee in mice. (I) Clinical scores of CIA mice receiving ILA (5, 20 mM) and IAA (5, 20 mM). (J) TNF-α, IL-6, and IL-1β levels in sera . (K, L) Representative flow cytometry plots (K) and quantification of Treg cells (L) in spleen of mice. (M) The expressions of Zo-1 and occludin in the small intestine of CIA mice. (N) Serum fluorescence intensity of FITC-conjugated dextran. Data are expressed as mean ± SEM. Statistical significance was assessed by one-way analysis of variance followed by Bonferroni post hoc tests. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.43430/abstract.

(Supplemental Figure 8). However, incubation with ILA significantly promoted the expression of these two proteins in Caco-2 cells treated with RA stool fluid (Figure 5E and G–K). Considering that ILA is an endogenic metabolite agonist of AhR, we further determined whether AhR is also activated in Caco-2 cells after ILA treatment. Cyp1a1 gene expression was dramatically increased in ILA-treated Caco-2 cells (Figure 5F). Moreover, the increased expression of Zo-1 and occludin induced by ILA could also be significantly reversed by treatment with the AhR antagonist CH223191 (Figure 5G and H). In addition, CH223191 blocked the ILA-induced increase in Zo-1 and occludin in fecal-treated cells (Figure 5I and J). Taken together, these results indicate that ILA regulates these two gut barrier proteins by activating AhR.

AhR activation is essential for ILA/IAA-mediated arthritis suppression. Both ex vivo and in vivo experiments have shown that ILA or IAA can affect Treg or intestinal barrier function via AhR activation. To test whether the beneficial effect of these two metabolites on arthritis is dependent on AhR activation, we treated mice with CH223191 combined with or without ILA or IAA. CH223191 itself did not affect CIA compared with in the CIA group (Figure 6A). However, the beneficial effects of ILA or IAA on CIA disease symptoms, including paw swelling, inflammatory cell infiltration, and bone erosion in the knees, were nearly completely eliminated by CH223191 addition (Figure 6A and B). Additionally, the increased proportion of Treg cells caused by ILA or IAA was significantly reduced in the spleens of the mice after CH223191 treatment (Figure 6C and D). Moreover, CIA mice treated with ILA showed increased dextran uptake and ZO-1 and occludin-positive staining in the small intestine, whereas CH223191 also significantly abolished these changes (Figure 6E and F). Taken together, these data strongly suggest that ILA or IAA might alleviate arthritis by enhancing Treg cell and gut barrier function in an AhR-dependent manner.

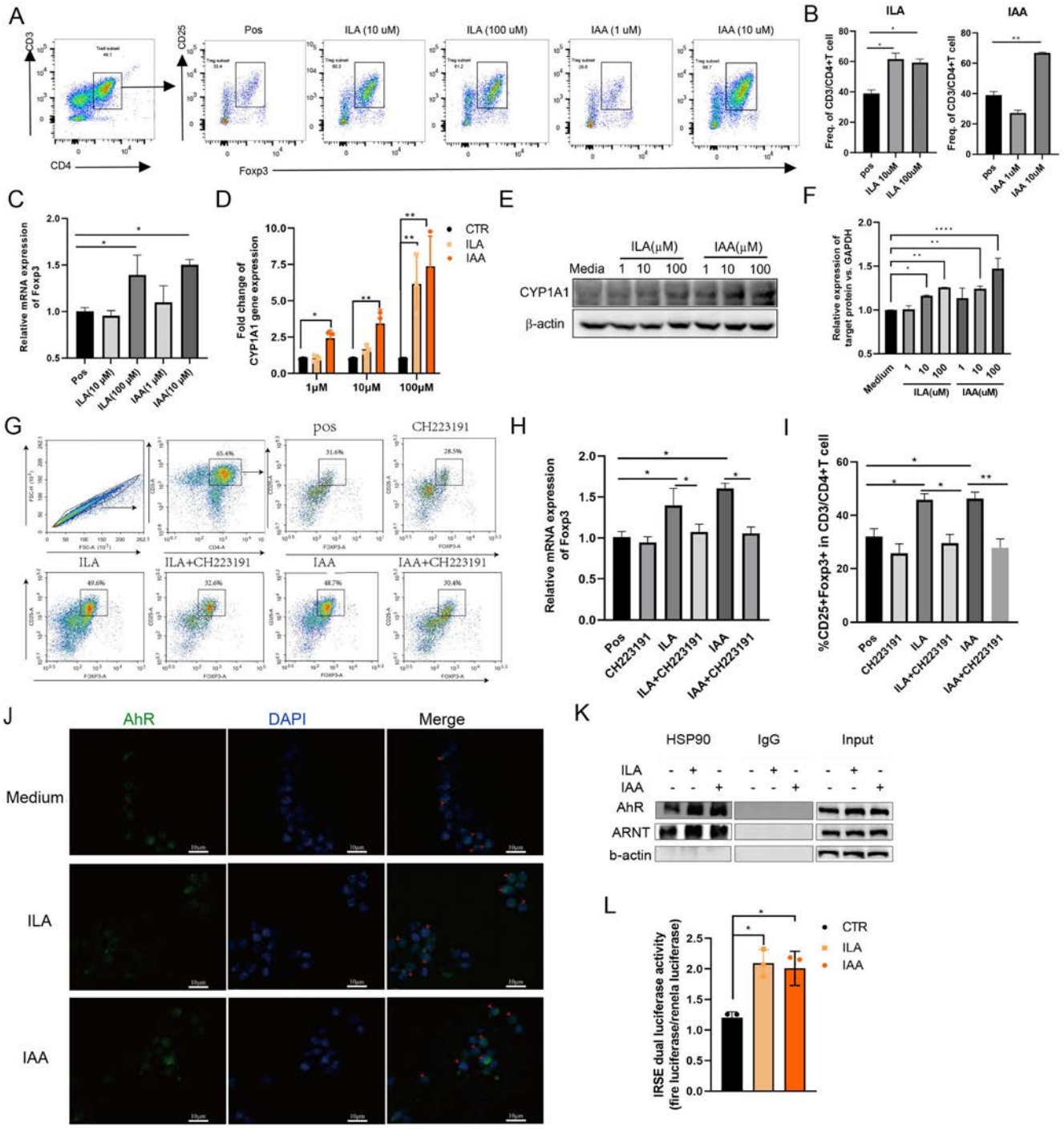


Figure 4. ILA and IAA promote T regulatory cell expansion via AhR activation. Representative flow cytometry plots (A) and quantification (B) of Treg cells in vitro after treatment with different doses of ILA and IAA. (C) Foxp3 gene expression in cells treated with ILA and IAA. (D) CYP1A1 messenger RNA expression in EL-4 cells. (E, F) Protein expression (E) and quantification (F) of CYP1A1. Representative flow cytometry plots (G) and quantification (I) of Treg cells and FoxP3 gene expression (H) in vitro after cells were incubated with ILA/IAA only or together with CH223191. (J) Splenic cells from mice were treated with ILA or IAA for one hour, and subjected to immunofluorescence staining (AhR, green, left column;DAPI, blue, middle column). (K) AhR and ARNT in the immunoprecipitate were visualized by the respective antibodies. (L) 293T cells were transfected with the XRE-luciferase construct and then were treated with ILA or IAA for 24 hours. Induced XRE-luciferase activity in 293T cells. Data were expressed as mean ± SEM. Statistical significance was assessed by one-way analysis of variance followed by Bonferroni post hoc tests. $^{*}P < 0.05$, $^{**}P < 0.01$; $^{***}P < 0.001$. $^{****}P < 0.0001$. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.43430/abstract.

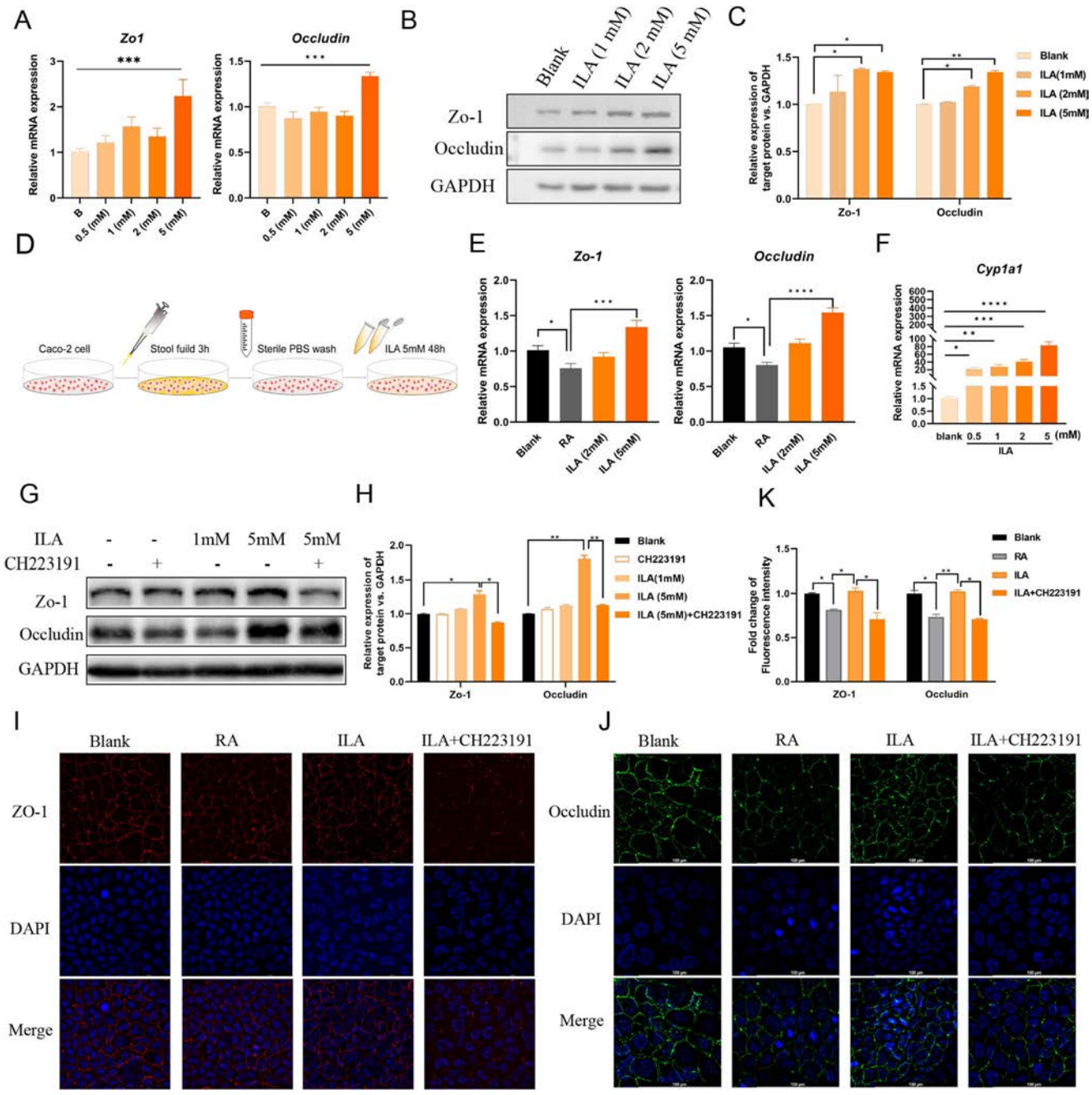


Figure 5. ILA enhances intestinal tight junctions via AhR activation. (A) Zo-1 and Occludin mRNA expressions in Caco-2 cells treated with ILA (0.5, 1, 2, 5 mM). Expression (B) and quantification (C) of Zo-1 and occludin in ILA-treated Caco-2 cells assessed by Western blot. (D) Schematic diagram of stool fluid isolated from HC or RA individuals incubated with Caco-2 cells, followed with the treatment with ILA in vitro. (E) RT-PCR analyses showing ILA repairing ZO-1 and occluding expression reduced by RA fecal supernatant. (F) The enhanced expression of downstream Cyp1a1 gene in Caco-2 cells treated by different concentrations of ILA. (G, H) Reversal of ILA-induced ZO-1 and occludin upregulation by AhR inhibition caused by ILA. (I–K) ILA-induced gut barrier repairing effect was suppressed by AhR inhibitor CH223191. Immunofluorescence staining (I, J) and quantification (K) of Zo1 (Red) and occludin (green) in Caco-2 cells. Data were expressed as mean ± SEM. Statistical significance was assessed by Student's *t*-test or one-way analysis of variance followed by Bonferroni post hoc test. $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.43430/abstract.

DISCUSSION

Multiple factors, including genes, smoking, infection, and gut microbiota dysbiosis, contribute to the breakdown of immune tolerance in the early stage of RA. However, the key step triggering autoimmune dysregulation may begin at the mucosal surfaces affected by the gut barrier dysfunction caused by intestinal microbiota dysbiosis. The understanding of the influence of the intestinal microbiota on host metabolism and the effects of metabolic

dysregulation on the phenotype and function of immune cells, such as T cells, macrophages, and B cells, provides critical insights into the pathogenesis. Therefore, identifying how metabolic dysregulation participates in immune dysregulation in RA is essential for elucidating the pathogenesis of RA and provides new therapeutic options. However, the role of specific metabolites in RA, particularly in the early stage of RA, remains largely unclear.

Here, we compared metabolite profiles between patients with PreRA and HCs and between patients with RA and HCs and identified 17 metabolites, some of which belong to the tryptophan and lipid metabolite category. The differences of tryptophan metabolism were not as obvious as lipid metabolism; however, FMT mouse experiments revealed more pronounced changes in tryptophan derivatives rather than oxidized lipids when comparing PreRA/RA-FMT mice with HC-FMT controls (Figure 2; Supplemental Figures 4 and 5). Gut microbiome plays a critical role in generating indole derivatives from tryptophan metabolism. Our initial interest was to elucidate how gut microbiota alteration in PreRA/RA contributes to immune dysregulation and disease initiation, so we applied targeted quantitative Trp metabolomics on serum from 53 individuals with PreRA, 30 with RA, and

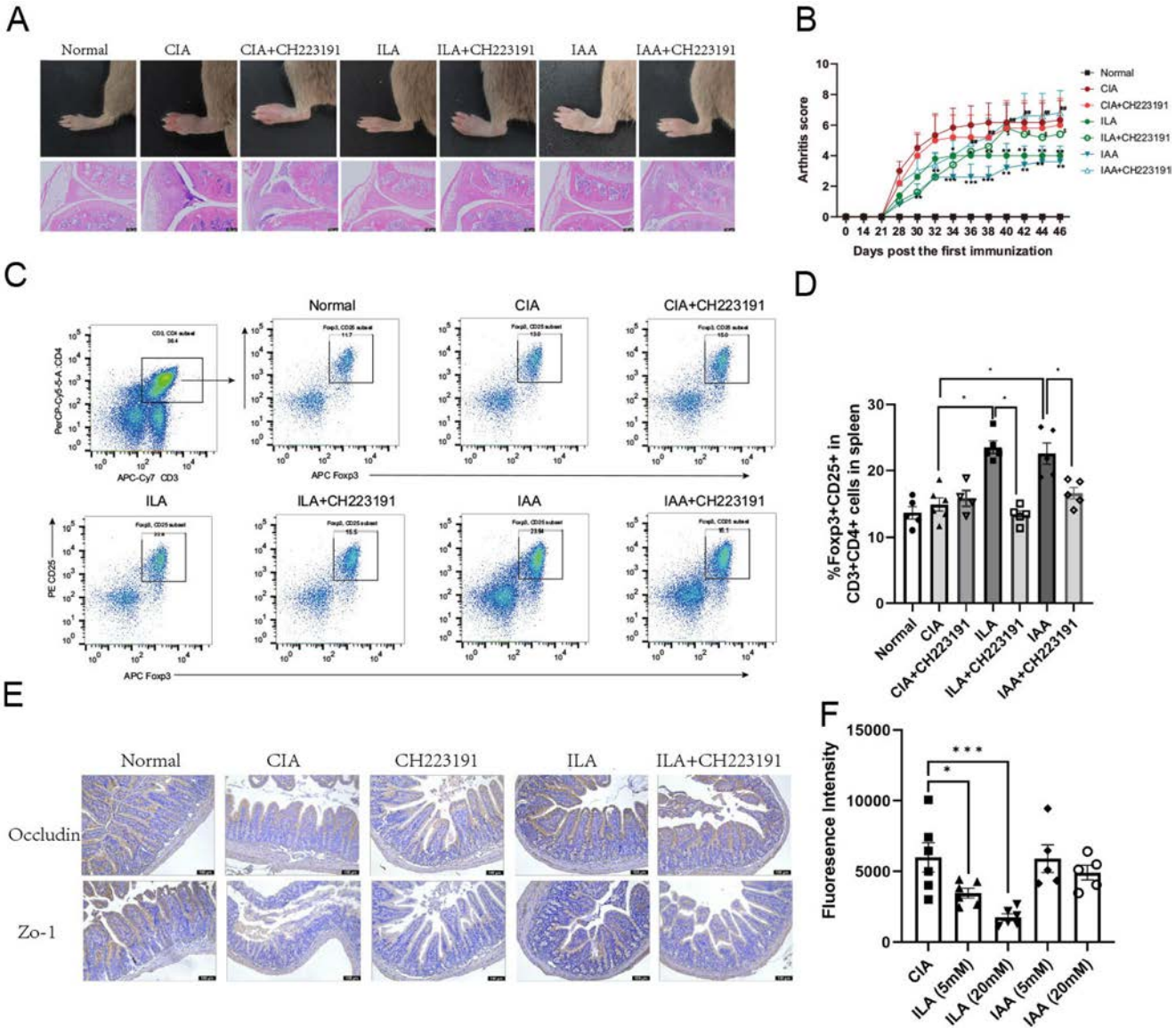


Figure 6. AhR inhibition blocks ILA- and IAA-induced improvement in arthritis of mice. (A) Representative images of paw swelling of mice and hematoxylin and eosin staining images of the knees. (B) Clinical scores of CIA mice (n = 6) and CIA mice receiving ILA (20 mM), IAA (5 mM), or AhR antagonist CH223191 (n = 8 per group). (C, D) Representative flow cytometry plots of Treg cells (C) and quantification of Treg cells (D) in spleen. (E) Histochemical staining of ZO-1 and occludin in the small intestine of mice. (F) Serum fluorescence intensity of fluorescein isothiocyanate–conjugated dextran. Data are expressed as mean ± SEM. Statistical significance was assessed by one-way analysis of variance followed by Bonferroni post hoc tests. $^{*}P < 0.05$, $^{**}P < 0.01$. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.43430/abstract.

38 HCs; we identified decreased circulating Trp concentrations and a high Kyn/Trp ratio, consistent with prior findings.[8,12] Moreover, the serotonin and indole metabolic pathways of tryptophan seemed to be decreased, as evidenced by the lower levels of 5-HIAA and ILA. Moulin et al reported lower levels of KYNA and XA acids and indole derivatives.[15] In this study, we did not observe similar reductions in our population, which may be due to racial differences and sample sizes. However, we did observe a decrease in the indole metabolic pathway, which is consistent with what Moulin et al reported.

Discrepancies also exist regarding the indole pathway of tryptophan. Seymour et al identified an increased tryptophan indole metabolism in CIA mice in a microbiome-dependent manner. However, the report of Moulin et al and our findings in human cohorts showed an opposite change in indole levels. Notably, we observed a similar change in tryptophan catabolism in the PreRA or RA fecal transfer mouse model: an increase in the Kyn metabolic pathway but a decrease in serotonin and indole pathway metabolites, with lower levels of 5-HIAA, ILA, and IAA. Because several *Lactobacillus* and *Bacteroides* species could catabolize Trp into ILA or IAA,[16,18] the reduced levels of *Lactobacillus*, *L reuteri*, and *B fragilis* bacteria in stool-transplanted mice likely underlie the decreased indole derivatives observed in our study. Moulin et al reported similar alterations in two experimental models of arthritis in mice, namely, CIA and the collagen antibody-induced arthritis (CAIA) mice. CIA mice exhibit most of the metabolic disorder characteristics associated with human rheumatoid arthritis. The levels of indole derivatives in their serum are decreased, while the levels of kynurenic acid and xanthine acid are reduced, while the level of quinolinic acid (QUIN) is increased.[15] The controversial findings regarding the level of indole may be due to differences in the detection methods, time points, and sample types used by the researchers.

There are conflicting findings concerning the roles of tryptophan metabolites in RA experimental arthritis models. 5-HIAA and 5-hydroxytryptophan reportedly improve arthritis severity in CIA mice.[35,36] In contrast, microbiome-dependent indole production plays an essential role in the development of CIA in mice.[14] Intra-articular injection of indole into rabbit knees can also induce RA-like synovitis.[37] Thus, specific tryptophan metabolites may play specific roles in the development of autoimmune arthritis. In our study, we verified that the RA gut microbiome exacerbates arthritis and decreases the number of indole pathways associated with tryptophan. Interestingly, supplementation with the indole derivatives ILA and IAA sufficiently suppressed the arthritis symptoms, indicating their potential as negative immunoregulatory factors on the development of arthritis in mice.

The breakdown of immune tolerance precedes the onset of clinical symptoms in patients with RA. We previously reported disturbances in the oral and intestinal flora in patients with PreRA.[1,38] Recent studies have pointed to three sites that may represent the initial site of breaking tolerance in RA: (1) the lungs, (2) the oral cavity, and (3) the intestinal tract.[39] Transplantation of fecal bacteria from PreRA or RA patients into mice induced intestinal barrier disruption and mucosal immune imbalance[1] (Figure 2). The intestinal mucosal barrier includes the chemical barrier, immune barrier and microbial barrier. The tight junctions among intestinal epithelial cells are composed of tight junction protein (ZO-1), sealing protein (claudin), and occludin, which form a physical mechanical barrier. The colonization of *Prevotella histicola* from the human gut into CIA mice can reduce intestinal permeability, which is associated with increases in ZO-1 and occludin expression in the jejunum, ileum and colon, and the severity of arthritis in mice is also significantly reduced.[40] Zonulin is a known regulator of tight junctions between cells. Zonulin protein reduces intestinal barrier function by releasing the binding of ZO1 and occludin in the tight junction complex. Recent studies have shown that the level of zonulin in the peripheral blood of patients with early RA is significantly increased, accompanied by increased intestinal permeability.[2] The protein level of zonulin in CIA mice also showed similar changes, and the joint inflammation of CIA mice was reduced and the number of osteoclasts was reduced after treatment with the zonulin antagonist larazotide, whereas the zonulin agonist AT-1002 aggravated their symptoms.[2] These findings underscore that a decrease in intestinal permeability may be the key event mediating the disruption of early immune tolerance in RA and that the enhancement of intestinal tight junctions plays an important role in the early treatment of RA. In our study, both RA and PreRA donors exhibited reduced gut microbiota diversity. Correspondingly, mice receiving fecal transplants from these donors also showed decreased microbial diversity. Principal coordinates analysis (PcoA) further confirmed distinct gut microbiome compositions in the recipient mice (Supplemental Figure 2E). The rapid death of some anaerobic gut microbiota during stool preparation preventing the identical of the specific changes in bacterial taxon abundances between donors and recipient mice; however, the majority of relatively high-abundance bacteria displayed similar abundance patterns across the HC, RA, and PreRA groups. More importantly, in this study, we did not observe increased zonulin expression in the serum (data not shown) but did detect decreases in Zo-1 and occludin expression in the small intestine. ILA is known to activate the AhR, Nrf2, and Stat1 signaling pathways and inhibit the release of IL-8 in intestinal epithelial cells and macrophages induced by LPS, TNF-α, or IL-1β.[41,42] In addition, ILA can affect the expression of genes related to intestinal cell morphology and development in humans and mice,[43] reprogramming of intraepithelial $CD4^+$ T helper cells, and polarization of mouse Th17 cells.[22] Here, we found that ILA directly enhanced Zo-1 and occludin expression in a Caco-2 cell culture system. More importantly, the reduction in Zo-1 and occludin in epithelial cells caused by RA stool fluid can also be reversed by ILA incubation, suggesting the protective role of ILA in intestinal permeability in RA.

Beyond intestinal barrier disruption, the disturbance of mucosal immune homeostasis caused by gut bacterial dysbiosis in the gut also affects RA initiation. Studies have shown that intestinal segmented filamentous bacteria (SFB) enhance Th17 cell activation, directly leading to the early onset of RA.[3] In addition to SFB, other bacteria, such as *Porphyromonas gingivalis*, skew the type II collagen (CII)-specific T cell response toward the Th17 phenotype in lymph nodes draining arthritic joints.[44] Treg cells in the gut are the main cell types that maintain the intestinal immune balance. In addition to being regulated by dentritic cells, Treg cells can also be directly regulated by a variety of metabolites, such as short-chain fatty acids, and increasing their proportion and function can significantly improve the symptoms of arthritis in experimental mice. Marietta et al isolated and cultured *P histicola* from the gut of healthy mice and reported that gavage of these symbiotic bacteria into RA-susceptible HLA-DQ8 mice increased the expression of antimicrobial peptidase and the expression of the intestinal epithelial close-binding protein ZO-1, reduced the permeability of the intestinal mucosa, promoted the production of intestinal Treg cells, and significantly reduced the occurrence of type II CIA.[40] Dietary interventions, such as adding acetic acid or propionate to drinking water or feeding resistant starch, also improved arthritis outcomes by increasing gut *Bacteroides* and Treg cell abundance.[45] In this study, both IAA and ILA suppressed arthritis symptoms in CIA mice by increasing the number of splenic Treg cells without affecting Th17 and Th1 cells in various lymphoid organs.

AhR is a ligand-activated heterodimeric transcription factor, and its agonists include endogenous and exogenous ligands, among which tryptophan metabolites are the main physiologic endogenous ligands of AhR. AhR is widely expressed in immune cells, epithelial cells, central glial cells, and neurons and plays an important role in cell differentiation and function.[27] Previous studies have shown that AhR can play an important role in the differentiation, maintenance, homing, and functional regulation of intestinal Treg cells through various mechanisms, and these roles and molecular mechanisms are closely related to the ligand type. The exogenous ligand 2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD) can trigger the differentiation of mouse $Foxp3^+$ T regulatory cells by promoting AhR-ARNT to form dimers and then bind to XRE reaction elements.[46] The impact of AhR activation on RA varies with ligand and cell type. The AhR agonist TCDD negatively regulates the proliferation and differentiation of osteoblasts in CIA mice.[47] Some exogenous AhR ligands, such as tetrandrine or norisopordine, can improve arthritis in mice by activating AhR receptors and increasing the number of intestinal Treg cells.[48,49] In our study, ILA and IAA exerted AhR-dependent protective effects in CIA mice. When CH223191 was given to CIA mice, it significantly reversed the ILA- and IAA-induced improvement in arthritis. Moreover, in the Treg cell differentiation system, ILA and IAA significantly induce CYP1a1 expression, causing AhR-ARNT dimer formation and binding to the XRE. Importantly, CH223191 significantly blocked ILA- and IAA-induced Treg expansion. In the Caco-2 cell culture system, CH223191 also reversed ILA-induced Zo-1 and occludin expression.

Our study has several limitations. First, the individuals with PreRA were not treated with any biologics/DMARDs and denied regular use of NSAIDs at the time of enrollment. Most patients with RA were undergoing DMARD, NSAID, or prednisone therapy. These drugs could be a potential confounder in the analysis of metabolites. Thus, this discrepancy may introduce confounding effects in metabolite analysis. Second, it is ideal to study the metabolite changes in both serum and stool samples from FMT-transferred mice. Meanwhile, due to the sample limitation, we only studied the Trp metabolism in the stool of mice. Although we observed the similar changes of Trp metabolism to that in serum of human patients with RA, it is still better to check the Trp metabolites in serum of FMT mice.

In summary, our findings reveal a novel mechanism by which alterations in tryptophan metabolism are implicated in the pathogenesis of RA. Supplementation with the reduced indole derivatives ILA and IAA has protective effects on CIA mice by enhancing Treg cell expansion and gut barrier function. This work expands the understanding of the roles of tryptophan metabolism in the development of autoimmune arthritis and highlight the therapeutic potential to prevent or treat RA in the early stage.

ACKNOWLEDGMENT

We would like to thank Qiaorong Huang for flow cytometry measurement and analysis (Research Core Facility of West China Hospital, Sichuan University, Chengdu, China).

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr Luo confirms that all authors have provided the final approval of the version to be published and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. Luo Y, Tong Y, Wu L, et al. Alteration of gut microbiota in individuals at high-risk for rheumatoid arthritis associated with disturbed metabolome and the initiation of arthritis through the triggering of mucosal immunity imbalance. Arthritis Rheumatol 2023;75(10):1736–1748.
2. Tajik N, Frech M, Schulz O, et al. Targeting zonulin and intestinal epithelial barrier function to prevent onset of arthritis. Nat Commun 2020;11(1):1995.
3. Evans-Marin H, Rogier R, Koralov SB, et al. Microbiota-dependent involvement of Th17 cells in murine models of inflammatory arthritis. Arthritis Rheumatol 2018;70(12):1971–1983.

4. Yu D, Du JP, Pu X, et al. The gut microbiome and metabolites are altered and interrelated in patients with rheumatoid arthritis. Front Cell Infect Microbiol 2022;11:763507.
5. Polinski KJ, Bemis EA, Yang F, et al. Association of lipid mediators with development of future incident inflammatory arthritis in an anti-citrullinated protein antibody-positive population. Arthritis Rheumatol 2021;73(6):955–962.
6. Luan H, Gu W, Li H, et al. Serum metabolomic and lipidomic profiling identifies diagnostic biomarkers for seropositive and seronegative rheumatoid arthritis patients. J Transl Med 2021;19(1):500.
7. McMillan M. The identification of a fluorescent reducing substance in the urine of patients with rheumatoid arthritis - the excretion of 3-hydroxyanthranilic acid in this and other conditions. J Clin Pathol 1960;13(2):140–148.
8. Schroecksnadel K, Winkler C, Duftner C, et al. Tryptophan degradation increases with stage in patients with rheumatoid arthritis. Clin Rheumatol 2006;25(3):334–337.
9. Widner B, Sepp N, Kowald E, et al. Enhanced tryptophan degradation in systemic lupus erythematosus. Immunobiology 2000;201(5): 621–630.
10. Åkesson K, Pettersson S, Ståhl S. Kynurenine pathway is altered in patients with SLE and associated with severe fatigue. Lupus Sci Med 2018;5(1):e000254.
11. Berlinberg AJ, Regner EH, Stahly A, et al. Multi 'omics analysis of intestinal tissue in ankylosing spondylitis identifies alterations in the tryptophan metabolism pathway. Front Immunol 2021;12:587119.
12. Schroecksnadel K, Kaser S, Ledochowski M, et al. Increased degradation of tryptophan in blood of patients with rheumatoid arthritis. J Rheumatol 2003;30(9):1935–1939.
13. Choi SC, Brown J, Gong M, et al. Gut microbiota dysbiosis and altered tryptophan catabolism contribute to autoimmunity in lupus-susceptible mice. Sci Transl Med 2020;12(551):eaax2220.
14. Seymour BJ, Trent B, Allen BE, et al. Microbiota-dependent indole production stimulates the development of collagen-induced arthritis in mice. J Clin Invest 2023;134(4):e167671.
15. Moulin D, Millard M, Taïeb M, et al. Counteracting tryptophan metabolism alterations as a new therapeutic strategy for rheumatoid arthritis. Ann Rheum Dis 2024;83(3):312–323.
16. Lee JH, Lee J. Indole as an intercellular signal in microbial communities. FEMS Microbiol Rev 2010;34(4):426–444.
17. Wu L, Luo Y. Bacterial quorum-sensing systems and their role in intestinal bacteria-host crosstalk. Front Microbiol 2021;12:611413.
18. Smith EA, Macfarlane GT. Enumeration of human colonic bacteria producing phenolic and indolic compounds: effects of pH, carbohydrate availability and retention time on dissimilatory aromatic amino acid metabolism. J Appl Bacteriol 1996;81(3):288–302.
19. Wikoff WR, Anfora AT, Liu J, et al. Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. Proc Natl Acad Sci U S A 2009;106(10):3698–3703.
20. Williams BB, Van Benschoten AH, Cimermancic P, et al. Discovery and characterization of gut microbiota decarboxylases that can produce the neurotransmitter tryptamine. Cell Host Microbe 2014;16(4): 495–503.
21. Dodd D, Spitzer MH, Van Treuren W, et al. A gut bacterial pathway metabolizes aromatic amino acids into nine circulating metabolites. Nature 2017;551(7682):648–652.
22. Cervantes-Barragan L, Chai JN, Tianero MD, et al. *Lactobacillus reuteri* induces gut intraepithelial $CD4^+CD8\alpha\alpha^+$ T cells. Science 2017; 357(6353):806–810.
23. Wilck N, Matus MG, Kearney SM, et al. Salt-responsive gut commensal modulates T_H17 axis and disease. Nature 2017;551(7682): 585–589.
24. Russell WR, Duncan SH, Scobbie L, et al. Major phenylpropanoid-derived metabolites in the human gut can arise from microbial fermentation of protein. Mol Nutr Food Res 2013;57(3):523–535.
25. Aragozzini F, Ferrari A, Pacini N, et al. Indole-3-lactic acid as a tryptophan metabolite produced by Bifidobacterium spp. Appl Environ Microbiol 1979;38(3):544–546.
26. Shinde R, McGaha TL. The aryl hydrocarbon receptor: connecting immunity to the microenvironment. Trends Immunol 2018;39(12): 1005–1020.
27. Rothhammer V, Quintana FJ. The aryl hydrocarbon receptor: an environmental sensor integrating immune responses in health and disease. Nat Rev Immunol 2019;19(3):184–197.
28. Krishnan S, Ding Y, Saeidi N, et al. Gut microbiota-derived tryptophan metabolites modulate inflammatory response in hepatocytes and macrophages. Cell Rep 2019;28(12):3285.
29. Aletaha D, Neogi T, Silman A, et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Ann Rheum Dis 2010;69(9):1580–1588.
30. Maeda Y, Kurakawa T, Umemoto E, et al. Dysbiosis contributes to arthritis development via activation of autoreactive T cells in the intestine. Arthritis Rheumatol 2016;68(11):2646–2661.
31. Wang G, Fan Y, Zhang G, et al. Microbiota-derived indoles alleviate intestinal inflammation and modulate microbiome by microbial cross-feeding. Microbiome 2024;12(1):59.
32. Hendrikx T, Duan Y, Wang Y, et al. Bacteria engineered to produce IL-22 in intestine induce expression of REG3G to reduce ethanol-induced liver disease in mice. Gut 2019;68(8):1504–1515.
33. Modoux M, Rolhion N, Lefevre JH, et al. Butyrate acts through HDAC inhibition to enhance aryl hydrocarbon receptor activation by gut microbiota-derived ligands. Gut Microbes 2022;14(1):2105637.
34. Li YY, Wang XJ, Su YL, et al. Baicalein ameliorates ulcerative colitis by improving intestinal epithelial barrier via AhR/IL-22 pathway in ILC3s. *Acta Pharmacol Sin* 2022;43(6):1495–1507. blood of patients with rheumatoid arthritis. J Rheumatol 2003;30(9):1935–1939.
35. Yang TH, Hsu PY, Meng M, et al. Supplement of 5-hydroxytryptophan before induction suppresses inflammation and collagen-induced arthritis. Arthritis Res Ther 2015;17:364.
36. Rosser EC, Piper CJM, Matei DE, et al. Microbiota-derived metabolites suppress arthritis by amplifying aryl-hydrocarbon receptor activation in regulatory B cells. Cell Metab 2020;31(4):837–851.e10.
37. Nakoneczna I, Forbes JC, Rogers KS. The arthritogenic effect of indole, skatole and other tryptophan metabolites in rabbits. Am J Pathol 1969;57(3):523–538.
38. Tong YL, Zheng LL, Qing PY, et al. Oral microbiota perturbations are linked to high risk for rheumatoid arthritis. Front Cell Infect Microbiol 2020;9:475.
39. Brusca SB, Abramson SB, Scher JU. Microbiome and mucosal inflammation as extra-articular triggers for rheumatoid arthritis and autoimmunity. Curr Opin Rheumatol 2014;26(1):101–107.
40. Marietta EV, Murray JA, Luckey DH, et al. Suppression of inflammatory arthritis by human gut-derived Prevotella histicola in humanized mice. Arthritis Rheumatol 2016;68(12):2878–2888.
41. Ehrlich AM, Pacheco AR, Henrick BM, et al. Indole-3-lactic acid associated with Bifidobacterium-dominated microbiota significantly decreases inflammation in intestinal epithelial cells. BMC Microbiol 2020;20(1):357.
42. Meng D, Sommella E, Salviati E, et al. Indole-3-lactic acid, a metabolite of tryptophan, secreted by Bifidobacterium longum subspecies infantis is anti-inflammatory in the immature intestine. Pediatr Res 2020;88(2):209–217.

43. Huang W, Cho KY, Meng D, et al. The impact of indole-3-lactic acid on immature intestinal innate immunity and development: a transcriptomic analysis. Sci Rep 2021;11(1):8088.
44. de Aquino SG, Abdollahi-Roodsaz S, Koenders MI, et al. Periodontal pathogens directly promote autoimmune experimental arthritis by inducing a TLR2- and IL-1-driven Th17 response. J Immunol 2014; 192(9):4103–4111.
45. Bai Y, Li Y, Marion T, et al. Resistant starch intake alleviates collagen-induced arthritis in mice by modulating gut microbiota and promoting concomitant propionate production. J Autoimmun 2021;116:102564.
46. Gandhi R, Kumar D, Burns EJ, et al. Activation of the aryl hydrocarbon receptor induces human type 1 regulatory T cell-like and Foxp3(+) regulatory T cells. Nat Immunol 2010;11(9):846–853.
47. Tong Y, Niu M, Du Y, et al. Aryl hydrocarbon receptor suppresses the osteogenesis of mesenchymal stem cells in collagen-induced arthritic mice through the inhibition of β-catenin. Exp Cell Res 2017;350(2): 349–357.
48. Yuan X, Tong B, Dou Y, et al. Tetrandrine ameliorates collagen-induced arthritis in mice by restoring the balance between Th17 and Treg cells via the aryl hydrocarbon receptor. Biochem Pharmacol 2016;101:87–99.
49. Wei ZF, Lv Q, Xia Y, et al. Norisoboldine, an anti-arthritis alkaloid isolated from Radix Linderae, attenuates osteoclast differentiation and inflammatory bone erosion in an aryl hydrocarbon receptor-dependent manner. Int J Biol Sci 2015;11(9):1113–1126.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis & Rheumatology
Vol. 78, No. 3, March 2026, pp 630–643
DOI 10.1002/art.43393

ZBTB20 Regulating Postnatal Articular Cartilage Development and Homeostasis: Implications for Early-Onset Osteoarthritis

Xianhua Ma,[1] Fei Jiang,[2] Zhenbang Qin,[1] Yuqing Zhang,[3,4,5,6,7,8] Chunchun Wei,[1] Shuang Han,[9] Yuxia Chen,[1] Hai Zhang,[1] Jiang Tao,[3,4,5,6,7,8] Zhifang Xie,[2] and Weiping J. Zhang[1,9]

Objective. Mutations in *ZBTB20*, which encodes a transcription factor, are linked to epiphyseal dysplasia and articular degeneration in humans. This study investigates the role of ZBTB20 in regulating articular cartilage growth and integrity during postnatal development and its implications for early-onset osteoarthritis (OA) in mice.

Methods. We assessed the spatiotemporal expression of ZBTB20 in articular cartilage using immunostaining and generated an inducible cartilage-specific *Zbtb20* knockout mouse model. The impacts of *Zbtb20* deletion on cartilage thickness, zonal organization, cellular proliferation, and apoptosis were analyzed. We employed histology, microcomputed tomography, in situ hybridization, RNA-sequencing (RNA-seq), and cleavage under targets and tagmentation (CUT&Tag) to evaluate structural and molecular changes in knees from six to eight male mice per group. *ZBTB20* expression in human OA cartilage was analyzed using publicly available datasets.

Results. ZBTB20 was expressed in postnatal developing and adult articular chondrocytes. Postnatal *Zbtb20* deletion resulted in progressive thickening of articular cartilage in knees, with a 1.9–fold increase at two months of age, particularly in the deep and calcified zones. This was accompanied by chondrocyte overproliferation and differentiation defects, leading to early-onset cartilage degeneration by six months. RNA-seq and CUT&Tag analyses revealed that ZBTB20 directly regulates a broad set of genes essential for cartilage growth, chondrocyte differentiation, and extracellular matrix organization. Moreover, *ZBTB20* expression was significantly reduced in aging-related OA cartilage in both mice and humans, and inducible deletion of *Zbtb20* in adult cartilage resulted in severe spontaneous OA-like changes in mice.

Conclusion. ZBTB20 is essential for postnatal articular cartilage development and homeostasis, with a protective role in aging-related OA progression.

INTRODUCTION

Articular cartilage undergoes significant structural and functional changes during postnatal development. Disruptions to this process can predispose individuals to early-onset osteoarthritis (OA).[1,2] However, the transcriptional regulatory networks that govern postnatal articular cartilage growth and maturation are still poorly understood.

At birth, mouse knee articular cartilage is thin, consisting of randomly oriented small cells within a sparse extracellular matrix (ECM). From postnatal week two, with the development of the secondary ossification center (SOC), the articular cartilage

Supported by the National Science and Technology Major Project (grant 2025ZD0549800), National Natural Science Foundation of China (grants 32271162, U23A20171 and 91849110) and R&D Program of Guangzhou National Laboratory (grant GZNL2023A02012).

[1]National Key Laboratory of Immunity and Inflammation, Naval Medical University, Shanghai, China; [2]Ministry of Education-Shanghai Key Laboratory of Children's Environmental Health, Institute of Early Life Health, Xinhua Hospital Affiliated to Shanghai Jiaotong University School of Medicine, Shanghai, China; [3]Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China; [4]College of Stomatology, Shanghai Jiao Tong University, Shanghai, China; [5]National Center for Stomatology, Shanghai, China; [6]National Clinical Research Center for Oral Diseases, Shanghai, China; [7]Shanghai Key Laboratory of Stomatology, Shanghai, China; [8]Shanghai Research Institute of Stomatology, Shanghai, China; [9]Chu Hsien-I Memorial Hospital & Tianjin Institute of Endocrinology, Tianjin Medical University, Tianjin, China.

Drs Ma and Jiang are co-first authors and contributed equally to this work. Drs Xie and W. J. Zhang are co-last authors and contributed equally to this work.

Additional supplementary information cited in this article can be found online in the Supporting Information section (http://onlinelibrary.wiley.com/doi/10.1002/art.43393).

Author disclosures and graphical abstract are available at https://onlinelibrary.wiley.com/doi/10.1002/art.43393.

Address correspondence via email to Zhifang Xie, MD, PhD, at zfxie70@hotmail.com or xiezhifang@xinhuamed.com.cn; or to Weiping J. Zhang, MD, PhD, at wzhang@smmu.edu.cn.

Submitted for publication May 23, 2025; accepted in revised form September 5, 2025.

becomes distinct from the growth plate.[3] By postnatal day 28 (P28), tibial articular cartilage thickness in mice reaches its peak and then decreases by two months, primarily due to changes in cellular proliferation, hypertrophy, and ECM accumulation.[4]

During this time, chondrocyte proliferation significantly declines, and the tissue acquires a zonal structure composed of superficial, transitional, deep, and calcified zones.[2,4] Concurrently, the ECM becomes more enriched and organized.[1,2] The mechanisms underlying the deceleration in chondrocyte proliferation and the development of properly thickened, organized mature tissue remain largely unclear.

ZBTB20 is a member of the BTB/POZ family of transcription factors, which is capable of acting both as a transcriptional repressor and activator. It plays a critical role in the development of various organs, including bones.[5–10] We previously demonstrated abundant expression of ZBTB20 in hypertrophic chondrocytes of the growth plate during late embryonic stages and in resting chondrocytes shortly after birth. Deletion of *Zbtb20* in chondrocytes delays endochondral ossification and impairs bone growth.[9] In humans, mutations in *ZBTB20* are associated with Primrose syndrome, which presents with skeletal abnormalities including epiphyseal dysplasia and articular degeneration,[11–14] suggesting a role for ZBTB20 in joint development and homeostasis.

In this study, we demonstrate that ZBTB20 regulates a broad set of genes essential for articular cartilage growth, chondrocyte differentiation, and ECM organization during postnatal life. Inducible deletion of *Zbtb20* in immature cartilage results in thickening and early-onset degeneration of articular cartilage.

MATERIALS AND METHODS

Animals. $Zbtb20^{Flox}$ mice have been constructed and backcrossed to C57BL/6J mice for at least five generations.[5,9] $Zbtb20^{Floxed/Agc1\text{-}CreER}$ mice were generated by crossing $Zbtb20^{Flox/Flox}$ mice with *Aggrecan (Agc1)-CreERT2* knockin mice (Jackson Laboratory, #019148).[15] To induce Cre-mediated inactivation of *Zbtb20*, $Zbtb20^{Floxed/Agc1\text{-}CreER}$ male mice and their sex-matched littermate wild-type (WT) control mice (Cre-negative or $Zbtb20^{+/+}$) received tamoxifen (TAM) (50 mg/kg body weight) by oral gavage for either 8 consecutive days starting at P10 or 10 consecutive days starting at 2.5 months. WT C57BL/6J mice, aged three months and two years, were obtained from GemPharmatech Co., Ltd. Knees from around six to eight mice per group were analyzed for each stage. In total, 153 animals were used. All animal experiments were conducted according to protocols (XIEC-NSFC-2021–097) approved by the Ethics Committee of Xinhua Hospital affiliated with Shanghai Jiaotong University School of Medicine.

Protocols. Protocols for immunofluorescent staining, in situ hybridization, histologic analyses, OA grading, microcomputed tomography analysis, 5-ethynyl-2'-deoxyuridine labeling, and terminal deoxynucleotidyl transferase dUTP nick-end labeling (TUNEL) assay are described in the Supplementary Methods and Supplementary Tables 1–2. Detailed methods for RNA-sequencing (RNA-seq), real-time quantitative reverse transcription polymerase chain reaction, and cleavage under targets and tagmentation analysis (Cut&Tag) are also available in the Supplementary Methods and Supplementary Table 3. The RNA-seq and Cut&Taq data are available in Gene Expression Omnibus (GSE246737) and China National Center for Bioinformation/Beijing Instiute of Genomics, Chinese Academy of Sciences (CRA02907), respectively.

Statistical analysis. No randomization methods were used to allocate samples. Sample sizes were empirically determined and are indicated in figure legends as *n*. No data were excluded. This was an exploratory study with no prespecified primary or secondary endpoints. Statistical analysis was performed using GraphPad Prism 7.0. Data with a sample size of at least seven per group were tested for Gaussian distribution. Normally distributed data were compared using an unpaired Student's *t*-test and are presented as dot plots with mean ± SD. Articular thickness was analyzed by two-way analysis of variance with Tukey's multiple comparisons test. Data with a sample size fewer than seven per group or non-normal distribution were compared using the Mann–Whitney test and are presented as dot plots with medians and interquartile ranges (IQR). A two-tailed $P < 0.05$ was considered significant. *P* values are shown in the graphs, with values <0.001 reported as $P < 0.001$. Summary statistics are provided in Supplementary Tables 4–20.

RESULTS

ZBTB20 is expressed by postnatal developing and adult articular chondrocytes. Our previous study demonstrated that ZBTB20 is specifically expressed in hypertrophic chondrocytes of the prenatal growth plate and later in the reserve zone chondrocytes shortly after birth.[9] In this study, we investigate the spatial and temporal expression patterns of ZBTB20 in articular cartilage. Immunostaining revealed no detectable ZBTB20 signals in knee articular chondrocytes at P4 (Supplementary Figure 1). However, by P10, ZBTB20 expression became evident (Figure 1A). At P21, when articular cartilage is organized into superficial, transitional, deep, and calcified zones, ZBTB20 was detected in DAPI-stained nuclei of chondrocytes across all zones, with the strongest in the deep and calcified zones, moderate in the transitional zone, and very weak in the superficial layer (Figure 1A). This expression pattern remained consistent at two months of age (Supplementary Figure 1), corresponding to the full maturation of articular cartilage. Thus,

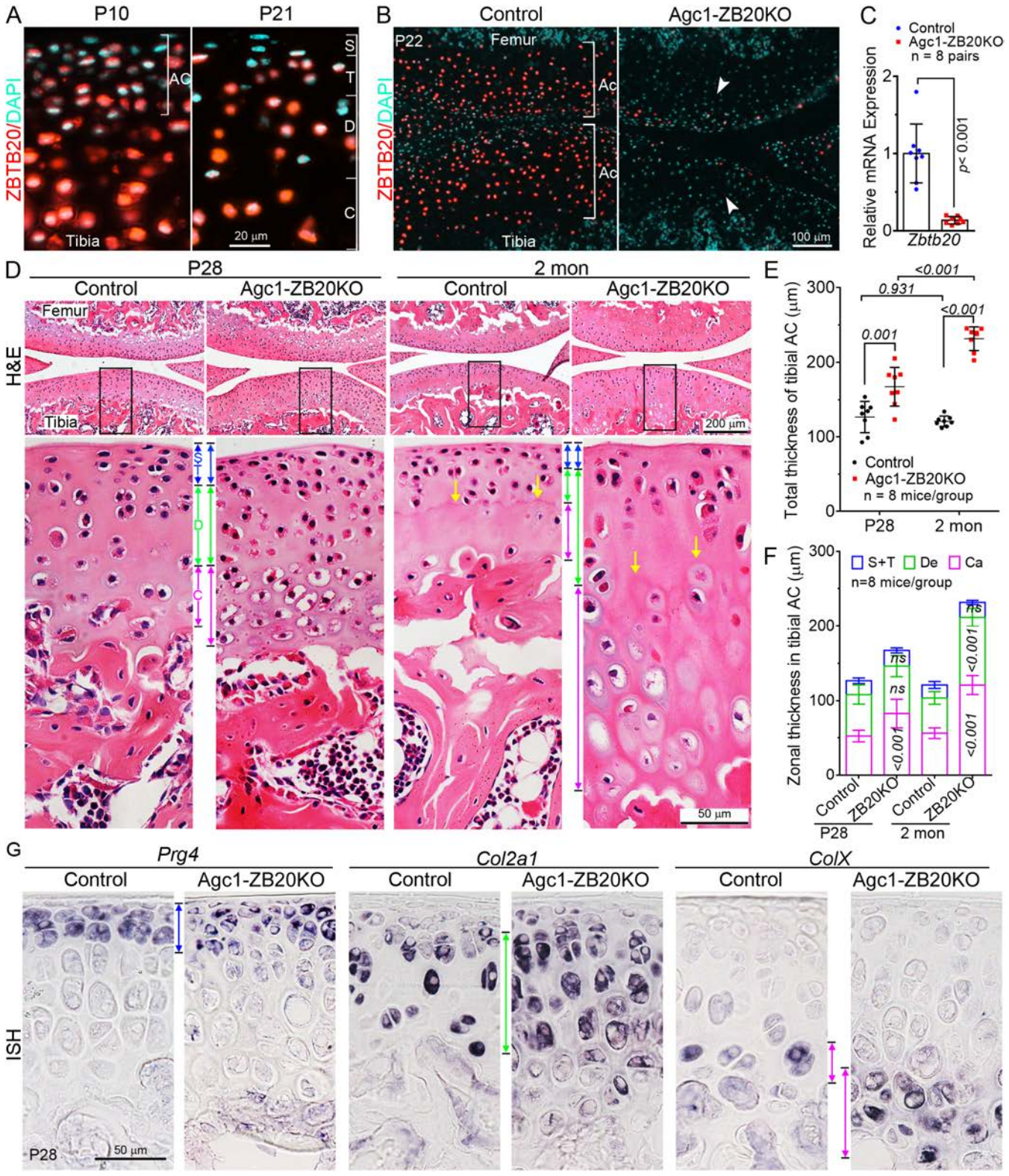


Figure 1. Legend on next page.

ZBTB20 is expressed in both postnatal developing and adult articular chondrocytes.

Postnatal deletion of *Zbtb20* leads to progressive thickening of the articular cartilage during juvenile-adolescent development. To reveal the role of ZBTB20 in articular cartilage, we developed an inducible cartilage-selective knockout mouse model (*Zbtb20*$^{Floxed/Agc1\text{-}CreER}$ mice, referred to as Agc1-ZB20KO) by crossing *Zbtb20*$^{Flox/Flox}$ mice[5] with *Aggrecan (Agc1)-CreERT2* knockin mice.[15] Cre-mediated inactivation of *Zbtb20* was induced by administering TAM for eight consecutive days starting at P10, coinciding with the initial detection of ZBTB20 in articular cartilage and following the initiation of SOC.[16] Immunostaining confirmed loss of ZBTB20 signals in over 90% of articular chondrocytes in Agc1-ZB20KO knee joints, whereas robust expression was observed in WT controls (Figure 1B). Chondrocytes in the mutant growth plate and meniscus also lost ZBTB20, whereas expression remained unchanged in intra-joint ligaments and articular capsule including synovial lining cells (Supplementary Figure 2), which is consistent with the previous report.[15] Real-time quantitative reverse transcription polymerase chain reaction (qRT-PCR) analysis showed an 86.5% reduction in *Zbtb20* messenger RNA (mRNA) in the articular cartilage of Agc1-ZB20KO femoral heads compared with WT controls (Figure 1C), confirming the efficiency and specificity of *Zbtb20* deletion in Agc1-ZB20KO chondrocytes.

We next assessed the temporal histologic changes in articular cartilage following *Zbtb20* deletion. Histologic morphometry of hematoxylin and eosin–stained knee joints revealed a 1.3–fold increase in overall articular cartilage thickness in the Agc1-ZB20KO proximal tibia compared with WT controls at P28 (Figure 1D and E), a time when articular cartilage reaches its maximal thickness and exhibits its zonal structure.[4] By two months of age, when articular cartilage normally matures, the thickness of the mutant articular cartilage had increased even further, reaching 1.9 times that of the control group (Figure 1D and E). Thus, the loss of ZBTB20 leads to progressive thickening of the articular cartilage during juvenile-adolescent development.

Zonal analysis revealed that at P28, only the calcified zone was significantly thicker in Agc1-ZB20KO cartilage compared with controls (Figure 1D and F). To further confirm this finding, we examined the expression patterns of zonal molecular markers at P28, including *Prg4* (a marker for the superficial and transitional zones),[17] *Col2a1* (a marker more abundantly expressed by deep layer chondrocytes),[18] and *ColX* (a marker for calcified layer chondrocytes).[19] In situ hybridization (ISH) revealed a significant expansion of *ColX* mRNA expression in the mutant proximal tibia compared with WT controls, whereas the expression domains of *Prg4* and *Col2a1* remained similar between the two groups (Figure 1G), which is consistent with the histologic observations.

By two months of age, both the deep and calcified zones showed significant increases in thickness in the mutant group, whereas the thicknesses of the superficial and transitional layers remained similar between the two groups at both time points (Figure 1D and F). Furthermore, the tidemark appeared less distinct and discontinuous in six of eight mutant samples, compared with the clear and continuous tidemark observed in all WT controls (Figure 1D). Additionally, Agc1-ZB20KO mice exhibited a decreased SOC area and delayed ossification of the growth plate at P28 or two months (Supplementary Figure 3), which is consistent with our previous report.[9] Collectively, these findings indicate that ZBTB20 negatively regulates the growth of articular cartilage thickness, particularly the deep and calcified zones, during juvenile-adolescent development.

Postnatal deletion of *Zbtb20* in articular cartilage leads to early-onset spontaneous cartilage damage. The articular defects observed in young Agc1-ZB20KO mice prompted us to investigate whether these defects persist and lead to articular cartilage degeneration in adults. To examine this, we conducted a histologic analysis of knee joints from both genotypic groups at six months of age, focusing on male mice to control for potential TAM-related effects. Safranin O/Fast Green-stained semiserial coronal sections revealed a 1.9–fold increase in overall articular cartilage thickness in Agc1-ZB20KO proximal tibia compared with WT controls (Figure 2A and B), confirming that articular cartilage thickening defects persist into adulthood in the mutant.

Indeed, Agc1-ZB20KO mice displayed significant early-onset articular cartilage degeneration at six months, particularly in the medial tibial plateau. This degeneration was marked by proteoglycan loss, as evidenced by reduced Safranin O staining that

Figure 1. Deletion of *Zbtb20* leads to progressively increased articular cartilage thickness: Immunofluorescence staining of Zbtb20 (red) counterstained with DAPI (turquoise) in tibial articular (Ac) and meniscus (Me) chondrocytes from WT (A) or Agc1-ZB20KO mice at indicated ages (B). n = 3/stage. (C) Real-time qRT-PCR showing significantly down-regulated *Zbtb20* mRNA in Agc1-ZB20KO articular cartilage at P22. n = 8/group. (D) Representative hematoxylin and eosin–stained knee sections showing increased thickness of articular cartilages from Agc1-ZB20KO mice at the indicated ages. The low panels are magnified views of respective boxed regions in the upper panels. Yellow arrows indicate tidemarks. n = 8/group for each stage. Bar graph showing the averaged thickness of overall tissue (E) or each zone (F) from the two groups at indicated ages. (G) ISH showing expression domains of *Prg4*, *Col2a1*, and *ColX* mRNA in articular cartilage at P28. n = 3/group. All values are presented as mean ± SD. Scale bar: 20 μm for (A), 100 μm for (B), 200 μm for (D _up panels), 50 μm for (D _low panels and G). See also Supplementary Figures 1–3 and Supplementary Tables 4–7. Ca, calcified zones; De, deep; ns, not significant; ISH, in situ hybridization; P28, postnatal day 28; mRNA, messenger RNA; qRT-PCR, quantitative reverse transcription polymerase chain reaction; S/T, superficial/transitional; WT, wild type.

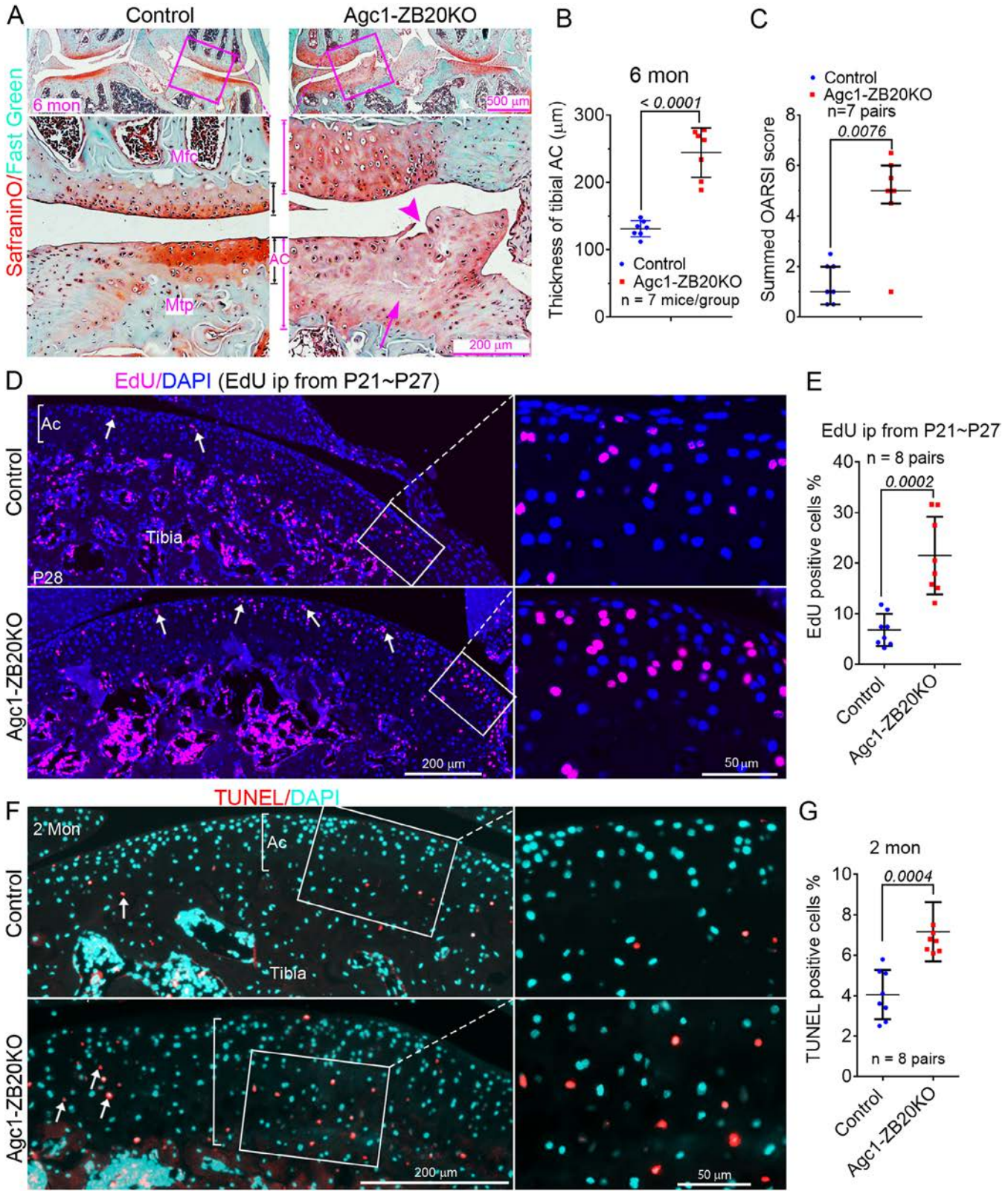


Figure 2. Legend on next page.

extended to more than 75% of the articular surface, and additional features included surface fibrillations, vertical clefts that extended into the deep or calcified layers, chondrocyte depletion and horizontal clefts within the deep zone, and cartilage overgrowth that was observed at the inner margin (Figure 2A). Accordingly, the mutant mice displayed significantly higher summed Osteoarthritis Research Society International (OARSI) scores compared with WT controls (Figure 2C). No notable histologic changes were observed in the intra-joint ligaments, articular capsule, or meniscus of six-month-old Agc1-ZB20KO mice (Figure 2A). These findings underscore the critical role of ZBTB20 in proper joint development, with its disruption resulting in joint instability that contributes to the early onset of cartilage degeneration.

Deletion of *Zbtb20* leads to increased chondrocyte proliferation and cell death in postnatal articular cartilage. The articular cartilage thickening defects observed in Agc1-ZB20KO mice suggest potential dysregulation in chondrocyte proliferation and/or cell death. To investigate chondrocyte proliferation, we employed a seven-day consecutive 5-ethynyl-2'-deoxyuridine (EdU) labeling strategy (twice-daily injections from P21 to P27) because articular chondrocytes show minimal proliferation after the second postnatal week.[4] The mutant mice displayed a significant increase in EdU-positive chondrocytes within the tibial articular cartilage compared with controls. These proliferating cells were primarily localized in the mutant transitional zone, with a higher abundance in the peripheral regions (Figure 2D–E), indicating excessive proliferation resulting from *Zbtb20* deletion.

Next, we assessed cell death in the articular cartilage of knee joints at P28 using the TUNEL assay. Few apoptotic cells were detected at this stage, with no significant difference between the mutant and control mice (Supplementary Figure 4). However, by two months of age, mutant tibial articular cartilage exhibited a significant increase in TUNEL-positive cells compared with WT controls (Figure 2F–G). These apoptotic cells were predominantly localized in the deep and calcified layers. These findings suggest early-onset chondrocyte degeneration following *Zbtb20* deletion, which is in line with the articular cartilage damage observed in six-month-old Agc1-ZB20KO knees.

Postnatal deletion of *Zbtb20* disrupts ECM homeostasis and articular chondrocyte differentiation. To gain insights into the molecular mechanisms of ZBTB20 action and identify potential downstream targets, we performed RNA-seq on articular cartilage from Agc1-ZB20KO and WT control femoral heads at P22. We conducted six biologic replicates per group and identified 195 significantly differentially expressed genes (DEGs) with a fold change of ≥2 and Q value ≤0.05, filtering out genes with fragments per kilobase per million mapped reads (FPKM) <2. Among these, 109 genes were significantly up-regulated, whereas 86 genes were down-regulated (Figure 3A, Supplementary File 1). We validated 11 selected DEGs using real-time qRT-PCR (Supplementary Figure 5A and B) based on their high fold changes or potential relevance to articular cartilage biology.

We then performed gene ontology (GO) analysis to analyze the pathways and biologic functions of the DEGs. Figure 3B shows the top 10 enriched biologic processes, with the ECM organization term being the most significant by Q value. ECM-related terms are also mostly enriched in the cellular component category (Supplementary Figure 5C). Within the 29 ECM-associated DEGs, 15 were up-regulated, and 14 were down-regulated (Figure 3C). It is well established that matrix metalloproteinases (MMPs) and their inhibitors (TIMPs) are pivotal regulators of ECM composition and turnover.[20,21] In Agc1-ZB20KO articular cartilage, *Mmp-2* and *Mmp-16* were up-regulated, whereas *Mmp-3*, *Timp3*, and *Timp4* were significantly down-regulated, suggesting an imbalanced expression of MMPs and TIMPs. We also noticed *Lox* (lysyl oxidase), encoding a key enzyme catalyzing the covalent cross-linking of ECM proteins,[22] was significantly down-regulated in Agc1-ZB20KO articular cartilage (Figure 3C).

We also observed dysregulated expression of several abundant noncollagenous ECM components, including the downregulation of *Ibsp* (integrin-binding sialoprotein), a marker for hypertrophic chondrocytes,[23] and the upregulation of *Epyc* (epiphycan) and *Tnc* (tenascin C) (Figure 3C), which are two

Figure 2. Deletion of *Zbtb20* leads to early-onset degeneration, increased chondrocyte proliferation, and cell death in articular cartilage. Safranin O/Fast Green-stained coronal knee sections of WT control and Agc1-ZB20KO mice at six months (A). The lower panels show magnified views of boxed regions in the Mtp articular cartilage. The arrowhead indicates a vertical cleft, whereas the arrow points to a chondrocyte-depleted area in Agc1-ZB20KO cartilage. Bar graphs show significantly increased tibial articular cartilage thickness (Ac) (B) and summed OARSI scores (C) in Agc1-ZB20KO knees (n = 7 mice/group). Increased EdU^+ (magenta) chondrocytes, counterstained with DAPI (blue), in Agc1-ZB20KO tibial articular cartilage at P28. The right panels are magnified views of boxed regions. EdU was administered twice daily from P21 to P27 (n = 8/group). (E) shows histomorphometry analysis of (D). Increased $TUNEL^+$ cells (red, arrows) counterstained with DAPI (turquoise) in Agc1-ZB20KO cartilage at two months (n = 8/group). (G) shows histomorphometry analysis of (F). Values are mean ± SD ([B], [E], [G]) or medians with 25–75% IQR ([C]). Scale bars are shown in each panel. See also Supplementary Figure 4 and Supplementary Tables 8–11. AC, articular; EdU, 5-ethynyl-2'-deoxyuridine; IQR, interquartile range; Mfc, medial femoral condyle; Mtp, medial tibial; OARSI, Osteoarthritis Research Society International; P28, postnatal day 28; TUNEL, terminal deoxynucleotidyl transferase dUTP nick-end labeling; WT, wild-type. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.43393/abstract.

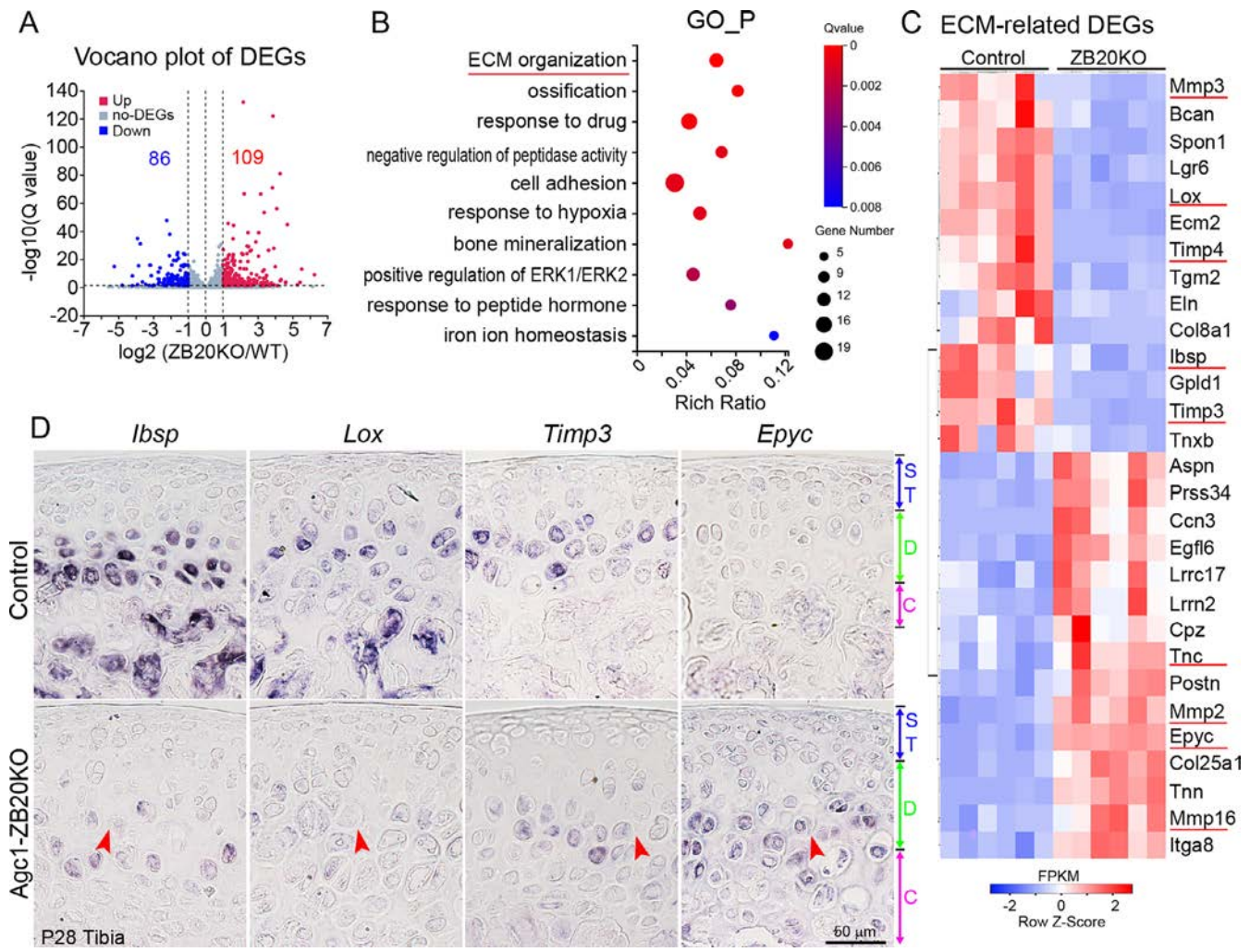


Figure 3. Postnatal deletion of *Zbtb20* disrupts ECM homeostasis and articular chondrocyte differentiation. (A) Volcano plot showing DEGs in Agc1-ZB20KO articular cartilage at P22, as detected by RNA-seq analysis (n = 6 pairs, fold change ≥2, Q value ≤0.05). (B) GO analysis of DEGs revealing the top 10 overrepresented biologic process terms. (C) Heatmap of ECM-related DEGs in Agc1-ZB20KO (KO) vs WT control articular cartilage. (D) ISH analysis confirmed dysregulation of ECM-related DEGs in the mutant tibial articular cartilage at P28 (indicated by red arrowheads). Zones shown: S/T, D, and C. n = 3 pairs. Scale bar: 50 μm. See also Supplementary Figure 5 and Supplementary Table 12–13. C, calcified; DEGs, differentially expressed genes; D, deep; ECM, extracellular matrix; FPKM, fragments per kilobase per million mapped reads; GO, gene ontology; ISH, in situ hybridization; P28, postnatal day 28; RNA-seq, RNA-sequencing; S/T, superficial/transitional; WT, wild type. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.43393/abstract.

proteoglycans typically expressed by neonatal articular chondrocytes but significantly down-regulated in juvenile articular chondrocytes.[24,25] The dysregulation of these genes suggests disruptions in ECM architecture and remodeling in *Zbtb20*-deficient articular cartilage.

The observed alterations in the mRNA expression of ECM-related genes also suggest differential defects in articular chondrocytes following ZBTB20 loss. We therefore performed ISH to examine the spatial expression patterns of several abundant DEGs and determine which zone might be most affected at P22. We found that *Ibsp*, *Lox*, and *Timp3* mRNA were predominantly expressed in the deep and calcified zones of WT control articular cartilage but were significantly down-regulated in the *Zbtb20*-deleted counterparts. In contrast, robust expression of *Epyc* mRNA was ectopically detected in the deep and calcified zones of *Zbtb20*-deleted articular cartilage (Figure 3D). These data suggest that the loss of ZBTB20 leads to differential defects in chondrocytes within the deep and calcified zones, which may contribute to the progressive thickening of these zones observed in *Zbtb20*-deficient articular cartilage.

Postnatal deletion of *Zbtb20* upregulates an imprinted gene network in juvenile-adolescent articular cartilage. To explore potential regulatory networks, we performed a protein-protein interaction (PPI) analysis on DEGs using the Search Tool for the Retrieval of Interacting Genes/Proteins database (https://string-db.org/). To increase the confidence of the interactions, we filtered out DEGs with low

expression levels (FPKM < 10) and set the minimum interaction score to 0.4. This analysis identified 41 nodes that were clustered into 12 distinct groups using the Markov cluster algorithm tool (Figure 4A). Among these, five small clusters were associated with ECM, which is consistent with the aforementioned enriched GO terms. The largest cluster contained 10 genes, with 4 up-regulated (*Jun*, *Foxo4*, *Cdkn1a*, and *Esr1*) and 6 down-regulated (*Dkk1*, *Mycl*, *Lox*, *Gpx3*, *Tgm2*, and *Stc2*) (Figure 4A and B).

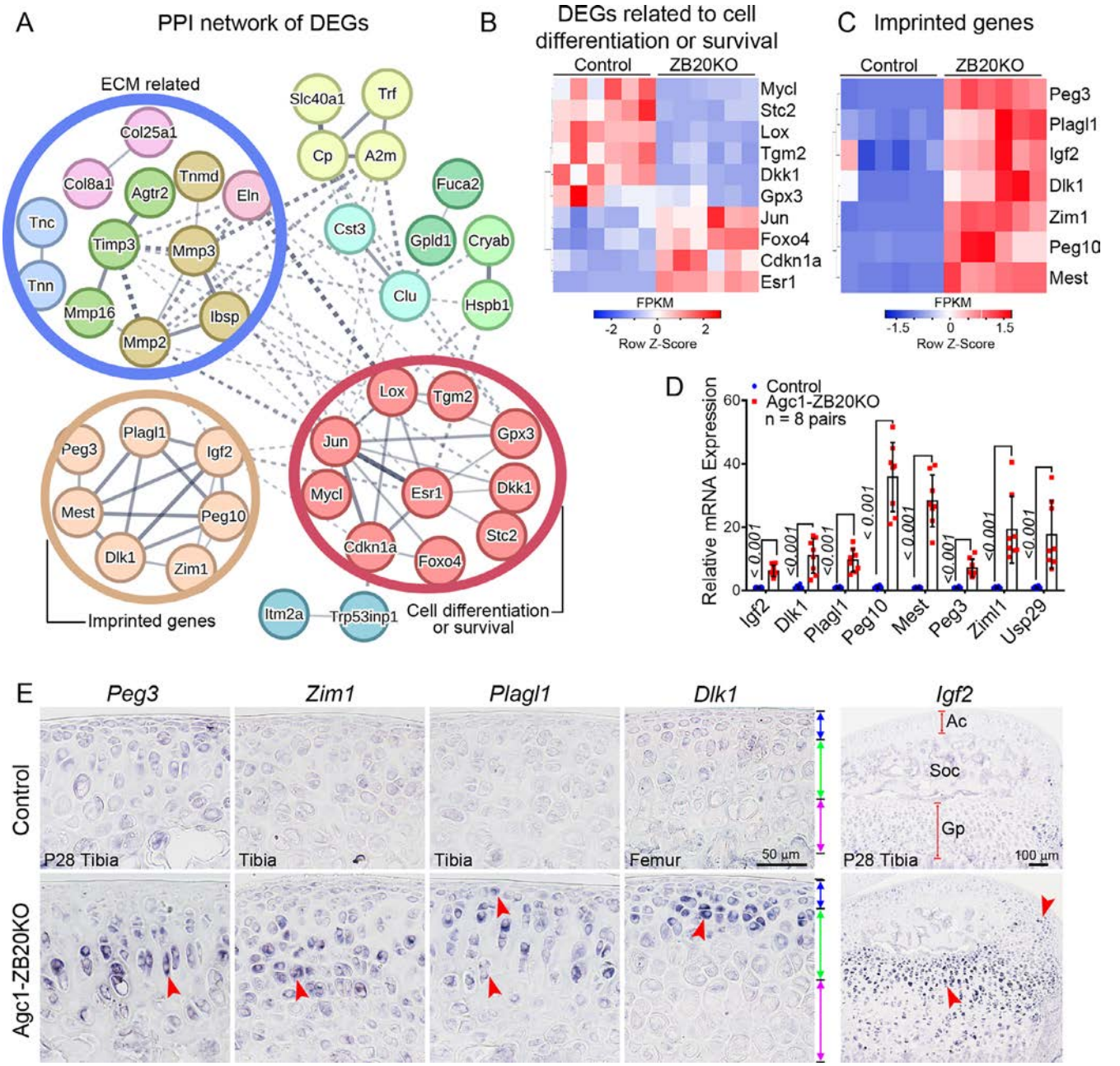


Figure 4. Postnatal deletion of *Zbtb20* upregulates an imprinted gene network in juvenile-adolescent articular cartilage. (A) PPI analysis identified protein interaction networks and functional clusters, including an imprinted gene network, among DEGs in Agc1-ZB20KO articular cartilage at P22. Heatmaps showing the largest (B) and second-largest (C) DEG clusters identified by PPI analysis. (D) Real-time qRT-PCR validation of up-regulated imprinted genes in articular cartilage at P28. n = 8 pairs. Data are presented as individual dot plots with mean ± SD. (E) ISH analysis confirmed the upregulation of imprinted genes in mutant cartilage (indicated by red arrows) at P28. The blue, green, and magenta double-headed arrows denote the approximate thickness of the superficial/transitional, deep, and calcified zones, respectively. n = 3 pairs. Scale bar: 50 μm (left panel), 100 μm (right panel). See also Supplementary Figure 6 and Supplementary Table 14. Ac, articular cartilage; DEGs, differentially expressed genes; FPKM, fragments per kilobase per million mapped reads; Gp, growth plate; ISH, in situ hybridization; mRNA, messenger RNA; P22, postnatal day 22; PPI, protein-protein interaction; qRT-PCR, quantitative reverse transcription polymerase chain reaction; SOC, second ossification center. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.43393/abstract.

These genes are primarily involved in chondrocyte differentiation or survival.[26–30]

The second-largest cluster consisted of seven imprinted genes, all of which were up-regulated in Agc1-ZB20KO chondrocytes. These genes—*Peg3*, *Zim1*, *Plagl1/Zac1*, *Dlk1*, *Mest*, *Igf2*, and *Peg10* (Figure 4A and C)—are known to form a coregulated imprinted gene network that regulates embryonic and postnatal growth.[31,32] Downregulation of this network is linked to growth deceleration in various tissues, including cartilage.[31–33] Real-time qRT-PCR confirmed the upregulation of these imprinted genes in Agc1-ZB20KO articular cartilage at P28 (Figure 4D).

We further used ISH to assess the spatial distribution of these genes in knee joints at P28. Although *Peg10* was not tested because of low expression levels, and *Mest* failed to yield consistent staining, the other imprinted genes showed significantly stronger signals in the mutant articular cartilage compared with controls (Figure 4E). Specifically, *Peg3* and *Zim1* were predominantly expressed in the mutant deep zone, *Plagl1* was broadly distributed across the mutant transitional and deep zones, and *Dlk1* was found in the mutant transitional and upper deep zones of the femora and to a lesser extent in the tibiae (Supplementary Figure 6). Increased *Igf2* expression was detected in the peripheral articular cartilage and in the growth plates of the mutants, and higher signals for *Peg3*, *Zim1*, and *Plagl1* were also observed (Supplementary Figure 6). Together, PPI analysis shows that ZBTB20 regulates a large set of genes associated with chondrocyte differentiation or survival and negatively controls the growth-promoting imprinted gene network in juvenile-adolescent cartilage.

ZBTB20 directly regulates a significant subset of genes critical for cartilage development and ECM homeostasis. To identify potential direct targets of ZBTB20, we performed CUT&Tag analysis[34] to map its genome-wide binding sites in articular cartilage. To obtain a sufficient quantity of primary chondrocytes, femoral heads from rats (aged at P22–P25) were used instead of mice. Two biologic replicates identified a total of 16,359 common ZBTB20-binding peaks after filtering out regions overlapping with IgG-negative controls using the SEARC (sparse enrichment analysis for cleavage under targets and release using nuclease) program (Supplementary File 2).[35] The binding peaks were predominantly located in intergenic regions (39.47%) and promoter regions (38.04%, defined as ±3 kb surrounding transcription start sites [TSSs]), followed by intronic regions (15.63%) (Figure 5A), suggesting ZBTB20 regulates gene transcription through either proximal and/or distal regulatory elements.

To further characterize ZBTB20 binding features, we conducted parallel CUT&Tag analyses for two histone modification markers—H3K27AC and H3K4ME1—using the same chondrocyte samples. These histone modifications are known to colocalize at enhancer regions and are recognized as chromatin signatures for enhancer identification.[36–38] Additionally, high levels of H3K27AC signals are detected at promoters of active genes, whereas H3K4ME1 marks both active and poised promoters.[39] Our integrative analysis revealed that 34.6% (5,665 of 16,359) of ZBTB20 peaks overlapped with both H3K27AC and H3K4ME1 peaks, marking enhancer regions, whereas 37.3% (6,094 of 16,359) colocalized with H3K27AC peaks alone, indicating active promoters. A smaller proportion of 5.5% (900 of 16,359) overlapped with H3K4ME1 peaks only, suggesting poised regulatory elements (Figure 5B). Further analysis of read distribution around the TSS showed that ZBTB20 and H3K27AC signals were centered on the TSS, whereas H3K4ME1 signals extended more broadly downstream (Figure 5C). These results indicate that ZBTB20 is enriched at both enhancer and promoter regions, implicating its role in regulating gene expression across multiple regulatory landscapes.

To link binding events to transcriptional changes, we integrated CUT&Tag data with the aforementioned DEGs identified by RNA-seq. Of the 195 DEGs, 43% (84 of 195) were found to have ZBTB20-binding sites (Figure 5D, Supplementary File 3), suggesting that ZBTB20 directly regulates a significant subset of DEGs. Among the up-regulated DEGs with ZBTB20-binding sites, five growth-promoting imprinted genes (*Zim1*, *Dlk1*, *Peg3*, *Igf2*, and *Plagl1*) were identified, along with other genes involved in chondrocyte differentiation, such as *Cdkn1a*, *Jun*, and *Tnc* (Figure 5D and E, Supplementary File 3). The down-regulated DEGs with ZBTB20-binding sites included several ECM-related genes, such as *Mmp3*, *Ibsp*, *Timp3*, and *Clu* (Figure 5D and E). These findings suggest that ZBTB20 directly regulates a broad set of genes critical for cartilage growth, chondrocyte differentiation, and ECM homeostasis.

Inducible deletion of *Zbtb20* in mature cartilage causes severe spontaneous OA with excessive osteophytes. Because ZBTB20 is also expressed in adult articular cartilage, we next asked if ZBTB20 is necessary for maintaining articular cartilage homeostasis in adulthood. To test this, we induced *Zbtb20* inactivation in Agc1-ZB20KO male mice through a 10-day consecutive TAM administration starting at two-and-a-half months of age. The specificity and efficiency of the *Cre*-mediated recombination were validated through immunostaining of ZBTB20 (Supplementary Figure 7).

To assess the joint health of Agc1-ZB20KO mice, we conducted microcomputed tomography (micro-CT) analysis followed by histologic analysis on their knees at one year of age, which is a stage considered to be middle-aged for mice. The micro-CT revealed extensive ectopic ossification and osteophytes in the mutant knee joints (Figure 6A–E). Consistently, histologic analysis revealed that the mutant mice exhibited extensive ectopic chondrogenesis and ossification in the synovium-capsule and cruciate ligaments, and large chondro-osteophytes in the outer or inner joint margins. This was accompanied by severe destruction of

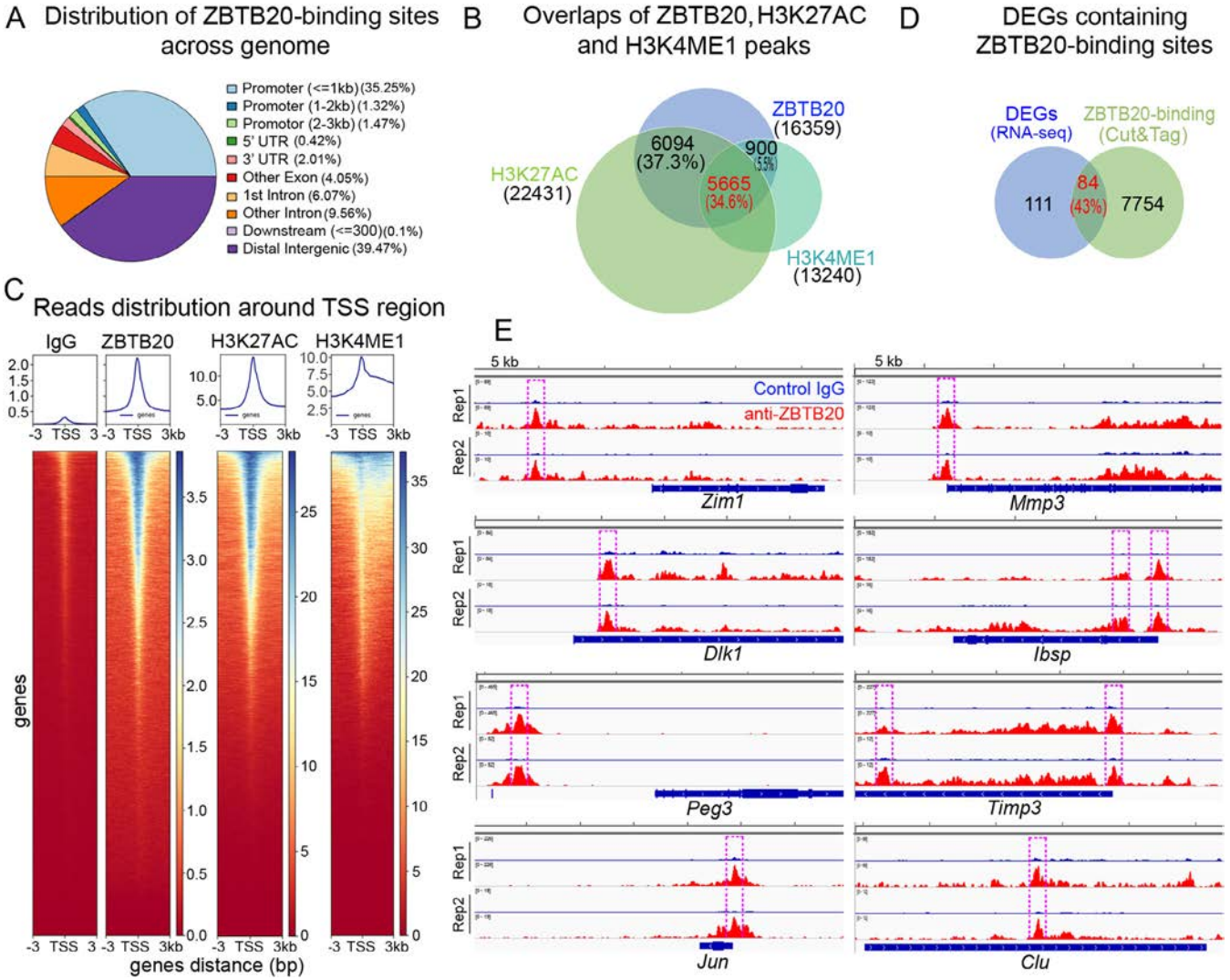


Figure 5. ZBTB20 directly regulates a significant subset of genes critical for cartilage development and ECM homeostasis. (A) Pie chart showing the distribution of ZBTB20-binding peaks across the genome, identified by CUT&Tag analysis in articular chondrocytes from rat femoral heads around P25 (two biologic replicates). (B) Venn diagram depicting common binding peaks among ZBTB20, H3K27AC, and H3K4ME1. (C) Heatmap of reads distribution of ZBTB20, H3K27AC, and H3K4ME1 within the –3 to +3 kb region flanking the TSS. (D) Venn diagram showing DEGs with ZBTB20-binding sites. (E) Integrative Genomics Viewer tracks showing ZBTB20-binding peaks (highlighted by magenta boxes) in the cis-regulatory regions of representative DEGs. Tracks for Control IgG (blue); for anti-ZBTB20 IgG (red). Source data are provided in Supplementary Datasets 2 and 3. CUT&Tag, cleavage under targets and tagmentation; DEGs, differentially expressed genes; ECM, extracellular matrix; P25, postnatal day 25; TSS, transcription start site. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.43393/abstract.

articular cartilage and synovial inflammation in the knee joints. Subchondral bone changes, including sclerosis and cysts, were occasionally observed along with severe cartilage lesions in some mutant joints. In contrast, WT control knee joints showed only mild cartilage damage (Figure 6F–G). Consequently, the mutant mice demonstrated significantly higher OARSI scores and grades of chondro-osteophyte formation compared with the control group (Figure 6H–J).

To investigate whether ZBTB20 expression alters during age-related cartilage degeneration, we examined histologic changes and ZBTB20 levels in knee articular cartilage from WT C57BL/6J mice at three months and two years. Histologic analysis and OARSI scoring revealed mild OA in two-year-old mice (Supplementary Figure 8). Immunostaining showed a 43% reduction in ZBTB20 immunofluorescence intensity in the articular cartilage of aged mice compared with younger controls (Figure 6K–L), supporting the link between ZBTB20 downregulation and age-related OA.

To explore similar changes in human OA, we analyzed publicly available datasets, including bulk RNA-seq dataset GSE168505 and two independent single-cell RNA-seq (scRNA-seq) datasets, GSE169454 and GSE255460.[40,41] GSE168505 and GSE169454 contained expression profiles of femoral condyle cartilage from three non-OA controls and four patients with

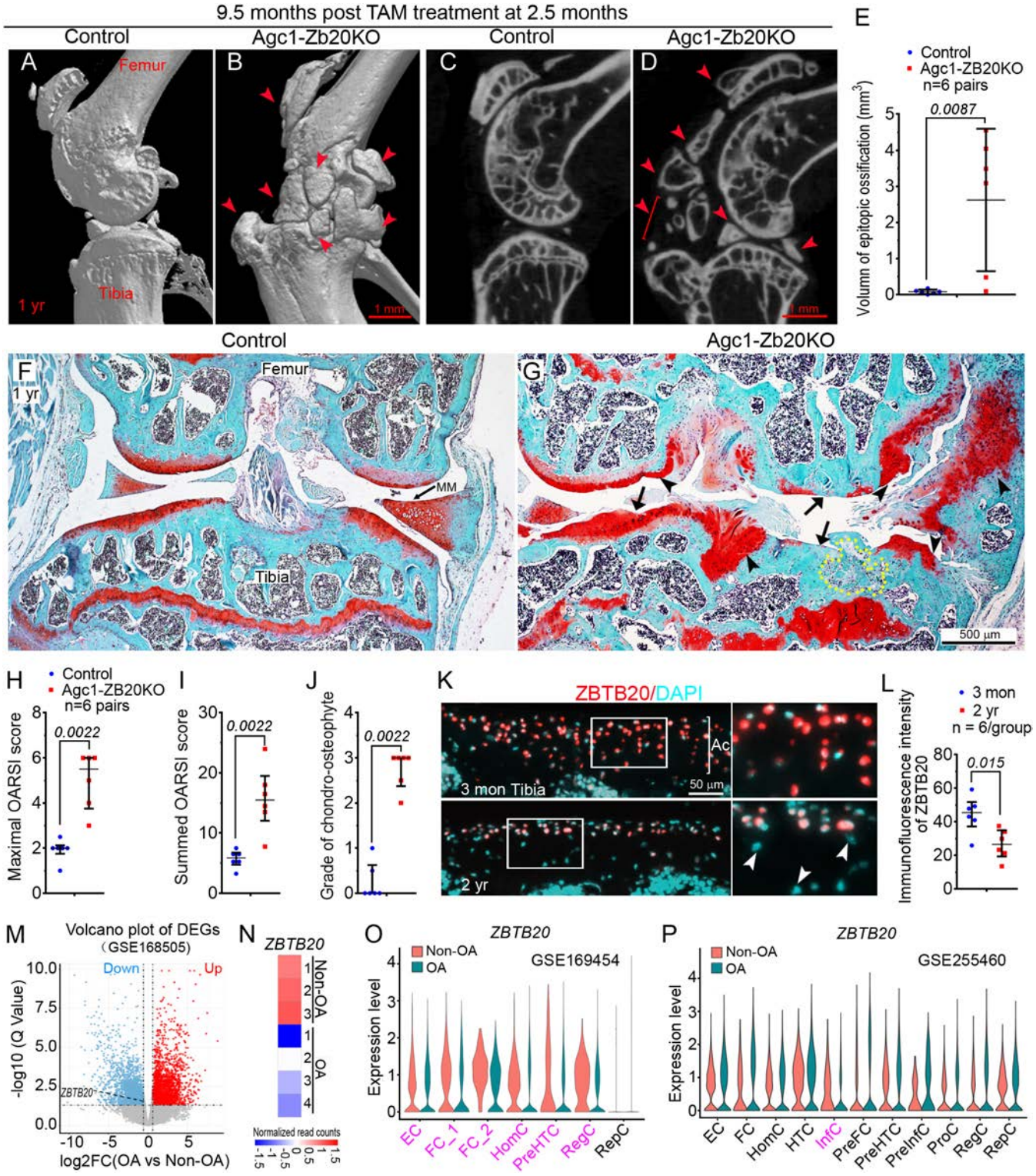


Figure 6. Legend on next page.

OA.[40] Bulk RNA-seq analysis revealed a modest downregulation of *ZBTB20* mRNA in OA cartilage compared with non-OA cartilage ($log_2FC = –0.6$, Q value = 0.029) (Figure 6M and N). Consistently, scRNA-seq data showed *ZBTB20* downregulation across six chondrocyte clusters in OA cartilage (Figure 6O).

To minimize bias from interpatient variability, we also analyzed GSE255460, an additional scRNA-seq dataset from femoral condyle cartilage of three non-OA controls (mean age 51.67 ± 4.9 years) and eight patients with OA (mean age 64.33 ± 7.6 years).[41] This analysis revealed significant downregulation of *ZBTB20* mRNA in inflammatory chondrocytes (InfC) from OA samples, with unchanged expression in other chondrocyte clusters (Figure 6P). The InfC cluster, enriched in OA samples, is believed to play a role in OA pathogenesis.[41] Furthermore, another scRNA-seq analysis comparing knee articular cartilage from healthy individuals (n = 6, mean age 35 ± 9 years) and those with OA (n = 6, mean age 71 ± 6 years) revealed significant downregulation of *ZBTB20* mRNA in a senescent pathogenic chondrocyte population.[42] Together, these results suggest that ZBTB20 plays a crucial role in maintaining joint integrity, and its downregulation is associated with age-related OA in both mice and humans.

DISCUSSION

This study reveals a novel role for ZBTB20 in postnatal articular cartilage development and homeostasis. We demonstrate that ZBTB20 regulates genes essential for articular development and ECM organization, particularly by suppressing the expression of a growth-promoting imprinted gene network during juvenile-adolescent stages. Loss of ZBTB20 leads to progressive thickening of the deep and calcified zones and early-onset articular cartilage damage. Furthermore, ZBTB20 plays a protective role in the progression of age-related OA.

A prominent phenotype observed in Agc1-ZB20KO mice was the abnormal thickening of articular cartilage, with a 1.9–fold increase in thickness at two months. Normally, cartilage thickness peaks around P28 and decreases by two months in mice,[4,43] but ZBTB20 loss disrupts this process. We propose that this is due to altered chondrocyte proliferation, differentiation, ECM architecture, and ossification.

Chondrocyte overproliferation was mainly observed in the mutant articular transitional zone, rather than the superficial zone. Transitional chondrocytes typically have a higher proliferation rate than superficial cells during postnatal life,[44,45] highlighting the importance of regulating proliferation in this zone for proper cartilage development. Increased proliferation alone, however, may not fully explain the thickening in the deep and calcified zones. This is more likely due to defects in chondrocyte differentiation and maturation. Previous studies suggest that articular chondrocytes undergo an endochondral ossification-like process during postnatal development, in which newly generated chondrocytes from the superficial/transitional zone differentiate into mature chondrocytes in the deep/calcified zone. These mature chondrocytes then undergo apoptosis or transdifferentiation at the chondro-osseous junction, contributing to the endochondral growth of underlying epiphyseal bones.[44,45] In *Zbtb20*-deficient knees, disrupted chondrocyte differentiation may delay this process, leading to articular thickening, and may partially contribute to a reduced SOC area. This is supported by the downregulation of hypertrophic markers (*Ibsp* and *Timp3*)[23,46] and upregulation of immature markers (*Epyc*, *Jun*, and *Tnc*).[24,25,47]

Alterations in ECM architecture and turnover may further contribute to this thickening. Future studies using zone-specific Cre lines could provide further insights into ZBTB20's role in cartilage zonal development and homeostasis.

We also observed upregulation of eight imprinted genes in *Zbtb20*-deficient articular cartilage, five of which (*Zim1*, *Dlk1*, *Peg3*, *Igf2*, and *Plagl1*) contain ZBTB20-binding sites in their cis-regulatory elements, suggesting direct transcriptional inhibition by ZBTB20. These genes, known to promote early growth, are down-regulated in cartilage around weaning.[31–33] Our findings imply that ZBTB20 coordinates the downregulation of these genes, which may be essential for cartilage maturation.

Induced deletion of *Zbtb20* in postnatal immature cartilage resulted in early-onset OA-like changes, likely due to developmental defects, including altered chondrocyte differentiation, increased apoptosis, and ECM abnormalities. In adult cartilage,

Figure 6. Inducible deletion of *Zbtb20* in mature cartilage causes severe spontaneous OA-like changes. Micro-CT analysis reveals exacerbated ectopic ossification and osteophyte formation (red arrowheads) in one-year-old Agc1-ZB20KO knees ([B], [D]) compared with WT controls ([A], [C]). (E) Bar graph showing significantly increased osteophyte volumes in the mutant compared to WT. Representative Safranin O/Fast Green-stained knee sections show severe cartilage destruction in Agc1-ZB20KO (G) compared to WT (F). Arrows indicate erosion, arrowheads indicate ectopic chondrogenesis/ossification, and the dotted circle indicates a subchondral bone cyst and microfracture. Graphs show significantly increased maximal (H) and summed OARSI scores (I), chondro-osteophyte formation (J) in Agc1-ZB20KO knees. n = 6 mice/group. Data are presented as median with 25–75% IQR. Representative immunofluorescence images (K) and a corresponding graph (L) show decreased ZBTB20 immunofluorescence intensity (indicated by arrowheads) in tibial articular cartilage of two-year-old WT mice. Nuclei were counterstained with DAPI (turquoise). Volcano plot (M) and heatmap (N) illustrating reduced expression of *ZBTB20* mRNA in human OA cartilage (GSE168505). (O, P) Violin plots show decreased *ZBTB20* expression in specific chondrocyte subsets (highlighted in magenta) from human OA knee cartilage (scRNA-seq datasets: GSE169454 and GSE255460). See also Supplementary Figures 7–8 and Supplementary Tables 15–20. IQR, interquartile range; MM, medial meniscus; micro-CT, microcomputed tomography; mRNA, messenger RNA; OA, osteoarthritis; OARSI, Osteoarthritis Research Society International; scRNA-seq, single-cell RNA sequencing; WT, wild-type. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.43393/abstract.

Zbtb20 deletion led to extensive ectopic chondrogenesis and ossification in the synovium-capsule and cruciate ligaments of mid-aged mice. No recombination activity was detected in cells within these regions, suggesting that these defects originate from *Zbtb20*-deleted chondrocytes. Our scRNA analysis, along with previous studies,[41,42] revealed that ZBTB20 is down-regulated in inflammatory and senescent pathogenic chondrocytes in human OA cartilage. Because inflammation is widely recognized as a key trigger for ectopic ossification,[48] we hypothesize that ZBTB20 loss induces cellular senescence and chronic inflammation in aging cartilage, promoting OA progression and ectopic ossification.

Our results are consistent with skeletal phenotypes observed in Primrose syndrome, which include epiphyseal abnormalities and articular destruction.[12–14] However, they contrast with a recent study showing that *Zbtb20* knockout in 10-week-old Col2a1-CreERT2 mice attenuated destabilization of medial meniscus–induced OA.[49] This suggests that ZBTB20 may have opposing effects in the progression of post-traumatic OA in young adults versus spontaneous OA in aging mice, indicating distinct molecular mechanisms in these models.[50] Further investigation is needed to clarify ZBTB20's context-dependent role in OA progression.

In conclusion, our study highlights ZBTB20 as essential for postnatal joint development and homeostasis. Along with our previous findings on its role in endochondral ossification and long bone growth,[9] these findings underscore its multifaceted roles in skeletal development and cartilage homeostasis.

ACKNOWLEDGMENTS

During the preparation of this work, the authors used ChatGPT and Grammarly in order to improve its clarity. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding authors, Drs Xie and W. J. Zhang confirm that all authors have provided the final approval of the version to be published and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. Pueyo Moliner A, Ito K, Zaucke F, et al. Restoring articular cartilage: insights from structure, composition and development. Nat Rev Rheumatol 2025;21(5):291–308.
2. Decker RS. Articular cartilage and joint development from embryogenesis to adulthood. Semin Cell Dev Biol 2017;62:50–56.
3. Xie M, Chagin AS. The epiphyseal secondary ossification center: evolution, development and function. Bone 2021;142:115701.
4. Decker RS, Um HB, Dyment NA, et al. Cell origin, volume and arrangement are drivers of articular cartilage formation, morphogenesis and response to injury in mouse limbs. Dev Biol 2017;426(1):56–68.
5. Xie Z, Zhang H, Tsai W, et al. Zinc finger protein ZBTB20 is a key repressor of alpha-fetoprotein gene transcription in liver. Proc Natl Acad Sci USA 2008;105(31):10859–10864.
6. Xie Z, Ma X, Ji W, et al. Zbtb20 is essential for the specification of CA1 field identity in the developing hippocampus. Proc Natl Acad Sci USA 2010;107(14):6510–6515.
7. Nielsen JV, Nielsen FH, Ismail R, et al. Hippocampus-like corticoneurogenesis induced by two isoforms of the BTB-zinc finger gene Zbtb20 in mice. Development 2007;134(6):1133–1140.
8. Cao D, Ma X, Cai J, et al. ZBTB20 is required for anterior pituitary development and lactotrope specification. Nat Commun 2016;7(1):11121.
9. Zhou G, Jiang X, Zhang H, et al. Zbtb20 regulates the terminal differentiation of hypertrophic chondrocytes via repression of Sox9. Development 2015;142(2):385–393.
10. Xie Z, Ma XH, Bai QF, et al. ZBTB20 is essential for cochlear maturation and hearing in mice. Proc Natl Acad Sci USA 2023;120(24):e2220867120.
11. Cordeddu V, Redeker B, Stellacci E, et al. Mutations in ZBTB20 cause Primrose syndrome. Nat Genet 2014;46(8):815–817.
12. Primrose DA. A slowly progressive degenerative condition characterized by mental deficiency, wasting of limb musculature and bone abnormalities, including ossification of the pinnae. J Ment Defic Res 1982;26(Pt 2):101–106.
13. Arora V, Leon E, Diaz J, et al. Unique skeletal manifestations in patients with Primrose syndrome. Eur J Med Genet 2020;63(8):103967.
14. Melis D, Carvalho D, Barbaro-Dieber T, et al. Primrose syndrome: characterization of the phenotype in 42 patients. Clin Genet 2020;97(6):890–901.
15. Henry SP, Jang CW, Deng JM, et al. Generation of aggrecan-CreERT2 knockin mice for inducible Cre activity in adult cartilage. Genesis 2009;47(12):805–814.
16. Blumer MJ, Longato S, Schwarzer C, et al. Bone development in the femoral epiphysis of mice: the role of cartilage canals and the fate of resting chondrocytes. Dev Dyn 2007;236(8):2077–2088.
17. Grogan SP, Duffy SF, Pauli C, et al. Zone-specific gene expression patterns in articular cartilage. Arthritis Rheum 2013;65(2):418–428.
18. Hidaka C, Cheng C, Alexandre D, et al. Maturational differences in superficial and deep zone articular chondrocytes. Cell Tissue Res 2006;323(1):127–135.
19. Gannon JM, Walker G, Fischer M, et al. Localization of type X collagen in canine growth plate and adult canine articular cartilage. J Orthop Res 1991;9(4):485–494.
20. Pulik Ł, Łęgosz P, Motyl G. Matrix metalloproteinases in rheumatoid arthritis and osteoarthritis: a state of the art review. Reumatologia 2023;61(3):191–201.
21. Mukherjee A, Das B. The role of inflammatory mediators and matrix metalloproteinases (MMPs) in the progression of osteoarthritis. Biomater Biosyst 2024;13:100090.
22. Vallet SD, Ricard-Blum S. Lysyl oxidases: from enzyme activity to extracellular matrix cross-links. Essays Biochem 2019;63(3):349–364.
23. Holm E, Aubin JE, Hunter GK, et al. Loss of bone sialoprotein leads to impaired endochondral bone development and mineralization. Bone 2015;71:145–154.

24. Wilson R, Norris EL, Brachvogel B, et al. Changes in the chondrocyte and extracellular matrix proteome during post-natal mouse cartilage development. Mol Cell Proteomics 2012;11(1):M111014159.
25. Mackie EJ, Ramsey S. Expression of tenascin in joint-associated tissues during development and postnatal growth. J Anat 1996;188(Pt 1):157–165.
26. Lin W, Xu L, Li G. Molecular insights into lysyl oxidases in cartilage regeneration and rejuvenation. Front Bioeng Biotechnol 2020;8:359.
27. He X, Ohba S, Hojo H, et al. AP-1 family members act with Sox9 to promote chondrocyte hypertrophy. Development 2016;143(16): 3012–3023.
28. Eelen G, Verlinden L, Maes C, et al. Forkhead box O transcription factors in chondrocytes regulate endochondral bone formation. J Steroid Biochem Mol Biol 2016;164:337–343.
29. Weng LH, Wang CJ, Ko JY, et al. Control of Dkk-1 ameliorates chondrocyte apoptosis, cartilage destruction, and subchondral bone deterioration in osteoarthritic knees. Arthritis Rheum 2010;62(5):1393–1402.
30. Fang C, Zhu S, Zhong R, et al. CDKN1A regulation on chondrogenic differentiation of human chondrocytes in osteoarthritis through single-cell and bulk sequencing analysis. Heliyon 2024;10(5):e27466.
31. Varrault A, Gueydan C, Delalbre A, et al. Zac1 regulates an imprinted gene network critically involved in the control of embryonic growth. Dev Cell 2006;11(5):711–722.
32. Lui JC, Finkielstain GP, Barnes KM, et al. An imprinted gene network that controls mammalian somatic growth is down-regulated during postnatal growth deceleration in multiple organs. Am J Physiol Regul Integr Comp Physiol 2008;295(1):R189–R196.
33. Andrade AC, Lui JC, Nilsson O. Temporal and spatial expression of a growth-regulated network of imprinted genes in growth plate. Pediatr Nephrol 2010;25(4):617–623.
34. Kaya-Okur HS, Wu SJ, Codomo CA, et al. CUT&Tag for efficient epigenomic profiling of small samples and single cells. Nat Commun 2019;10(1):1930.
35. Meers MP, Tenenbaum D, Henikoff S. Peak calling by Sparse Enrichment Analysis for CUT&RUN chromatin profiling. Epigenetics Chromatin 2019;12(1):42.
36. Creyghton MP, Cheng AW, Welstead GG, et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc Natl Acad Sci USA 2010;107(50):21931–21936.
37. Wang Z, Zang C, Rosenfeld JA, et al. Combinatorial patterns of histone acetylations and methylations in the human genome. Nat Genet 2008;40(7):897–903.
38. Heintzman ND, Stuart RK, Hon G, et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet 2007;39(3):311–318.
39. Bae S, Lesch BJ. H3K4me1 Distribution predicts transcription state and poising at promoters. Front Cell Dev Biol 2020;8:289.
40. Fu W, Hettinghouse A, Chen Y, et al. 14-3-3 epsilon is an intracellular component of TNFR2 receptor complex and its activation protects against osteoarthritis. Ann Rheum Dis 2021;80(12):1615–1627.
41. Fan Y, Bian X, Meng X, et al. Unveiling inflammatory and prehypertrophic cell populations as key contributors to knee cartilage degeneration in osteoarthritis using multi-omics data integration. Ann Rheum Dis 2024;83(7):926–944.
42. Swahn H, Li K, Duffy T, et al. Senescent cell population with ZEB1 transcription factor as its main regulator promotes osteoarthritis in cartilage and meniscus. Ann Rheum Dis 2023;82(3):403–415.
43. Rux D, Decker RS, Koyama E, et al. Joints in the appendicular skeleton: developmental mechanisms and evolutionary influences. Curr Top Dev Biol 2019;133:119–151.
44. Hunziker EB, Kapfinger E, Geiss J. The structural architecture of adult mammalian articular cartilage evolves by a synchronized process of tissue resorption and neoformation during postnatal development. Osteoarthritis Cartilage 2007;15(4):403–413.
45. Li L, Newton PT, Bouderlique T, et al. Superficial cells are self-renewing chondrocyte progenitors, which form the articular cartilage in juvenile mice. FASEB J 2017;31(3):1067–1084.
46. Apte SS, Hayashi K, Seldin MF, et al. Gene encoding a novel murine tissue inhibitor of metalloproteinases (TIMP), TIMP-3, is expressed in developing mouse epithelia, cartilage, and muscle, and is located on mouse chromosome 10. Dev Dyn 1994;200(3):177–197.
47. Kan A, Tabin CJ. c-Jun is required for the specification of joint cell fates. Genes Dev 2013;27(5):514–524.
48. Song JH, Liu MY, Ma YX, et al. Inflammation-associated ectopic mineralization. Fundam Res 2022;3(6):1025–1038.
49. Hao X, Zhao J, Jia L, et al. LATS1-modulated ZBTB20 perturbing cartilage matrix homeostasis contributes to early-stage osteoarthritis. Bone Res 2025;13(1):33.
50. Masson AO, Krawetz RJ. Understanding cartilage protection in OA and injury: a spectrum of possibilities. BMC Musculoskelet Disord 2020;21(1):432.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis & Rheumatology
Vol. 78, No. 3, March 2026, pp 644–653
DOI 10.1002/art.43397

The Influence of Race, Ethnicity and Historical Redlining on Psoriatic Disease Burden and Clinical Outcomes

Sharon Dowell,[1] Brittany Banbury,[2] Christopher Jenkins,[3,4,5] Emily E. Holladay,[6] Fenglong Xie,[6,7] Jingyi Zhang,[6] Grace C. Wright,[8] Jeffrey R. Curtis,[6,7] and Gail S. Kerr[3,4,9]

Objective. The study objective was to evaluate the association of neighborhood deprivation and structural barriers with disease burden in racial and ethnic subsets of patients with psoriatic arthritis (PsA).

Methods. Patients with PsA in the Rheumatology Informatics System for Effectiveness registry of the American College of Rheumatology with reported race, region, and three or more years of follow-up were evaluated. Demographic factors, disease activity measures, social deprivation defined by the area deprivation index, and therapeutic agents were stratified by race. Subgroup analyses were conducted to examine demographic and clinical characteristics across residential areas graded by the Home Owners' Loan Corporation (HOLC), from "best" (HOLC 1—predominantly White residents) to "hazardous" (HOLC 4 or redlined—predominantly Black residents) based on investment risk.

Results. The cohort included 21,429 predominantly female (57.7%) and obese (56.1% body mass index >30) patients with PsA, with median age of 55 (12.8) years. High social deprivation was prevalent among Black patients (25.7 % vs 2.3% Asian, 12.5% White, and 17.3% other), as was high disease activity (HDA; 40.2% vs 25.8% Asian, 29.6% White, and 33.5% other). Approximately 7% of patients with PsA lived in HOLC-graded districts. Smoking, obesity, high social deprivation, federal insurance, and HDA were more prevalent in patients in HOLC 4 areas compared to HOLC 1 areas. HOLC 4 patients also had longer median [interquartile range] periods of HDA (105.0 [0–690] person-days) and fewer days in remission (1.0 [0–5,457] person-days).

Conclusion. In the United States, Black patients with PsA have prevalent HDA and high social deprivation. Additionally, the enduring effects of structural racism appear to negatively influence PsA disease characteristics of patients living in historically redlined areas.

INTRODUCTION

Asian, Black, multiracial, and Hispanic White patients with psoriatic disease (PD) have greater comorbidity, more extensive skin disease, and higher disease activity with more frequent axial involvement.[1–4] Compared to their White counterparts, these ethnic minority individuals more often have diagnostic delays and a longer time to initiation of disease-modifying antirheumatic drugs (DMARDs) along with less use of biologic medications and worse clinical outcomes.[2,5] Both socioeconomic status (SES) and neighborhood social deprivation contribute to these health disparities in PD, although there are conflicting findings regarding the high prevalence of psoriatic arthritis (PsA) in both high[6] and low[7] socioeconomic groups. SES is influenced by a complex interplay of several factors, which range from the individual level (education, occupation, and income) to the broader structural and policy environment (housing, health care, and social security policies) in which people live and work. Neighborhood-level factors such as housing quality, public services, environmental conditions, and social cohesion further influence individual outcomes

Supported as an investigator-initiated study funded by AbbVie Inc. Infrastructure support for Rheumatology Informatics System for Effectiveness registry analyses at The University of Alabama at Birmingham and The Foundation for Advancing Science, Technology, Education and Research was provided in part by NIH (P30-AR-072583).

[1]Northeast Arkansas Baptist Memorial Hospital, Jonesboro; [2]Mount Sinai Beth Israel, New York, New York; [3]Washington Veterans Affairs Medical Center, Washington, DC; [4]Georgetown University, Washington, DC; [5]Keck School of Medicine-Los Angeles General Medical Center, Los Angeles, California; [6]University of Alabama at Birmingham; [7]Foundation for Advancing Science, Technology, Education, and Research, Birmingham, Alabama; [8]Association of Women in Rheumatology, New York, New York; [9]Howard University College of Medicine, Washington, DC.

Additional supplementary information cited in this article can be found online in the Supporting Information section (http://onlinelibrary.wiley.com/doi/10.1002/art.43397).

Author disclosures and graphical abstract are available at https://onlinelibrary.wiley.com/doi/10.1002/art.43397.

Address correspondence via email to Sharon Dowell, MD, at sharon.dowell@awirgroup.org.

Submitted for publication February 25, 2025; accepted in revised form August 6, 2025.

by either amplifying or mitigating disadvantage.[8] Although multiple social factors may contribute to PsA outcomes, the association of neighborhood-level deprivation on disease expression in PsA has not been studied.

In the United States, the spectrum of bias and prejudices based on race is intertwined with legislation that has affected health and patient care. Furthermore, practices established by the federal Home Owners' Loan Corporation (HOLC) in the 1930s to promote affordable homeownership disproportionately disadvantaged minority populations[9,10] with the sequelae of limited education and employment opportunities and enduring low SES. Private banks adopted directives from the HOLC that examined the racial and ethnic makeup of neighborhoods and income data to determine risk for mortgage lending.[11,12] Areas that had predominantly Black residents and low income often were deemed at high risk for mortgage lending and shaded red on the HOLC maps, in a practice referred to as "redlining." These redlined neighborhoods, today home to more than 8 million people,[13] remain financially disadvantaged[11] and, by extension, have compromised access to health care and a lack of expanded health care support.[14,15] The toll of these policies in these affected areas is underscored by the higher rates of chronic diseases such as heart disease, diabetes, and asthma, and more frequent adverse clinical outcomes.[16,17] This is an added contributor to the overall area deprivation and the sequela of health care disparity.

A recent report delineates the impact of historical redlining on increased health care resource utilization.[18] Literature describing the impact of neighborhood deprivation and structural barriers in the United States on PD characteristics and disease-specific outcomes is limited. In this study, we aimed to evaluate the association of neighborhood social deprivation, captured by the area deprivation index (ADI) and HOLC historical redlining maps,[10,19] with PsA disease burden and treatment among racial and ethnic subsets of patients with PsA. We hypothesized that ethnic minority individuals had greater PsA disease burden, less biologic DMARD (bDMARD) use, higher social vulnerability as measured by the ADI, and a greater likelihood to live in a HOLC-identified high risk area. We further evaluated the impact of residence in historically redlined areas on PsA disease burden.

METHODS

Study population. A retrospective cohort analysis of the Rheumatology Informatics System for Effectiveness (RISE) registry data[20] was performed. Patients >18 years of age, with documented race and region of residence, two or more rheumatology provider visits between October 01, 2016, and September 30, 2020 (the PsA case ascertainment period), and at least three years of follow-up receiving care at active RISE sites were included, with an analysis cutoff date of September 30, 2023. PsA diagnosis was confirmed by the presence of two or more International Classification of Diseases, Ninth (ICD-9) or Tenth Revision, Clinical Modification (ICD-10-CM) codes (ICD-9 696.0 and/or ICD-10-CM L40.5*) separated by at least 28 days and occurring within 365 days, with current or previous use of PsA-related medication prescriptions or procedures (injection or infusion of PsA medications) up to 12 months before the index date. The index date was defined as the second ambulatory visit with a PsA diagnosis code in the case ascertainment period. Case identification algorithms requiring two or more ICD PsA diagnosis codes from rheumatology providers and prescription of disease-specific therapy have been shown to have high positive predictive values for PsA.[21,22] PsA medications included methotrexate and other conventional synthetic DMARD agents; bDMARD agents (tumor necrosis factor alpha inhibitors, a costimulation modulator, interleukin-17 inhibitors, interleukin-12/23 inhibitors); and targeted synthetic DMARD (tsDMARD) agents including a phosphodiesterase 4 inhibitor, apremilast, and Janus kinase inhibitors (tofacitinib and upadacitinib). Patients with ICD-10-CM codes for rheumatoid arthritis, ankylosing spondylitis, systemic lupus erythematosus, systemic vasculitis, or gout at any time before the index date were excluded.

Variables. Specific variables included self-reported race and ethnicity, neighborhood SES,[19] age at index date, body mass index (BMI; >30 indicates obesity), comorbidities, tobacco use, PsA disease characteristics (Multidimensional Health Assessment Questionnaire, pain, physician global assessment, patient global assessment, Routine Assessment of Patient Index Data 3 [RAPID3],[23] erythrocyte sedimentation rate, c-reactive protein), insurance coverage and type (government [Medicare, Medicaid], private/commercial), and medication use including glucocorticoids (lower-dose steroid <10 mg, higher-dose ≥10 mg of daily prednisone or equivalent), nonsteroidal anti-inflammatory drugs, and opioids. Sociodemographics, PsA disease status, and medications were extracted at index. Medication use, including baseline therapy and initiation of therapy during follow-up, were recorded and used to assess burden of disease and provider adherence to treatment guidelines for active PsA.[24,25] High disease activity (HDA) was defined as a RAPID3 score of >12 (0–30). We used RAPID3 scores and the trapezoid rule to estimate the amount of time spent in disease activity categories during follow-up. Comorbidity was assessed using the RxRisk Comorbidity Index, a validated pharmacy-based measure of comorbidity that uses documented prescription medicines to estimate the number of conditions treated out of 46 disease entities.[26] High comorbidity was further defined as an RxRisk Comorbidity Index above the 90th percentile. Distance to care was defined as the median distance between the centroid of the patients' home address zip code and practice site zip code.

Social deprivation and HOLC. Socioeconomic deprivation experienced according to neighborhood was assessed using

the ADI, with vulnerable patients defined by an ADI ≥80 (the most disadvantaged 20% of neighborhoods). The ADI is a validated neighborhood level score of socioeconomic disadvantage based on several census-derived variables (from income, education, housing, and employment domains) and ranges from 1 to 100, with an ADI ≥80 (ie, the 80th percentile) reflecting a previously established indicator of high social deprivation.[19]

Characteristics of patients with PsA were stratified by HOLC score. We converted nine-digit zip codes to census tract five-digit Federal Information Processing Series codes to use the University of Michigan's 2020 census tract HOLC redlining scores.[27] HOLC scores were reported from 1 to 4, with 1 denoting an A grade or best and 4 denoting a D grade or hazardous residence (redlined). Redlining scores were then used to stratify PsA clinical characteristics to evaluate associations with geographic areas that have experienced structural racism. Historically, only approximately 16% of all 2020 US Census Tracts received these scores.[28,29] The proportion of US census tracts does not account for population sizes.

Statistical analyses. *Hypothesis 1: Ethnic minority patients have more severe PsA disease activity (Hypothesis 1a) and less use of b/tsDMARDs for PsA (Hypothesis 1b).* Demographics, disease activity, social deprivation defined by high ADI (≥80), and use of biologic and tsDMARD PsA treatments were examined overall and stratified by race. Ordinal alternating logistic regression was used to evaluate the association between demographic variables and both baseline use of PsA medications and follow-up (over 36 months) initiation of new PsA medications. The dependent variable in these models was the categorical number of new PsA medications used and classified during baseline as three or more, exactly two, or exactly one PsA medication, and during the three-year follow-up, it was classified as two or more, exactly one, and no (0) new PsA medications newly initiated. Ordinal regression was used and controlled for the clustering of patients within physician practice. We controlled for other potentially confounding factors including age, sex, and (for the longitudinal model) the number of baseline PsA medications. We then used ordinary least squares regression and generalized estimating equations to evaluate the association between race and ethnicity and the proportion of days (as a percentage) spent in HDA during the three years of follow-up. We evaluated crude models evaluating only race and ethnicity that accounted for the effect of clustering of patients with PsA within physician practice. We then sequentially adjusted for potentially confounding factors to identify which factors were most influential in the race and ethnicity associations. These additional factors included insurance type (government [Medicare, Medicaid], private/commercial), geographic region, and sex. We additionally evaluated the bivariate association with each potential confounder (sex, comorbidity, insurance type, geographic region) to evaluate which factors most strongly influenced the effect estimate with race observed.

Hypothesis 2: Living in a historically "hazardous" (HOLC 4) area is associated with Black and Hispanic White race and ethnicity and associated with higher disease activity than living in a historically "best" (HOLC 1) area. A subgroup analysis assessing demographic and clinical characteristics of patients with PsA stratified by historical residential security grades was also conducted. Although the legacy of HOLC neighborhood grading persists, its effects are most pronounced in HOLC 1 (best) and HOLC 4 (hazardous) areas,[30] which we selected for comparison in this study. Patients who lived in areas without an HOLC score were excluded from the analysis. For disease activity, the analysis was further restricted to those patients with RAPID3 at baseline and at least three distinct RAPID3 measurements in the disease activity category during the follow-up period. We assessed the descriptive statistics (ie, mean ± SD and mean [interquartile range; IQR]) for medications used, visits with a rheumatologist, PsA medications per visit, and person-days in disease activity category stratified by HOLC category. *P* values were provided using Kruskal-Wallis rank sum and Pearson's χ^2 tests.

RESULTS

The cohort comprised 21,429 patients with PsA (Supplemental Table 1—Attrition Table). The median (IQR) age was 56 (47–65) years, with 57.7% of female sex, 52.2% in Southern residence, and 56.1% characterized as obese (BMI >30) (Table 1). Of the total participants, 595 individuals (3.2%) identified as Hispanic, representing 25% of the "other" group. Black patients with PsA constituted a small minority of the cohort (n = 388, 1.8%), were predominantly female (73.5%), resided in the South (77.1%), and had obesity (73.2%) compared to other ethnic groups. Black patients with PsA also had a greater prevalence of HDA (40.2% vs Asian: 25.8%, other: 33.5%, White: 29.6%), high comorbidity burden (mean ± SD RxRisk 8.7 ± 4.1 vs Asian: 5.9 ± 3.5, other: 7.1 ± 3.9, and White patients: 7.9 ± 4.2), and high social deprivation (ADI ≥80th percentile) (25.7% vs Asian: 2.3%, other 17.3%, White patients 12.5%). With regard to PsA therapies, Black patients had greater use of oral (12.6% vs Asian: 9.8%, other: 7.7%, and White patients: 10.2%) and injectable glucocorticoids (30.9% vs Asian: 19.0%, other: 21.2% and White patients: 20.7%) and less use of bDMARDs (54.9% vs Asian: 59.9%, other: 57.8%, and White patients: 60.4%).

Over three years of follow-up, patients with PsA had a median (IQR) of 8.0 (6.0–10.0) visits with their rheumatologist and a median (IQR) of 25 (0.0–599.5) person-days in HDA and 147 (0.0–513.5) person-days in moderate disease activity (Table 2). Black and other patients with PsA spent a greater proportion of time in HDA (Black: 37.5%, other: 32.1%) compared to White (27.1%) or Asian (26.3%) patients with PsA ($P < 0.0001$). Of patients with a RAPID3 measurement at the index date and three or more during follow-up (n = 5,999), the median number of new PsA medications in the three years following the index date was 1.0 (0.0–2.0). Most patients (67.4%) with HDA

Table 1. Baseline characteristics of patients with PsA in the RISE registry, by race*

Characteristic	Overall, n = 21,429	Asian, n = 327	Black, n = 388	Other, n = 987	White, n = 19,727	*P* value[a]
Age, median (IQR), y	56.0 (47.0–65.0)	52.0 (41.0–61.0)	57.0 (48.0–65.0)	53.0 (43.0–61.0)	57.0 (47.0–65.0)	0.00
Female	12,354 (57.7)	168 (51.4)	285 (73.5)	559 (56.6)	11,342 (57.5)	0.00
Region						0.00
Midwest	5,059 (23.6)	37 (11.3)	60 (15.5)	315 (31.9)	4,647 (23.6)	
Northeast	3,138 (14.6)	30 (9.2)	20 (5.2)	71 (7.2)	3,017 (15.3)	
South	11,195 (52.2)	153 (46.8)	299 (77.1)	458 (46.4)	10,285 (52.1)	
West	2,037 (9.5)	107 (32.7)	9 (2.3)	143 (14.5)	1,778 (9.0)	
Ethnicity						0.00
Hispanic	595 (3.2)	4 (1.6)	5 (1.5)	206 (25.0)	380 (2.2)	
Non-Hispanic	18,105 (96.8)	254 (98.4)	334 (98.5)	618 (75.0)	16,899 (97.8)	
Unknown	2,729	69	49	163	2,448	
Insurance						0.00
Medicaid	737 (3.6)	6 (1.9)	34 (9.0)	52 (6.0)	645 (3.4)	
Medicare	8,124 (39.6)	63 (20.3)	160 (42.6)	257 (29.7)	7,644 (40.3)	
Private	10,989 (53.6)	209 (67.4)	175 (46.5)	533 (61.5)	10,072 (53.2)	
Other insurance	648 (3.2)	32 (10.3)	7 (1.9)	24 (2.8)	585 (3.1)	
Unknown	931	17	12	121	781	
Ever smoker	6,196 (32.6)	65 (20.8)	122 (33.9)	255 (30.1)	5,754 (32.9)	0.00
Unknown	2,419	14	28	139	2,238	
BMI category						0.00
Underweight, <18.5	259 (1.6)	11 (4.2)	1 (0.3)	11 (1.5)	236 (1.6)	
Healthy, 18.5–25	2,189 (13.3)	68 (25.7)	26 (8.7)	81 (11.2)	2,014 (13.3)	
Overweight, 25–30	4,747 (28.9)	115 (43.4)	53 (17.8)	213 (29.3)	4,366 (28.9)	
Obesity, ≥30	9,211 (56.1)	71 (26.8)	218 (73.2)	421 (58.0)	8,501 (56.2)	
Unknown	5,023	62	90	261	4,610	
RAPID3 (last on or before index date)[b]						0.03
Remission	2,239 (26.6)	46 (34.8)	33 (20.1)	96 (23.0)	2,064 (26.8)	
Low	1,485 (17.6)	21 (15.9)	25 (15.2)	76 (18.2)	1,363 (17.7)	
Moderate	2,171 (25.8)	31 (23.5)	40 (24.4)	106 (25.4)	1,994 (25.9)	
High (≥12)	2,519 (29.9)	34 (25.8)	66 (40.2)	140 (33.5)	2,279 (29.6)	
Unknown	13,015	195	224	569	12,027	
csDMARD[c]	11,600 (54.1)	178 (54.4)	225 (58.0)	527 (53.4)	10,670 (54.1)	0.46
tsDMARD[c]	3,134 (14.6)	48 (14.7)	67 (17.3)	180 (18.2)	2,839 (14.4)	0.00
bDMARD[c]	14,085 (65.7)	218 (66.7)	233 (60.1)	635 (64.3)	12,999 (65.9)	0.08
bDMARD or JAK[c,d]	14,227 (66.4)	223 (68.2)	238 (61.3)	642 (65.0)	13,124 (66.5)	0.12
DMARD hierarchy[c,e]						0.01
tsDMARDs	3,134 (14.6)	48 (14.7)	67 (17.3)	180 (18.2)	2,839 (14.4)	
bDMARDs	12,903 (60.2)	196 (59.9)	213 (54.9)	570 (57.8)	11,924 (60.4)	
cDMARDs	5,392 (25.2)	83 (25.4)	108 (27.8)	237 (24.0)	4,964 (25.2)	
Opioid[c]	1,851 (8.6)	10 (3.1)	36 (9.3)	101 (10.2)	1,704 (8.6)	0.00
NSAID[c]	10,228 (47.7)	133 (40.7)	191 (49.2)	467 (47.3)	9,437 (47.8)	0.07
Oral steroid[c]	2,171 (10.1)	32 (9.8)	49 (12.6)	76 (7.7)	2,014 (10.2)	0.03
Injectable glucocorticoids[c]	4,466 (20.8)	62 (19.0)	120 (30.9)	209 (21.2)	4,075 (20.7)	0.00
RxRisk[f]						0.00
Mean ± SD	7.8 ± 4.1	5.9 ± 3.5	8.7 ± 4.1	7.1 ± 3.9	7.9 ± 4.2	
Median (IQR)	7.0 (5.0–11.0)	5.0 (3.0–8.0)	8.0 (5.0–12.0)	7.0 (4.0–10.0)	7.0 (5.0–11.0)	
ADI ≥80	2,585 (12.8)	7 (2.3)	92 (25.7)	158 (17.3)	2,328 (12.5)	0.00
Unknown	1,220	20	30	75	1,095	
HOLC						0.02
1 (best)	270 (12.1)	0 (0.0)	8 (11.0)	10 (7.4)	252 (12.6)	
2	380 (17.0)	4 (12.1)	12 (16.4)	17 (12.5)	347 (17.4)	
3	558 (24.9)	12 (36.4)	19 (26.0)	27 (19.9)	500 (25.1)	
4 (most hazardous)	1,029 (46.0)	17 (51.5)	34 (46.6)	82 (60.3)	896 (44.9)	
HOLC not classified	19,192	294	315	851	17,732	

* Values are n (%) unless otherwise indicated. csDMARDs include cyclosporine, leflunomide, methotrexate, and sulfasalazine. tsDMARDs include apremilast, baricitinib, tofacitinib, upadacitinib, and deucravacitinib. bDMARDs include abatacept, adalimumab, brodalumab, certolizumab, etanercept, golimumab, guselkumab, infliximab, ixekizumab, risankizumab, secukinumab, ustekinumab, bimekizumab, and tildrakizumab. ADI, area deprivation index; bDMARD, biologic DMARD; BMI, body mass index; csDMARD, conventional synthetic DMARD; DMARD, disease-modifying antirheumatic drug; HOLC, Home Owners' Loan Corporation; IQR, interquartile range; NSAID, nonsteroidal anti-inflammatory drug; PsA, psoriatic arthritis; RAPID3, routine assessment of patient index data 3; RISE, Rheumatology Informatics System for Effectiveness; tsDMARD, targeted synthetic DMARD.

[a] Kruskal-Wallis rank sum test; Pearson's χ^2 test.

[b] Remission: RAPID3 0–3; low: 3–6; moderate: 6–12; high: 12–30.

[c] Baseline: 0–12 months before index date.

[d] Includes bDMARDs, upadacitinib, tofacitinib, and baricitinib.

[e] Hierarchy: tsDMARD > bDMARD > csDMARD.

[f] History: on or before index date using all available data. Highest RxRisk quartile: history/baseline RxRisk >90th percentile, RxRisk ≥11.

Table 2. Health care utilization and PsA medication changes and over 36 months in patients with PsA with baseline patient-reported outcome data (n = 5,599)*

Characteristic	Overall, n = 5,599	Asian, n = 85	Black, n = 114	Other, n = 293[a]	White, n = 5,107	*P* value[b]
Time in various disease activity states (hypothesis 1a)						
Person-days in high disease activity						0.00
Mean ± SD	302.0 ± 411.2	287.6 ± 406.8	410.9 ± 454.7	351.9 ± 406.6	297.0 ± 410.1	
Median (IQR)	25.0 (0.0–599.5)	25.0 (0.0–482.0)	204.0 (0.0–934.2)	151.0 (0.0–698.0)	18.0 (0.0–578.0)	
Time in high disease activity, %	27.58	26.30	37.50	32.10	27.10	
Unique patients contributing to high disease activity, n (%)	2,968 (53.0)	45 (52.9)	69 (60.5)	186 (63.5)	2,668 (52.2)	0.00
Person-days in moderate disease activity						0.05
Mean ± SD	287.1 ± 332.2	214.2 ± 294.0	268.1 ± 301.8	303.9 ± 298.6	287.8 ± 335.1	
Median (IQR)	147.0 (0.0–513.5)	77.0 (0.0–348.0)	157.5 (0.0–446.2)	259.0 (0.0–501.0)	143.0 (0.0–517.0)	
Time in moderate disease activity, %	26.22	19.60	24.50	27.70	26.30	
Unique patients contributing to moderate disease activity, n (%)	3,738 (66.8)	51 (60.0)	82 (71.9)	215 (73.4)	3,390 (66.4)	0.03
Person-days in low disease activity						0.88
Mean ± SD	216.3 ± 292.4	154.3 ± 209.5	199.1 ± 262.6	197.2 ± 270.6	218.8 ± 295.3	
Median (IQR)	70.0 (0.0–347.5)	82.0 (0.0–224.0)	78.5 (0.0–309.8)	59.0 (0.0–312.0)	70.0 (0.0–351.0)	
Time in low disease activity, %	19.75	14.10	18.20	18.00	20.00	
Unique patients contributing to low disease activity, n (%)	3,416 (61.0)	58 (68.2)	73 (64.0)	179 (61.1)	3,106 (60.8)	0.50
Person-days in remission						0.01
Mean ± SD	289.5 ± 405.8	439.0 ± 460.1	216.8 ± 335.0	242.0 ± 373.2	291.4 ± 407.4	
Median (IQR)	1.0 (0.0–587.5)	270.0 (0.0–932.0)	2.5 (0.0–326.0)	1.0 (0.0–386.0)	1.0 (0.0–595.5)	
Time in remission, %	26.44	40.10	19.80	22.10	26.60	
Unique patients contribute to remission disease activity, n (%)	2,859 (51.1)	54 (63.5)	59 (51.8)	147 (50.2)	2,599 (50.9)	0.14
Total days, mean ± SD	1,095 ± 0	1,095 ± 0	1,095 ± 0	1,095 ± 0	1,095 ± 0	0.99
Initiation of PsA medications (Hypothesis 1a)						
No. new PsA medications						0.03
Mean ± SD	1.02 ± 1.33	0.96 ± 1.31	1.00 ± 1.22	1.18 ± 1.32	1.01 ± 1.33	
Median (IQR)	1.0 (0.0–2.0)	0.0 (0.0–1.0)	1.0 (0.0–2.0)	1.0 (0.0–2.0)	1.0 (0.0–2.0)	
No. PsA visits						0.00
Mean ± SD	8.42 ± 3.7	8.12 ± 3.7	9.05 ± 4.8	7.07 ± 3.7	8.49 ± 3.6	
Median (IQR)	8.00 (6.0–10.0)	8.00 (5.0–10.0)	9.00 (6.2–11.0)	7.00 (5.0–9.0)	8.00 (6.0–11.0)	
No. new PsA medications per office visit						0.00
Mean ± SD	0.1 ± 0.2	0.1 ± 0.2	0.1 ± 0.2	0.2 ± 0.3	0.1 ± 0.2	
Median (IQR)	0.1 (0.0–0.2)	0.0 (0.0–0.2)	0.1 (0.0–0.2)	0.1 (0.0–0.3)	0.1 (0.0–0.2)	
Undefined	18	0	1	0	17	

* Number of PsA visits was defined as Current Procedural Terminology 99XXX codes occurring after index. Patients who only had a visit >1,095 days after index are categorized as undefined for the assessment of medications per visit. Number of new PsA medications over 36 months was defined as the number of new PsA medications used during the 36-month follow-up. A new medication is defined as one that was never used during baseline. Subgroup of Table 1 patients with PsA, baseline RAPID3, and 3 distinct RAPID3 scores during the follow-up period by time in disease activity category. IQR, interquartile range; PsA, psoriatic arthritis.
[a] Other race includes American Indian or Alaska Native, multiracial, Native Hawaiian or other Pacific Islander, no determined Office of Management and Budget race classification.
[b] Kruskal-Wallis rank sum test.

did not have a change in PsA medication in the subsequent 12 months, irrespective of race.

Hypothesis 1a: association between race and ethnicity and disease activity over three years. Black patients were more likely to spend a greater proportion of days in the three-year follow-up in HDA (Table 2). In unadjusted results, they spent 10.4% (95% confidence interval [CI] 1.3–19.6) more of those three years in HDA (P = 0.03). After adjusting for potential confounding factors, race no longer had a significant impact on the proportion of follow-up time in HDA. The factors that were most strongly associated with effect (based upon point estimates) were insurance type (28.0%), sex (9.6%), and ethnicity (Hispanic: 6.9%) (Supplemental Table 2a).

Multivariable adjustment to identify potential explanatory factors revealed that patients insured by Medicaid or Medicare had greater odds of baseline HDA compared to commercially insured patients at baseline (adjusted odds ratio [aOR] 4.30, 95% CI: 2.80–6.58 and aOR 1.88, 95% CI: 1.51–2.34, respectively) and spent a greater proportion of time in HDA during follow-up (27.55%, 95% CI: 19.88–35.21 and 8.64%, 95% CI: 4.23–13.05), controlling for baseline RAPID3 and RxRisk. In the final model, ADI was added to assess social deprivation and time in HDA. Patients with an ADI ≥80th percentile spent approximately 10% more time in HDA (9.60%, 95% CI: 5.34–13.85) compared to those with an ADI <80th percentile. An interaction term was added between ADI and race. The interaction was not statistically significant (Type III P = 0.61), suggesting that the association between ADI and HDA does not vary across racial groups (results not shown in table).

Hypothesis 1b: association between race and ethnicity and PsA medication use. In multivariable-adjusted results (Supplemental Table 2a and 2b), Black patients were 21% less likely to have been prescribed more PsA medications during the baseline period (aOR: 0.79, 95% CI: 0.64–0.97) compared to White patients, and Hispanic patients were 16% less likely compared to non-Hispanic patients (aOR: 0.84, 95% CI: 0.71–1.00). After controlling for these baseline differences including RxRisk comorbidity score, the odds of initiating a new PsA medication over the three-year follow-up period were not different for Black and Hispanic patients relative to their White and non-Hispanic counterparts (aOR: 1.01, 95% CI: 0.85–1.20 and aOR: 1.14, 95% CI: 0.97–1.35, respectively). However, Black patients, along with those classified as other, had more use of apremilast and opioids (Table 1).

Medicaid- and Medicare-insured patients had greater odds of using more PsA medications during baseline, referent to those with commercial insurance (aOR 1.19, 95% CI: 1.02–1.38 and aOR 1.19, 95% CI: 1.11–1.28), and Medicaid coverage was associated with a 19% greater likelihood of new PsA medication use during follow-up (aOR: 1.19, 95% CI: 1.02–1.40) compared to commercially insured patients, controlling for baseline number of PsA medications. When adjusting for baseline RxRisk and number of medications used during the baseline period, this result was no longer significant.

Hypothesis 2 associations between residence in historically redlined area, race and ethnicity, and HDA. Although a majority of study individuals did not reside in areas defined by HOLC zoning, among those who did (6.7%, n = 1,433), racial distribution varied significantly between HOLC scores 1 and 4 (Table 3). Hispanic and other patients had a greater proportion of residence in HOLC 4 compared to HOLC 1 (6.3% vs 1.2% and 12.0% vs 3.7%, respectively). Conversely, White patients comprised a lesser proportion of residency in HOLC 4 (80.4%) compared to HOLC 1 (93.3%). Black patients had the greatest prevalence of living in HOLC 4 areas compared to all areas cross classified by race (3.4% vs White: 0.9%, other: 2.7%, and Asian: 1.2%, results not shown), and a similar trend was seen for Hispanic patients (2.2% vs Non-Hispanic: 1.1%). Compared to HOLC 1, smoking (39.6% vs 28.0), obese BMI (49.4% vs 45.8%), ADI ≥80th percentile (35.2% vs 4.1%), and HDA (34.7% vs 14.4%) were more prevalent in HOLC 4 patients. tsDMARDs (predominantly apremilast) were used more frequently in HOLC 4 compared to HOLC 1 (13.3% vs 9.6%), but the opposite was true for bDMARDs (61.3% vs 63.3%). HOLC 4 patients traveled a mean ± SD of 39.4 ± 208.2 miles for their health care commutes, nearly four times longer than the 10.3 ± 58.2 miles traveled by HOLC 1 patients. Finally, the median (IQR) number of patients per provider was statistically greater in HOLC 4 compared to HOLC 1 (2,442.0 [1,586.0–2,882.3] vs 2,046.7 [1,528.2–2,501.3]), respectively.

Follow-up periods for disease activity were absent for more than 50% of patients in each HOLC category. Available data on 376 patients indicated that patients living in HOLC 4 regions spent more person-days in HDA compared to HOLC 1 (median [IQR]: 105.0 [0.0–689.8] vs 0.0 [0.0–172.5]) and fewer person-days in remission (median [IQR]: 1.0 [0.0–456.5] vs 139.0 [0.0–839.5]) (Supplemental Table 3).

DISCUSSION

Our findings demonstrate that the health care disparities observed in other rheumatic diseases are also present in PsA. Black patients with PsA had a higher burden of comorbidity (including obesity), the greatest proportion of social deprivation, and, over the three years of follow-up, spent more time in HDA than other race groups. They also had more frequent use of Medicare and Medicaid, less use of bDMARDs, and, with those classified as other, more use of apremilast and opioids. Additionally, our analysis is one of the first to report the residual impact of historical redlining on regional health outcomes for diverse patients with PD. Patients residing in historically redlined areas spent more time in HDA and had the highest prevalence of social deprivation. The present-day sequela of historical redlining influences the health disparity for Black, other race, and Hispanic patients who reside in these areas.

Obesity, a well-established contributor to the inflammatory pathogenesis of PD,[31,32] was more prevalent in Black patients and individuals living in historically redlined (HOLC 4) neighborhoods. The co-occurrence of increased disease risk factors and adverse social determinants of health creates a vulnerability for increased disease incidence and poor health outcomes.[33,34] Poor walkability, fewer recreational facilities, and less access to safe green spaces for exercise is a shared template for HOLC 4 areas, which, in addition to fewer options for healthy food, contribute to higher rates of obesity and other comorbid disease.[35–37]

Table 3. Characteristics of patients with PsA included in the RISE registry by residential HOLC score 1–4 at index*

Characteristic	Overall, n = 1,433	HOLC 1 (best), n = 270[a]	HOLC 2, n = 380	HOLC 3, n = 558	HOLC 4 (most hazardous), n = 225	*P* value[a]
Age, median (IQR), y	55.0 (45.0–64.0)	57.0 (47.0–65.8)	55.0 (44.8–64.0)	56.0 (45.0–64.0)	54.0 (42.0–62.0)	0.02
Female	766 (53.5)	135 (50.0)	209 (55.0)	310 (55.6)	112 (49.8)	0.28
Race						0.00
Asian	20 (1.4)	0 (0.0)	4 (1.1)	12 (2.2)	4 (1.8)	
Black	52 (3.6)	8 (3.0)	12 (3.2)	19 (3.4)	13 (5.8)	
Other[b]	81 (5.7)	10 (3.7)	17 (4.5)	27 (4.8)	27 (12.0)	
White	1,280 (89.3)	252 (93.3)	347 (91.3)	500 (89.6)	181 (80.4)	
Ethnicity						0.03
Hispanic	53 (4.1)	3 (1.2)	12 (3.6)	25 (4.9)	13 (6.3)	
Non-Hispanic	1,242 (95.9)	243 (98.8)	325 (96.4)	481 (95.1)	193 (93.7)	
Unknown	138	24	43	52	19	
Insurance						0.05
Medicaid	77 (5.6)	5 (1.9)	15 (4.1)	40 (7.5)	17 (8.1)	
Medicare	480 (35.1)	102 (39.2)	128 (35.4)	181 (34.0)	69 (32.7)	
Private	745 (54.5)	142 (54.6)	202 (55.8)	283 (53.1)	118 (55.9)	
Other insurance	64 (4.7)	11 (4.2)	17 (4.7)	29 (5.4)	7 (3.3)	
Unknown	67	10	18	25	14	
Ever smoker	450 (35.3)	68 (28.0)	115 (33.7)	191 (38.2)	76 (39.6)	0.02
Unknown	157	27	39	58	33	
BMI						0.07
Underweight, <18.5	19 (1.8)	3 (1.6)	3 (1.0)	12 (2.8)	1 (0.6)	
Healthy, 18.5–25	168 (15.5)	37 (19.3)	42 (14.4)	63 (14.9)	26 (14.8)	
Overweight, 25–30	333 (30.8)	64 (33.3)	97 (33.3)	110 (26.0)	62 (35.2)	
Obesity, ≥30	562 (51.9)	88 (45.8)	149 (51.2)	238 (56.3)	87 (49.4)	
Unknown	351	78	89	135	49	
RAPID3 (last on or before index date)[c]						0.06
Remission	158 (30.7)	38 (34.2)	33 (28.0)	59 (31.1)	28 (29.5)	
Low	81 (15.8)	20 (18.0)	23 (19.5)	27 (14.2)	11 (11.6)	
Moderate	132 (25.7)	37 (33.3)	27 (22.9)	45 (23.7)	23 (24.2)	
High (≥12)	143 (27.8)	16 (14.4)	35 (29.7)	59 (31.1)	33 (34.7)	
Unknown	919	159	262	368	130	
csDMARD[d]	768 (53.6)	146 (54.1)	216 (56.8)	289 (51.8)	117 (52.0)	0.46
tsDMARD[d]	195 (13.6)	26 (9.6)	48 (12.6)	91 (16.3)	30 (13.3)	0.06
bDMARD[d]	925 (64.5)	180 (66.7)	245 (64.5)	351 (62.9)	149 (66.2)	0.69
bDMARD or JAKi[d]	933 (65.1)	181 (67.0)	248 (65.3)	355 (63.6)	149 (66.2)	0.77
DMARD Hierarchy[d,e]						0.26
tsDMARDs	195 (13.6)	26 (9.6)	48 (12.6)	91 (16.3)	30 (13.3)	
bDMARDs	864 (60.3)	171 (63.3)	228 (60.0)	327 (58.6)	138 (61.3)	
cDMARDs	374 (26.1)	73 (27.0)	104 (27.4)	140 (25.1)	57 (25.3)	
Opioid[d]	113 (7.9)	16 (5.9)	31 (8.2)	49 (8.8)	17 (7.6)	0.55
NSAID[d]	662 (46.2)	125 (46.3)	178 (46.8)	254 (45.5)	105 (46.7)	0.98
Oral steroid[d]	134 (9.4)	28 (10.4)	39 (10.3)	46 (8.2)	21 (9.3)	0.68
Injectable glucocorticoids[d]	290 (20.2)	53 (19.6)	84 (22.1)	102 (18.3)	51 (22.7)	0.39
RxRisk comorbidity score[f]						0.05
Mean ± SD	7.3 ± 4.1	7.2 ± 3.9	7.3 ± 4.1	7.7 ± 4.3	6.7 ± 3.8	
Median (IQR)	7.0 (4.0–10.0)	6.0 (4.0–10.0)	7.0 (4.0–10.0)	7.0 (4.0–11.0)	7.0 (4.0–9.0)	
No. patients per provider, median (IQR)	2,046.7 (1,502.2–2,592.7)	2,046.7 (1,528.2–2,501.3)	2,046.7 (1,528.2–2,441.0)	2,046.7 (1,221.7–2,592.7)	2,442.0 (1,586.0–2,882.3)	0.00
Driving distance						0.00
Mean ± SD	21.5 ± 140.9	10.3 ± 58.2	27.7 ± 193.5	15.3 ± 79.3	39.4 ± 208.2	
Median (IQR)	6.2 (3.5–10.3)	4.3 (2.9–7.7)	6.1 (3.4–9.9)	6.9 (3.9–10.6)	8.1 (4.5–16.7)	
ADI ≥80	258 (18.3)	11 (4.1)	34 (9.0)	136 (24.7)	77 (35.2)	0.00
Unknown[g]	21	3	4	8	6	

* Values are shown as n (%) unless otherwise specified. ADI, area deprivation index; bDMARD, biologic DMARD; BMI, body mass index; csDMARD, conventional synthetic DMARD; DMARD, disease-modifying antirheumatic drug; HOLC, Home Owners' Loan Corporation; IQR, interquartile range; NSAID, nonsteroidal anti-inflammatory drug; PsA, psoriatic arthritis; RAPID3, routine assessment of patient index data 3; RISE, Rheumatology Informatics System for Effectiveness; tsDMARD, targeted synthetic DMARD.

[a] Kruskal-Wallis rank sum test; Pearson's χ^2 test.

[b] Race "other" includes American Indian or Alaska Native, multiracial, Native Hawaiian or other Pacific, no determinate Office of Management and Budget race classification.

[c] Remission: RAPID3 0–3, low: 3–6; moderate: 6–12; high: 12–30.

[d] Baseline: 0–12 months before index date.

[e] Hierarchy: tsDMARD > bDMARD > csDMARD.

[f] History: on or before index date using all available data.

[g] Unknown ADI due to no ADI for 9-digit zip code.

HOLC 4 residents also experience increased exposure to environmental stressors such as gun violence, home instability, employment uncertainty, and pollutant exposure, which may contribute to documented stress-related exacerbations of autoimmune disease states[38] and mental health disorders (frequently seen in PD).[32] These exposures could lead to increased pain perception and prevalent HDA states.[17] Additional studies have demonstrated higher odds of late-stage cancer diagnosis, chronic kidney disease, asthma, and diabetes in redlined neighborhoods,[17,39,40] potentially affecting the response to therapy, initiation, and continuation of DMARD therapy in patients with rheumatic disease. Obesity has been associated with a lower probability of achieving sustained remission, or even minimal disease activity states in PD, irrespective of therapy used.[31] Whereas obesity may already exacerbate disease and impair DMARD efficacy, prevalent comorbidity negatively influences use of bDMARDs in PD, with both factors playing a potential role in the observed high PsA disease activity,[4] as was observed in our study.

Additional factors are implicated, as living in neighborhoods with concentrated disadvantages limits opportunities for upward mobility, such as education and stable employment options.[41] Low-income Black patients with chronic disease are more often unemployed and receive Medicaid, or are uninsured, compared to their middle income counterparts who have stable employment with private insurance.[42] Attaining financial stability facilitates access to health care and health-promoting behaviors. Black and Hispanic patients in our study were more often in HDA but without PsA medications at baseline, suggesting delayed diagnosis and delays in starting targeted therapy, which ultimately affects treatment outcomes. Poor health care infrastructure in historically redlined neighborhoods affects delivery of equitable care. Because "safety net" facilities and community health centers often lack adequate resources and suffer from poor retention of health care providers,[14] patients may have to travel further for rheumatology care. Additionally, pharmacies in historically redlined neighborhoods are often scarce,[43,44] and those that are present may lack the ability to dispense or coordinate specialized medications, compared to pharmacies in less segregated areas.[14] Home delivery may improve access but is limited by lack of a supportive delivery infrastructure, safety concerns, and digital access gaps in these neighborhoods.

Although some historically redlined districts have undergone selective redevelopment, many individuals residing in these neighborhoods continue to live in poorly maintained housing.[11,45] It is possible that these poor conditions may limit the ability to refrigerate and store biologic medications, influence treatment options, and explain the higher use of tsDMARDs (apremilast) over bDMARDs among RISE patients living in historically redlined areas.[46,47] Apremilast is not considered first-line therapy in current professional treatment guidelines for PsA and is generally reserved for oligoarticular, mild PsA or existing contraindications to bDMARD therapy,[48,49] limited variables to evaluate in our study. Consequently, predominant use of tsDMARDs such as apremilast and less frequent use of bDMARDs[50] may prevent the achievement of remission.

An unexpected finding of our study was the relatively infrequent PsA medication change for all patients over three years of follow-up, regardless of geographic location, race, or ethnicity, and even for those with persistent HDA, suggesting a gap in quality of care. The greater odds of starting a new PsA medication for Medicaid patients but higher proportion of time spent in HDA during follow-up may reflect refractory and more severe disease in this patient population or, as mentioned previously, limited options or patient acceptance for preferred medications. These factors may explain the greater use of opioids in Black patients, regardless of residence. Hence, there is the need for greater support and use of appropriate DMARD therapies.

Our study is limited by its cross-sectional view, a high level of missingness for longitudinal data, and with a very small portion of the cohort living in redlined areas. Furthermore, because the RISE registry is largely subscribed from community-based, nonacademic practices and lacks individual patient socioeconomic data, its findings may have limited generalizability. The reliance on zip code–level data to calculate distance to care may not accurately reflect precise travel distances or capture patient mobility among zip codes during the study period. The concepts of ADI and HOLC 4 are interlinked; hence residence in HOLC 4 region may more reflect the impact of SES than geography. Further studies would need to assess HOLC 4 patients' access, distance traveled, and number of rheumatology visits related to disease activity states to preferred rheumatology care, as well as the role of patient preference on type and recommended therapy.

Strengths of this report include it being one of the first to report the association between structural racism and the characteristics and outcomes of a rheumatic disease. These findings provide new perspectives for identifying areas for intervention, investment, and development. Our study is also one of the first to show the association between historically redlined areas and worse health outcomes for all ethnicities with PD, emphasizing the need for policies focused on enhancing health infrastructure in these areas, and for all ethnic groups.

The effects of systemic racism may be another contributing component to ethnic and racial variation in severity and management of PD. Our study highlights that geographic location—further defined by structural racism—may have a significant effect on patient characteristics, outcomes, and quality of life. There is much yet to be done to rectify the enduring and multifaceted health consequences of residing in historically redlined areas.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software,

investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr Dowell confirms that all authors have provided the final approval of the version to be published and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

ROLE OF THE STUDY SPONSOR

Abbvie Inc had no role in the study design, or in the collection, analysis or interprtation of the data, the writing of the manuscript or decision to submit the manuscript for publication. Publication of this article was not contingent upon approval by Abbvie Inc.

REFERENCES

1. Haberman RH, Ahmed T, Um S, et al. Racial and ethnic determinants of psoriatic arthritis phenotypes and disease activity. Rheumatology (Oxford) 2025;64(2):574–580.
2. Kerr GS, Qaiyumi S, Richards J, et al. Psoriasis and psoriatic arthritis in African-American patients--the need to measure disease burden. Clin Rheumatol 2015;34(10):1753–1759.
3. Ogdie A, Matthias W, Thielen RJ, et al. Racial differences in prevalence and treatment for psoriatic arthritis and ankylosing spondylitis by insurance coverage in the USA. Rheumatol Ther 2021;8(4):1725–1739.
4. Ross Y, Jaleel S, Magrey M. Racial disparities in comorbidities of patients with psoriatic arthritis. Rheumatol Int 2023;43(8):1525–1529.
5. Patel AA, Ferrante SA, Lin I, et al. Racial and ethnic disparities in treatment initiation among patients with newly diagnosed psoriatic arthritis: a retrospective medicaid claims database study. Rheumatol Ther 2023;10(5):1241–1253.
6. Haddad A, Elkayam PC, Stein N, et al. Epidemiological trends in psoriatic arthritis: a comprehensive population-based study. Arthritis Res Ther 2024;26(1):108.
7. Kerola AM, Sexton J, Wibetoe G, et al. Incidence, sociodemographic factors and treatment penetration of rheumatoid arthritis and psoriatic arthritis in Norway. Semin Arthritis Rheum 2021;51(5):1081–1088.
8. Dahlgren G, Whitehead M. Policies and Strategies to Promote Social Equity in Health. Institute for Future Studies; 1991.
9. Mehdipanah R, McVay KR, Schulz AJ. Historic redlining practices and contemporary determinants of health in the Detroit metropolitan area. Am J Public Health 2023;113(S1):S49–S57.
10. Nelson RK, Winling L, Marciano R, et al. Mapping inequality: redlining in New Deal America. Accessed November 1, 2024. https://dsl.richmond.edu/panorama/redlining
11. Egede LE, Walker RJ, Campbell JA, et al. Modern day consequences of historic redlining: finding a path forward. J Gen Intern Med 2023;38(6):1534–1537.
12. Swope CB, Hernández D. Housing as a determinant of health equity: a conceptual model. Soc Sci Med 2019;243:112571.
13. Meier HC, Mitchell BC. Tracing the legacy of redlining: a new method for tracking the origins of housing segregation. National Community Reinvestment Coalition website. 2022. Accessed December 3, 2025. https://ncrc.org/redlining-score.
14. White K, Haas JS, Williams DR. Elucidating the role of place in health care disparities: the example of racial/ethnic residential segregation. Health Serv Res 2012;47(3 Pt 2):1278–1299.
15. Swope CB, Hernández D, Cushing LJ. The relationship of historical redlining with present-day neighborhood environmental and health outcomes: a scoping review and conceptual model. J Urban Health 2022;99(6):959–983.
16. Motairek I, Lee EK, Janus S, et al. Historical neighborhood redlining and contemporary cardiometabolic risk. J Am Coll Cardiol 2022;80(2):171–175.
17. Lee EK, Donley G, Ciesielski TH, et al. Health outcomes in redlined versus non-redlined neighborhoods: a systematic review and meta-analysis. Soc Sci Med 2022;294:114696.
18. Yang S, Santacroce L, Collins JE, et al. Association of historical redlining and present-day racialised economic segregation with health-care utilisation among individuals with rheumatic conditions in Massachusetts and surrounding areas of the USA: a retrospective cohort study. Lancet Rheumatol 2025;7(1):e33–e43.
19. Kind AJH, Buckingham WR. Making neighborhood-disadvantage metrics accessible — the neighborhood atlas. N Engl J Med 2018;378(26):2456–2458.
20. Yazdany J, Bansback N, Clowse M, et al. Rheumatology Informatics System for Effectiveness: a national informatics-enabled registry for quality improvement. Arthritis Care Res (Hoboken) 2016;68(12):1866–1873.
21. Singh JA, Holmgren AR, Noorbaloochi S. Accuracy of Veterans Administration databases for a diagnosis of rheumatoid arthritis. Arthritis Rheum 2004;51(6):952–957.
22. Kim SY, Servi A, Polinski JM, et al. Validation of rheumatoid arthritis diagnoses in health care utilization data. Arthritis Res Ther 2011;13(1):R32.
23. Coates LC, Tillett W, Shaddick G, et al. Value of the routine assessment of patient index data 3 in patients with psoriatic arthritis: results from a tight-control clinical trial and an observational cohort. Arthritis Care Res (Hoboken) 2018;70(8):1198–1205.
24. Singh JA, Guyatt G, Ogdie A, et al. Special Article: 2018 American College of Rheumatology/National Psoriasis Foundation Guideline for the Treatment of Psoriatic Arthritis. Arthritis Rheumatol 2019;71(1):5–32.
25. Coates LC, Soriano ER, Corp N, et al; GRAPPA Treatment Recommendations domain subcommittees. Group for Research and Assessment of Psoriasis and Psoriatic Arthritis (GRAPPA): updated treatment recommendations for psoriatic arthritis 2021. Nat Rev Rheumatol 2022;18(8):465–479.
26. Sloan KL, Sales AE, Liu CF, et al. Construction and characteristics of the RxRisk-V: a VA-adapted pharmacy-based case-mix instrument. Med Care 2003;41(6):761–774.
27. Meier HC, Mitchell BC. Historic Redlining Scores for 2010 and 2020 US Census Tracts. Inter-university Consortium for Political and Social Research; 2021.
28. Meier HC, Mitchell BC. *Historic Redlining Indicator for* 2000, *2010, and 2020 US Census Tracts*. Inter-university Consortium for Political and Social Research; 2023.
29. Tallies. US Census Bureau website. Accessed May 20, 2025. https://www.census.gov/geographies/reference-files/time-series/geo/tallies.html
30. Mitchell B, Franco J. HOLC "redlining" maps: The persistent structure of segregation and economic inequality. National Community Reinvestment Coalition; 2018.
31. Ogdie A, Schwartzman S, Husni ME. Recognizing and managing comorbidities in psoriatic arthritis. Curr Opin Rheumatol 2015;27(2):118–126.
32. Perez-Chada LM, Merola JF. Comorbidities associated with psoriatic arthritis: review and update. Clin Immunol 2020;214:108397.
33. Vo A, Tao Y, Li Y, et al. The association between social determinants of health and population health outcomes: ecological analysis. JMIR Public Health Surveill 2023;9:e44070.

34. Cockerham WC, Hamby BW, Oates GR. The social determinants of chronic disease. Am J Prev Med 2017;52(1):S5–S12.

35. Sallis JF, Glanz K. Physical activity and food environments: solutions to the obesity epidemic. Milbank Q 2009;87(1):123–154.

36. Hill JO, Peters JC. Environmental contributions to the obesity epidemic. Science 1998;280(5368):1371–1374.

37. Mujahid MS, Gao X, Tabb LP, et al. Historical redlining and cardiovascular health: The Multi-Ethnic Study of Atherosclerosis. Proc Natl Acad Sci USA 2021;118(51):e2110986118.

38. Porcelli B, Pozza A, Bizzaro N, et al. Association between stressful life events and autoimmune diseases: a systematic review and meta-analysis of retrospective case–control studies. Autoimmun Rev 2016;15(4):325–334.

39. Krieger N, Wright E, Chen JT, et al. Cancer stage at diagnosis, historical redlining, and current neighborhood characteristics: breast, cervical, lung, and colorectal cancers, Massachusetts, 2001-2015. Am J Epidemiol 2020;189(10):1065–1075.

40. Nguyen KH, Buckle-Rashid R, Thorsness R, et al. Structural racism, historical redlining, and incidence of kidney failure in US cities, 2012-2019. J Am Soc Nephrol 2023;34(9):1493–1503.

41. Faber JW. Cashing in on distress: the expansion of fringe financial institutions during the great recession. Urban Aff Rev 2018;54(4): 663–696.

42. Becker G, Newsom E. Socioeconomic status and dissatisfaction with health care among chronically ill African Americans. Am J Public Health 2003;93(5):742–748.

43. Guadamuz JS, Alexander GC, Kanter GP, et al. More US pharmacies closed than opened in 2018-21; independent pharmacies, those in Black, Latinx communities most at risk. Health Aff (Millwood) 2024; 43(12):1703–1711.

44. Appolon G, Tang S, Gabriel N, et al. Association between redlining and spatial access to pharmacies. JAMA Netw Open 2023;6(8): e2327315.

45. Geller AB, Polsky DE, Burke SP, eds. Neighborhood and built environment. In: Federal Policy to Advance Racial, Ethnic, and Tribal Health Equity. National Academies; 2023.

46. Bailey ZD, Krieger N, Agénor M, et al. Structural racism and health inequities in the USA: evidence and interventions. Lancet 2017; 389(10077):1453–1463.

47. Yu YB, Briggs KT, Taraban MB, et al. Grand challenges in pharmaceutical research series: ridding the cold chain for biologics. Pharm Res 2021;38(1):3–7.

48. Singh JA, Furst DE, Bharat A, et al. 2012 Update of the 2008 American College of Rheumatology recommendations for the use of disease-modifying antirheumatic drugs and biologic agents in the treatment of rheumatoid arthritis. Arthritis Care Res 2012;64(5): 625–639.

49. Gossec L, Baraliakos X, Kerschbaumer A, et al. EULAR recommendations for the management of psoriatic arthritis with pharmacological therapies: 2019 update. Ann Rheum Dis 2020;79(6):700–712.

50. Zhang H, Wen J, Alexander GC, et al. Comparative effectiveness of biologics and targeted therapies for psoriatic arthritis. RMD Open 2021;7(1):e001399.

Arthritis & Rheumatology
Vol. 78, No. 3, March 2026, pp 654–664
DOI 10.1002/art.70044

American College of Rheumatology
Empowering Rheumatology Professionals

Glucagon-Like Peptide 1 Receptor Agonists, Sodium-Glucose Cotransporter 2 Inhibitors, and Risk of Autoimmune Rheumatic Diseases

Derin Karacabeyli, MD,[1,2,3] Diane Lacaille, MD, MHSc,[1,2,3] Na Lu, MPH,[3] Hui Xie, PhD,[3,4] and J. Antonio Aviña-Zubieta, MD, PhD[1,2,3]

Objective. Glucagon-like peptide 1 receptor agonists (GLP-1RAs) and sodium-glucose cotransporter 2 inhibitors (SGLT2i) facilitate weight loss and exhibit immunomodulatory effects, but their impact on the risk of developing autoimmune rheumatic diseases (ARDs) is unclear. We compared ARD incidence following initiation of GLP-1RAs or SGLT2i versus a weight-neutral comparator (dipeptidyl peptidase 4 inhibitors [DPP4i]).

Methods. We performed a population-based cohort study using administrative health data from a Canadian province with universal health care. We included adults with type 2 diabetes (T2D) and no prior ARD initiating a GLP-1RA, SGLT2i, or DPP4i between January 1, 2014, and December 31, 2022. Incident ARD cases, including rheumatoid arthritis, psoriatic disease, axial spondyloarthritis, and systemic ARDs (SARDs; systemic lupus erythematosus, systemic sclerosis, Sjögren disease, idiopathic inflammatory myopathies, and systemic vasculitides), were identified using validated algorithms. Propensity score (PS) weighting was used to balance cohorts at treatment initiation. Then hazard ratios (HRs) were estimated using Cox regression.

Results. Among 229,300 adults, 49,514 initiated GLP-1RAs, 101,925 initiated SGLT2i, and 77,861 initiated DPP4i. After PS weighting, ARD incidence per 10,000 person-years was 29.1 (95% confidence interval [CI] 23.5–35.5) with GLP-1RAs, 24.4 (95% CI 19.8–29.7) with SGLT2i, and 27.3 (95% CI 22.1–33.4) with DPP4i. Mean follow-up was 1.3 to 1.6 years. Relative to DPP4i, adjusted HRs (aHRs) of ARD were 1.04 (95% CI 0.81–1.33) with GLP-1RAs and 0.93 (95% CI 0.75–1.16) with SGLT2i. Risk of SARDs, but not other diseases, was lower with SGLT2i versus DPP4i (aHR 0.51 [95% CI 0.31–0.84]).

Conclusion. Neither GLP-1RA nor SGLT2i treatment was associated with increased or decreased ARD risk versus DPP4i in adults with T2D; however, SGLT2i use was associated with significantly lower risk of SARDs, warranting further study.

INTRODUCTION

Over 500 million adults worldwide have diabetes mellitus, and type 2 diabetes (T2D) accounts for >90% of cases.[1] Although developed as antidiabetic medications, glucagon-like peptide 1 receptor agonists (GLP-1RAs) and sodium-glucose cotransporter 2 inhibitors (SGLT2i) have become increasingly used following discoveries of various pleiotropic effects.[2] Both drug

The following data sets were used in this study: Consolidation, Hospital Separations, Medical Services Plan, Pharmanet, National Ambulatory Care Reporting System, and Vital Statistics-Death. Further information regarding these data sets can be found by visiting the PopData project webpage at https://my.popdata.bc.ca/project_listings/14-131/collection_approval_dates. All inferences, opinions, and conclusions drawn in this publication are those of the authors, and do not reflect the opinions or policies of the Data Steward(s).

Supported by Preventing Complications from Inflammatory Skin, Joint and Bowel Conditions (PRECISION), a team grant funded by the Canadian Institutes of Health Research (grant THC-135235). Dr Karacabeyli's work was supported by the University of British Columbia (UBC) Clinician Investigator Program and a Canadian Institutes of Health Research Canada Graduate Scholarship–Master's Award. Dr Lacaille's work was supported by the Mary Pack Chair in Rheumatology Research from UBC and the Arthritis Society Canada. Dr Xie's work was supported by the Milan Ilich/Merck Chair in Statistics for Arthritis and Musculoskeletal Diseases from Simon Fraser University. Dr Aviña-Zubieta is a British Columbia Lupus Society Research Scholar and a Walter & Marilyn Booth Research Scholar.

[1]Division of Rheumatology, University of British Columbia, Vancouver, British Columbia, Canada; [2]Experimental Medicine, Department of Medicine, University of British Columbia, Vancouver, British Columbia, Canada; [3]Arthritis Research Canada, Vancouver, British Columbia, Canada; [4]Faculty of Health Sciences, Simon Fraser University, Burnaby, British Columbia, Canada.

We are unable to share the data publicly for ethical and legal reasons. Population Data BC (PopData), a pan-provincial, multi-institutional platform, is the custodian of these data. Researchers may access these data by working directly with PopData and following PopData processes. Upon approval, data will be provisioned by PopData and released on the PopData Secure Research Environment. Researchers may contact PopData at dataaccess@popdata.bc.ca regarding access to the data.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/art.70044).

Author disclosures and graphical abstract are available at https://onlinelibrary.wiley.com/doi/10.1002/art.70044.

Address correspondence via email to J. Antonio Aviña-Zubieta, MD, PhD at azubieta@arthritisresearch.ca.

Submitted for publication June 28, 2025; accepted in revised form December 8, 2025.

classes promote weight loss, lower the risk of major adverse cardiovascular and renal events, and improve survival.[3–5] Randomized controlled trials have reported placebo-subtracted weight loss of 3% to 18% with GLP-1RAs, and 2% with SGLT2i.[6,7] Beyond their established benefits, GLP-1RAs and SGLT2i have demonstrated positive immunomodulatory effects in preclinical studies.[6,7] However, the clinical implications of these effects remain uncertain. Excess adiposity is a known risk factor for autoimmune rheumatic diseases (ARDs) such as rheumatoid arthritis (RA), psoriatic arthritis (PsA), and axial spondyloarthritis (axSpA).[8–10] Based on this, we hypothesized that GLP-1RAs and SGLT2i may reduce risk of developing ARDs through weight loss–dependent and independent immunomodulatory mechanisms.

To date, studies have yielded conflicting results. A retrospective cohort study observed lower risk of certain systemic ARDs (SARDs) among users of GLP-1RAs; however, the study lacked an active comparator and did not adjust for key confounders, including comorbidities and health care use.[11] In contrast, two cases of new RA-like syndromes after GLP-1RA exposure, with symptom resolution after drug discontinuation, have been reported.[12,13] Evidence regarding SGLT2i and autoimmune disease risk is also limited and inconsistent. Three cohort studies investigating the association between SGLT2i and psoriasis reported discordant findings: two reported no difference in psoriasis risk with SGLT2i exposure, whereas one observed an 8% increased risk relative to dipeptidyl peptidase 4 inhibitors (DPP4i).[14–16]

In light of existing uncertainty, we conducted a population-based cohort study with separate analyses comparing incidence of ARDs following treatment initiation with GLP-1RAs, or SGLT2i versus DPP4i. DPP4i served as an active comparator given their similar indication as second-line therapy for T2D, their weight-neutral profile, and their lack of known association with increased ARD risk.[17–21]

PATIENTS AND METHODS

Study design. We performed a propensity score (PS)–weighted, population-based new-user cohort study using administrative health data from British Columbia (BC), Canada, a province with a universal health care system. The objective was to compare the incidence of ARDs following initiation of GLP-1RAs or SGLT2i relative to DPP4i in adults with T2D.

Data sources and population. The source population included all residents of BC between January 1, 2014, and December 31, 2022 (population range: ~4.7–5.4 million).[22] Health data were obtained through Population Data BC, which provides individual-level linkage across several administrative databases, including physician billing claims (one diagnostic code per visit) and hospital admissions and discharges (up to 25 diagnostic codes) since 1990,[23,24] emergency department visits since 2011,[25] vital statistics,[26] and all pharmacy-dispensed medications regardless of source of funding (including dose, date, and quantity dispensed) since 1996.[27]

Cohort definition. We identified adults (≥18 years old) with T2D and no ARD diagnosis in the previous 10 years who initiated a new treatment course of a GLP-1RA, SGLT2i, or DPP4i between January 1, 2014, and December 31, 2022, and observed them until December 31, 2023 (Supplementary Figure S1). The index date was the date of drug initiation. Drugs included are listed in Supplementary Table S1. People initiating a new treatment course were required to have a 12-month washout from all three drug classes; otherwise, they were excluded. Patients with fewer than two years of continuous health plan enrollment before the index date were excluded to ensure that no prevalent users moving into the province were included. Diagnoses of T2D and ARDs, including RA, psoriatic disease, axSpA, and SARDs (including systemic lupus erythematosus [SLE], systemic sclerosis [SSc], Sjögren disease, idiopathic inflammatory myopathies, giant cell arteritis, antineutrophil cytoplasmic antibody–associated vasculitis, Takayasu arteritis, and polyarteritis nodosa) were ascertained using validated *International Classification of Diseases, Ninth or Tenth Revision* (ICD-9 or ICD-10)–based algorithms[28–33] (Supplementary Table S2). These algorithms have specificities >90% and positive predictive values between 78% and 98% in validation studies.[28–33]

The unit of analysis was a treatment course. Patients were observed from the index date (ie, beginning of the treatment course) until first ARD diagnosis, medication discontinuation, switch/add-on of a comparator drug, death, loss of health plan enrollment (for people moving out of the province), or December 31, 2023, whichever occurred first. The exposure medication was considered discontinued if 60 days elapsed after the expiration date of the last prescription's supply without a refill. A patient could contribute more than one treatment course but only after a >12-month washout period of all three drug classes.

Outcomes. The primary outcome was incidence of ARDs (composite of RA, psoriatic disease, axSpA, and SARDs). Secondary outcomes included each ARD subtype individually. Psoriatic disease included both psoriasis and PsA, as BC's administrative data do not allow for distinguishing the two.

Covariates. Baseline covariates were selected based on their potential to influence treatment selection or ARD diagnosis (Table 1). These included sociodemographic factors, comorbidities (including overall comorbidity burden using the Romano modification of the Charlson Comorbidity Index for administrative data),[34] immunomodulatory and autoimmune disease–triggering medications, and health care use patterns. Diabetes complications and duration, prior use of antihyperglycemic agents, prior

Table 1. Baseline characteristics among patients with type 2 diabetes initiating GLP-1RAs, SGLT2i, or DPP4i*

Variable list	Before matching weights				After matching weights[a]			
	GLP-1RA (n = 49,514)	SGLT2i (n = 101,925)	DPP4i (n = 77,861)	SMD	GLP-1RA (n = 24,560.4)	SGLT2i (n = 24,648.1)	DPP4i (n = 24,495.3)	SMD
Demographics								
Age, mean (SD), y	56.0 (12.7)	61.7 (12.2)	64.4 (13.0)	0.44	59.9 (11.7)	59.9 (12.0)	59.3 (13.0)	0.03
Male, n (%) or %[a]	18,239 (36.8)	62,814 (61.6)	45,147 (58.0)	0.34	48.0	46.9	48.4	0.019
Neighborhood income quintile, n (%) or %[a]				0.163				0.008
1 = lowest	9,197 (18.6)	23,380 (22.9)	19,253 (24.7)		20.6	20.5	20.4	
2	9,483 (19.2)	22,777 (22.3)	18,276 (23.5)		20.6	20.8	20.7	
3	9,875 (19.9)	20,703 (20.3)	15,445 (19.8)		20.4	20.4	20.4	
4	10,189 (20.6)	18,201 (17.9)	12,853 (16.5)		19.0	19.0	19.2	
5 = highest	10,203 (20.6)	15,826 (15.5)	11,138 (14.3)		17.9	17.8	17.9	
Unknown	567 (1.1)	1,038 (1.0)	896 (1.2)		1.5	1.4	1.5	
Health region, n (%) or %[a]				0.101				0.017
Interior	6,285 (12.7)	14,172 (13.9)	9,902 (12.7)		12.9	12.4	12.7	
Fraser	22,532 (45.5)	45,894 (45.0)	36,291 (46.6)		46.6	47.0	47.0	
Vancouver Coastal	10,147 (20.5)	21,643 (21.2)	18,726 (24.1)		21.8	22.4	21.8	
Vancouver Island	7,654 (15.5)	14,123 (13.9)	8,924 (11.5)		13.4	13	13.2	
Northern	2,668 (5.4)	5,696 (5.6)	3,619 (4.6)		4.7	4.6	4.7	
Unknown	228 (0.5)	397 (0.4)	399 (0.5)		0.6	0.6	0.6	
Mean Charlson Comorbidity Index (SD)	1.8 (2.6)	1.8 (2.6)	1.9 (2.8)	0.014	1.6 (2.6)	1.6 (2.5)	1.5 (2.5)	0.022
Comorbidities, n (%) or %[a]								
Obesity[b]	10,396 (21.0)	5,859 (5.7)	2,638 (3.4)	0.377	6.9	6.8	6.9	0.004
MI[b]	2,577 (5.2)	9,811 (9.6)	5,864 (7.5)	0.113	5.8	5.5	5.4	0.011
Stroke or TIA[b]	3,865 (7.8)	9,610 (9.4)	8,182 (10.5)	0.063	7.8	7.6	7.4	0.011
VTE[b]	1,286 (2.6)	2,171 (2.1)	1,592 (2.0)	0.024	2.3	2.2	2.1	0.009
Cancer[b]	8,446 (17.1)	15,050 (14.8)	11,306 (14.5)	0.046	13.8	13.6	13.4	0.009
Hypertension[c]	12,347 (24.9)	25,059 (24.6)	17,203 (22.1)	0.045	21.4	21.4	20.2	0.021
CHF[c]	1,672 (3.4)	7,518 (7.4)	4,258 (5.5)	0.119	3.5	4.4	3.8	0.031
IHD[c]	3,896 (7.9)	12,421 (12.2)	7,042 (9.0)	0.096	7.9	8.6	7.4	0.031
CKD[c]	4,197 (8.5)	10,192 (10.0)	9,787 (12.6)	0.089	9.6	9.2	8.8	0.019
Chronic liver disease[c]	508 (1.0)	452 (0.4)	283 (0.4)	0.054	0.6	0.4	0.5	0.014
IBD[c]	126 (0.3)	200 (0.2)	161 (0.2)	0.008	0.2	0.2	0.3	0.008
Thyroid disease[c]	2,706 (5.5)	2,696 (2.6)	1903 (2.4)	0.104	3.2	2.9	2.8	0.018
Infection[c]	14,831 (30.0)	24,855 (24.4)	21,266 (27.3)	0.084	29.1	26.5	27.0	0.039
COPD[c]	4,247 (8.6)	7,164 (7.0)	5,125 (6.6)	0.05	7.1	6.5	6.2	0.025
Mean diabetes duration (SD), y	10.7 (8.6)	12.2 (7.9)	12.5 (7.6)	0.147	11.8 (8.3)	11.6 (7.9)	11.4 (7.8)	0.034
No. of unique oral diabetes drugs per patient (SD)[c]	1.0 (0.9)	1.5 (0.8)	1.5 (0.8)	0.378	1.2 (0.9)	1.3 (0.8)	1.3 (0.8)	0.092
Diabetic complications,[b] n (%) or %[a]								
Diabetic nephropathy	2,681 (5.4)	5,598 (5.5)	7,803 (10.0)	0.116	7.0	6.5	6.3	0.019
Diabetic retinopathy	6,917 (14.0)	20,909 (20.5)	16,982 (21.8)	0.137	18.3	17.5	16.9	0.024
Diabetic neuropathy	9,591 (19.4)	15,776 (15.5)	10,806 (13.9)	0.099	17.4	17.0	16.5	0.015
Medications,[c] n (%) or %[a]								
Any diabetes drug	30,753 (62.1)	89,762 (88.1)	69,454 (89.2)	0.443	83.0	85.6	84.9	0.012
Insulin	11,255 (22.7)	22,782 (22.4)	11,991 (15.4)	0.125	24.8	22.5	21.8	0.048
Metformin	26,337 (53.2)	80,609 (79.1)	61,847 (79.4)	0.385	76.0	77.1	76.0	0.011
Sulfonylurea	10,987 (22.2)	43,905 (43.1)	39,545 (50.8)	0.411	30.2	29.2	30.0	0.014
TZD	315 (0.6)	927 (0.9)	1,181 (1.5)	0.057	1.0	0.8	0.7	0.019
Statin	1,562 (3.2)	4,412 (4.3)	4,251 (5.5)	0.076	4.1	4.1	4.1	0.001
Glucocorticoid	2,778 (5.6)	4,230 (4.2)	3,526 (4.5)	0.045	4.4	4.3	4.3	0.004
NSAID	3,392 (6.9)	5,487 (5.4)	4,312 (5.5)	0.041	5.1	5.1	5.2	0.005
PPI	14,609 (29.5)	24,848 (24.4)	18,939 (24.3)	0.078	27.2	26.8	26.3	0.014
Loop diuretic	3,727 (7.5)	8,596 (8.4)	6,915 (8.9)	0.033	8.3	7.7	7.4	0.021
Thiazide diuretic	9,670 (19.5)	20,682 (20.3)	16,044 (20.6)	0.018	20.6	20.4	20.0	0.01
Hydralazine	428 (0.9)	1,002 (1.0)	1,414 (1.8)	0.055	1.0	0.9	1.5	0.035
Procainamide	0	0	0	<0.001	0	0	0	<0.001

(Continued)

Table 1. (*Cont'd*)

Variable list	Before matching weights				After matching weights[a]			
	GLP-1RA (n = 49,514)	SGLT2i (n = 101,925)	DPP4i (n = 77,861)	SMD	GLP-1RA (n = 24,560.4)	SGLT2i (n = 24,648.1)	DPP4i (n = 24,495.3)	SMD
Minocycline	293 (0.6)	330 (0.3)	217 (0.3)	0.032	0.5	0.4	0.3	0.027
Phenytoin	34 (0.1)	153 (0.2)	165 (0.2)	0.026	0.1	0.1	0.1	0.011
Propylthiouracil	<10 (0.0)	14 (0.0)	<10 (0.0)	0.002	0	0	0	0.006
Methimazole	96 (0.2)	227 (0.2)	158 (0.2)	0.004	0.1	0.3	0.2	0.019
Mean health care use (SD)[c]								
MSP claims/visits	23.0 (17.1)	20.9 (16.0)	22.2 (20.9)	0.077	22.0 (16.2)	21.3 (17.3)	21.1 (20.6)	0.034
Hospitalizations	0.4 (1.0)	0.4 (0.9)	0.5 (1.1)	0.063	0.4 (1.1)	0.4 (0.9)	0.4 (0.9)	0.02
Hospitalizations due to diabetes	0.4 (0.5)	0.6 (0.5)	0.6 (0.5)	0.268	0.5 (0.5)	0.5 (0.5)	0.5 (0.5)	0.033
Emergency department visits	0.5 (1.6)	0.5 (1.3)	0.6 (1.6)	0.034	0.5 (1.7)	0.5 (1.4)	0.5 (1.3)	0.011

* CHF, congestive heart failure; CKD, chronic kidney disease; COPD, chronic obstructive pulmonary disease; DPP4i, dipeptidyl peptidase 4 inhibitor; GLP-1RA, glucagon-like peptide 1 receptor agonist; IBD, inflammatory bowel disease; IHD, ischemic heart disease; MI, myocardial infarction; MSP, Medical Services Plan; NSAID, nonsteroidal anti-inflammatory drug; PPI, proton pump inhibitor; SGLT2i, sodium-glucose cotransporter 2 inhibitor; SMD, standardized mean difference; TIA, transient ischemic attack; TZD, thiazolidinedione; VTE, venous thromboembolism.
[a] n represents weighted samples sizes and is therefore fractional. Distribution of categorical covariates in weighted cohorts are presented as percentages alone for clarity.
[b] Measured since 1990 or enrollment.
[c] Measured over one year before index date.

vascular events, malignancy, and other autoimmune or inflammatory conditions were also included. These covariates were measured in the 12 months before the index date, except for select chronic conditions, which were captured from enrollment onward. We also controlled for secular trends in prescribing patterns by including calendar year of treatment initiation in the multivariable logistic regression model for calculating PS.

Statistical analysis. We used matching weights (MW), a PS-based approach, to balance baseline covariates across treatment groups.[35] This approach is an extension of inverse probability of treatment weighting (IPTW). Unlike IPTW, individuals are down-weighted to achieve balance in baseline characteristics, avoiding the risk of inflating weights for a small number of patients with good PS overlap. Unlike PS matching, MW retains all patients; outliers are down-weighted rather than excluded. This approach has demonstrated robust performance in settings characterized by rare outcomes, limited covariate overlap, or unequal exposure distributions.[35] Standardized mean differences (SMDs) were calculated before and after MW were applied to assess balance in baseline characteristics.

We calculated incidence rates (IRs) of ARD for each treatment group before and after applying MW. We then fitted Cox proportional hazards models in unweighted and matching-weighted cohorts to estimate hazard ratios (HRs) and corresponding 95% confidence intervals (CIs) that were unadjusted and adjusted for baseline covariates, respectively. We accounted for the competing risk of death using the Fine–Gray subdistribution hazards model, which can inform the direction and size of effects of medication exposures on ARD cumulative incidence (ie, the absolute risk of ARD in the presence of competing risk).[36] Treatment courses were treated as independent in MW and Cox regression models. Because patients could contribute multiple treatment courses, we applied robust SEs clustered by individual. After identifying incident cases of RA, psoriatic disease, axSpA, and SARDs, we calculated frequencies of disease-modifying antirheumatic drug (DMARD) use and DMARD or systemic glucocorticoid use in the 12 months following these diagnoses.

We performed three sensitivity analyses. The first extended the grace period for prescription refills by 30 days (ie, a gap of

Table 2. Incidence of ARDs following initiation of GLP-1RAs or SGLT2i vs DPP4i*

ARD incidence	GLP-1RA (n = 49,514)	SGLT2i (n = 101,925)	DPP4i (n = 77,861)
Incident cases, n	225	355	267
Mean follow-up, y	1.27	1.64	1.81
Unadjusted IR, per 10,000 PY	35.9	21.2	18.9
Matching-weighted IR, per 10,000 PY	29.1	24.4	27.3
Unadjusted HR (95% CI)	1.55 (1.30–1.86)	1.05 (0.89–1.23)	1.0 (ref)
Matching-weighted HR (95% CI)	1.04 (0.81–1.33)	0.93 (0.75–1.16)	1.0 (ref)

* ARD, autoimmune rheumatic disease; CI, confidence interval; DPP4i, dipeptidyl peptidase 4 inhibitor; GLP-1RA, glucagon-like peptide 1 receptor agonist; HR, hazard ratio; IR, incidence rate; PY, person-years; SGLT2i, sodium-glucose cotransporter 2 inhibitor.

up to 90 days between prescriptions was allowed before the exposure medication was considered discontinued). The second continued follow-up for two years after the index date, regardless of whether the medication was continued, to mimic an intention-to-treat analysis. In the third, instead of using MW, we performed 1:1 PS matching to balance baseline covariates between groups. In separate analyses, we matched GLP-1RA initiators to DPP4i initiators and SGLT2i initiators to DPP4i initiators and then estimated risk of ARD with GLP-1RAs or SGLT2i versus DPP4i.

All P values were two sided, and $P < 0.05$ was considered significant for all tests. All statistical analyses were performed with SAS software, version 9.4 (SAS Institute, Inc).

Ethical considerations. No patient partners were involved in the design or conduct of this study. All procedures were performed in compliance with relevant laws and institutional guidelines and have been approved by the University of British Columbia's Behavioral Research Ethics Board (H15-00887). This study exclusively used deidentified administrative data, so the need for informed consent was waived by the ethics committee. All data were fully anonymized before the study team accessed them. This study followed recommendations of the STrengthening the Reporting of OBservational studies in Epidemiology initiative for reporting cohort studies.

RESULTS

The study population included 229,300 adults initiating a new treatment course of either a GLP-1RA (n = 49,514), an SGLT2i (n = 101,925), or a DPP4i (n = 77,861). Of all treatment courses observed, 20% (n = 45,298) were from individuals who

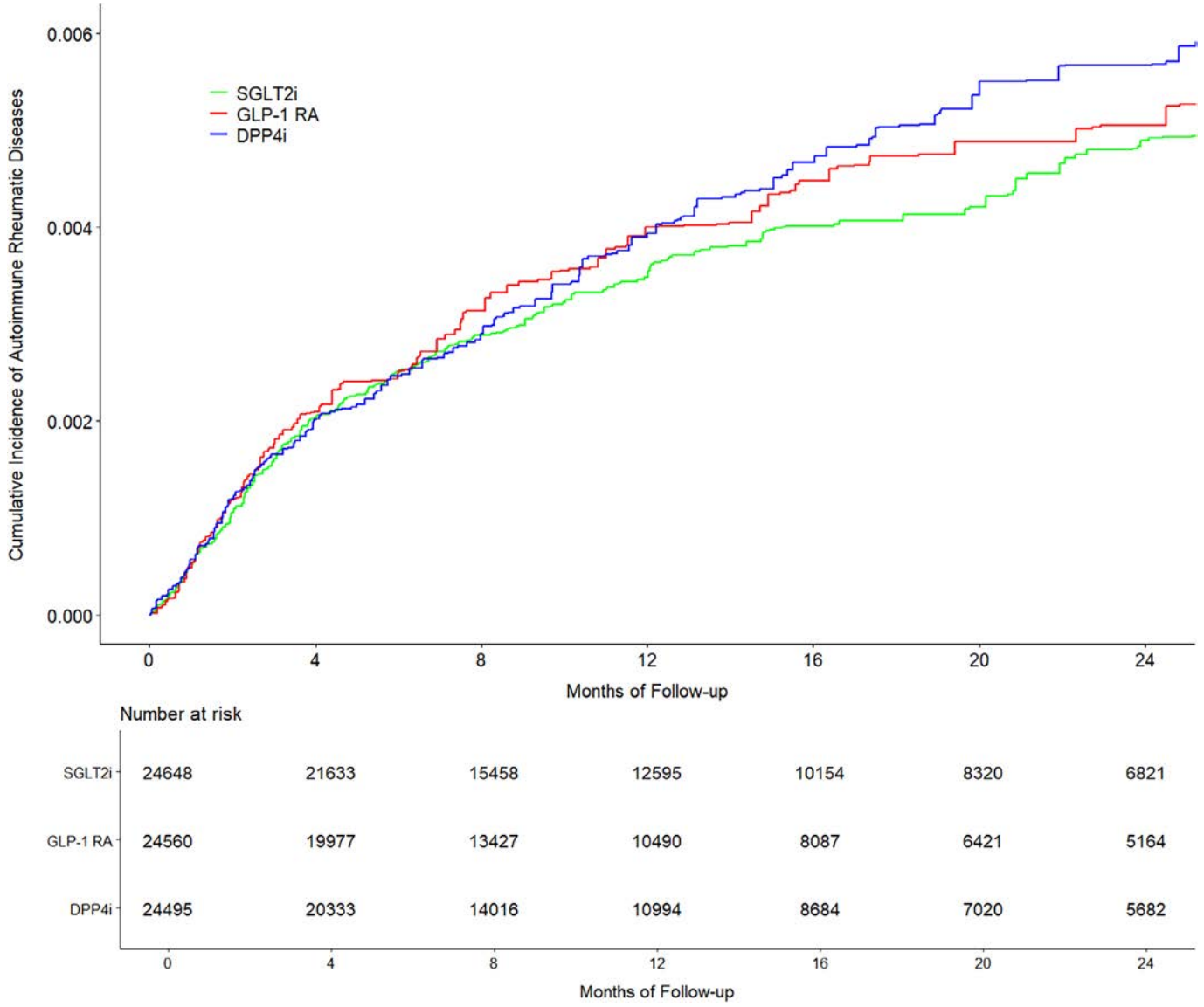


Figure 1. Cumulative incidence of autoimmune rheumatic diseases. DPP4i, dipeptidyl peptidase 4 inhibitor; GLP-1RA, glucagon-like peptide 1 receptor agonist; SGLT2i, sodium-glucose cotransporter 2 inhibitor.

were re-exposed after a >12-month washout period of all three drug classes. Baseline characteristics are summarized in Table 1. Before MW, groups differed in many baseline characteristics. The GLP-1RA cohort, for example, was younger and had a greater proportion of female patients, individuals living in high-income neighborhoods, and people with obesity. After applying MW, covariates were well balanced across all groups (all SMDs < 0.1). The weighted cohorts had mean ages ranging from 59.3 to 59.9 years, and between 46.9% and 48.4% were male. After weighting, mean follow-up duration was 1.3 years for GLP-1RA users, 1.6 years for SGLT2i users, and 1.4 years for DPP4i users.

Overall incidence of ARD. Associations between medication exposures and ARD incidence are presented in Table 2. Among GLP-1RA users, 225 incident cases were identified (unadjusted IR 35.9 per 10,000 person-years), compared with 267 cases among DPP4i users (unadjusted IR 18.9 per 10,000 person-years). The unadjusted HR for ARD risk with GLP-1RA versus DPP4i was 1.55 (95% CI 1.30–1.86); however, after balancing baseline covariates by applying MW, the association was no longer significant (HR 1.04, 95% CI 0.81–1.33).

Among SGLT2i users, 355 incident ARD diagnoses were observed, yielding an unadjusted IR of 21.2 cases per 10,000 person-years. The risk of developing an ARD was not significantly different between SGLT2i and DPP4i users in either unadjusted (HR 1.05, 95% CI 0.89–1.23) or adjusted analyses (HR 0.93, 95% CI 0.75–1.16). Figure 1 illustrates cumulative incidence of ARDs with each exposure.

Results of sensitivity analyses were consistent with the primary findings. Relative to DPP4i, no significant difference in ARD risk was observed with GLP-1RAs or SGLT2i when the exposure window was extended by 30 days (Supplementary Table S3), when follow-up was extended to two years post index date to mimic an intention-to-treat approach (Supplementary Table S4), and when 1:1 PS matching was used instead of MW (Supplementary Tables S5 and S6).

Incidence of individual ARD subtypes. *Psoriatic disease, RA, and axSpA.* Unadjusted and matching-weighted IRs and HRs for psoriatic disease, RA, and axSpA are presented in Table 3. In unadjusted analyses, psoriatic disease risk was significantly higher with both GLP-1RAs (HR 1.93, 95% CI 1.40–2.65) and SGLT2i (HR 1.45, 95% CI 1.09–1.91) relative to DPP4i.

Table 3. Incidence of RA, psoriatic disease, SARDs, and axSpA following initiation of GLP-1RAs or SGLT2i vs DPP4i*

	GLP-1RA (n = 49,514)	SGLT2i (n = 101,925)	DPP4i (n = 77,861)
Psoriatic disease incidence			
Incident cases, n	79	142	80
Mean follow-up, y	1.27	1.64[a]	1.82
Unadjusted IR, per 10,000 PY	12.5	8.4	5.6
Matching-weighted IR, per 10,000 PY	11.1	9.8	8.8
Unadjusted HR (95% CI)	1.93 (1.40–2.65)	1.45 (1.09–1.91)	1.0 (ref)
Matching-weighted HR (95% CI)	1.18 (0.77–1.82)	1.21 (0.84–1.75)	1.0 (ref)
RA incidence			
Incident cases, n	89	139	120
Mean follow-up, y	1.27	1.64[a]	1.82
Unadjusted IR, per 10,000 PY	14.1	8.2	8.4
Matching-weighted IR, per 10,000 PY	11.1	9.1	11.3
Unadjusted HR (95% CI)	1.39 (1.05–1.83)	0.92 (0.72–1.18)	1.0 (ref)
Matching-weighted HR (95% CI)	1.02 (0.69–1.51)	0.76 (0.55–1.05)	1.0 (ref)
SARD incidence			
Incident cases, n	44	51	58
Mean follow-up, years	1.27	1.65[a]	1.82
Unadjusted IR, per 10,000 PY	7.0	3.0	4.0
Matching-weighted IR, per 10,000 PY	5.5	3.1	6.2
Unadjusted HR (95% CI)	1.31 (0.89–1.94)	0.68 (0.47–0.99)	1.0 (ref)
Matching-weighted HR (95% CI)	0.85 (0.49–1.48)	0.51 (0.31–0.84)	1.0 (ref)
axSpA incidence			
Incident cases, n	14	24	14
Mean follow-up, y	1.27	1.65[a]	1.82
Unadjusted IR, per 10,000 PY	2.2	1.4	1.0
Matching-weighted IR, per 10,000 PY	1.8	2.1	1.1
Unadjusted HR (95% CI)	1.77 (0.85–3.71)	1.33 (0.69–2.59)	1.0 (ref)
Matching-weighted HR (95% CI)	1.42 (0.55–3.69)	2.11 (0.82–5.66)	1.0 (ref)

* axSpA, axial spondyloarthritis; CI, confidence interval; DPP4i, dipeptidyl peptidase 4 inhibitor; GLP-1RA, glucagon-like peptide 1 receptor agonist; HR, hazard ratio; IR, incidence rate; PY, person-years; RA, rheumatoid arthritis; SARD, systemic autoimmune rheumatic disease; SGLT2i, sodium-glucose cotransporter 2 inhibitor.
[a] Slight differences in mean follow-up time reflect outcome-specific censoring.

Significant differences were no longer present after balancing baseline covariates with MW. Cumulative incidence functions for RA and psoriatic disease, stratified by drug class, are presented in Supplementary Figures S2 and S3, respectively. In the 12 months following their diagnosis, 55% of people with RA, 32% of people with psoriatic disease, and 21% of people with axSpA received conventional synthetic, biologic, or targeted synthetic DMARDs (Supplementary Table S7).

SARDs. We identified 153 incident SARDs, and 60 (39%) met the case definition with consecutive 710 ICD-9 codes. Because this three-digit code represents the category of diffuse connective tissue diseases (including SLE, SSc, Sjögren disease, idiopathic inflammatory myopathies, mixed connective tissue disease, and undifferentiated connective tissue disease), a specific SARD was not identifiable in these cases. Supplementary Table S8 summarizes the frequencies of specified SARDs based on ICD-10 and/or four-digit ICD-9 codes. Among individuals newly diagnosed with SARDs, 103 (67%) received DMARDs or glucocorticoids in the 12 months following their diagnosis (Supplementary Table S7).

There were 44, 51, and 58 incident cases of SARDs among new users of GLP-1RAs, SGLT2i, and DPP4i, respectively, corresponding to unadjusted IRs of 7.0, 3.0, and 4.0 per 10,000 person-years (Table 3). After weighting, IRs were 5.5, 3.1, and 6.2 per 10,000 person-years, respectively. Cumulative incidence of SARDs with SGLT2i, GLP-1RAs, and DPP4i is displayed in Figure 2. Weighted analysis showed a significantly lower risk of SARDs with SGLT2i relative to DPP4i exposure (HR 0.51, 95%

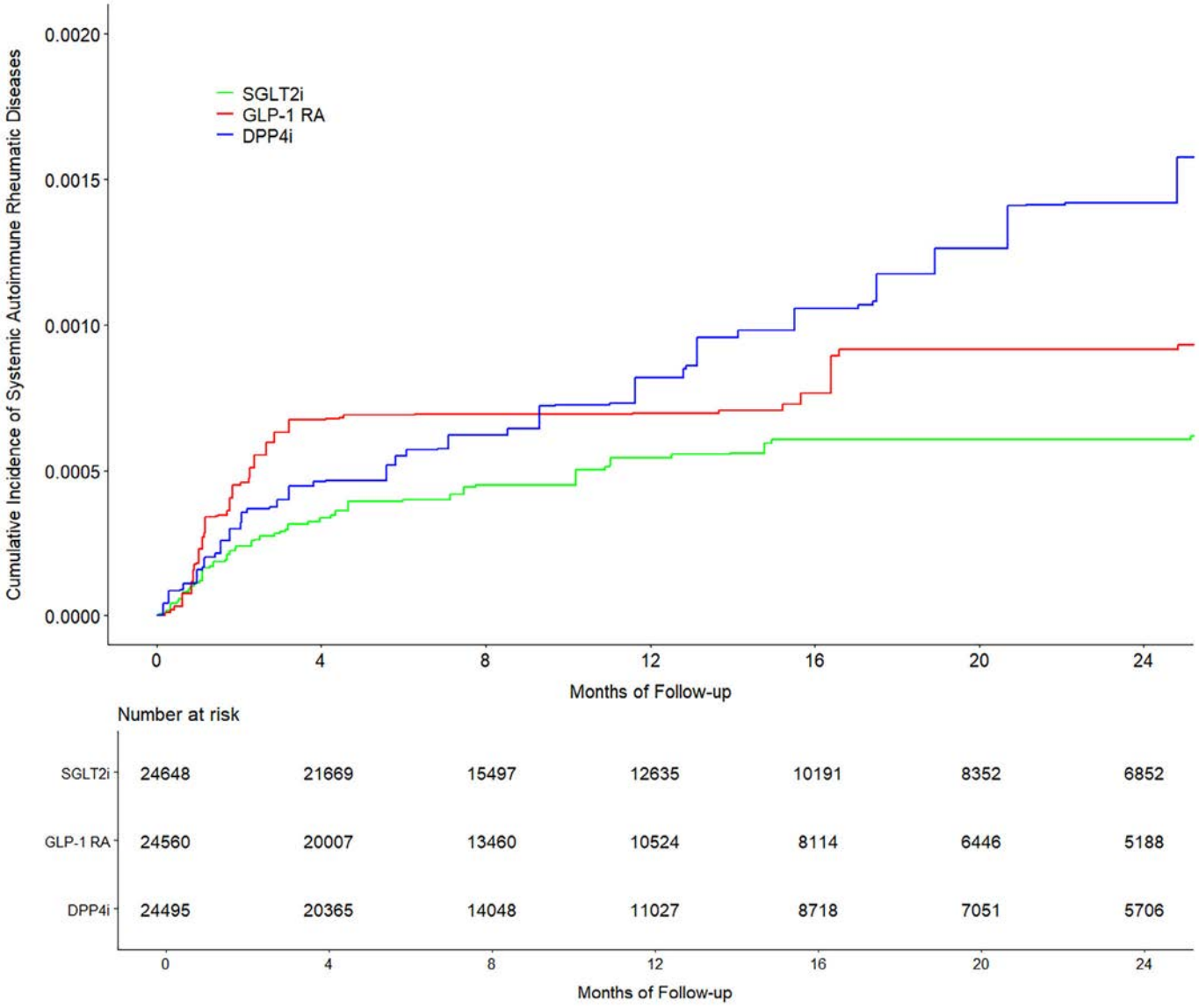


Figure 2. Cumulative incidence of systemic autoimmune rheumatic diseases. DPP4i, dipeptidyl peptidase 4 inhibitor; GLP-1RA, glucagon-like peptide 1 receptor agonist; SGLT2i, sodium-glucose cotransporter-2 inhibitor. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.70044/abstract.

CI 0.31–0.84). No significant difference was observed between GLP-1RA and DPP4i users. The risk difference between SGLT2i and DPP4i exposure was –3.0 cases per 10,000 person-years (95% CI –10.3 to 4.3).

DISCUSSION

In this large population-based cohort study of over 200,000 adults with T2D, we observed no difference in the risk of developing psoriatic disease, RA, axSpA, or ARDs overall among those newly treated with GLP-1RAs or SGLT2i relative to DPP4i. However, initiating SGLT2i was associated with a nearly 50% lower risk of incident SARDs relative to DPP4i. Although unadjusted analyses suggested an increased risk of psoriatic disease, RA, and ARDs overall with GLP-1RA exposure, these associations were no longer evident after controlling for baseline imbalances by applying MW, suggesting that the initial differences were attributable to confounding.

These findings contribute to a limited but growing body of literature evaluating ARD risk following GLP-1RA or SGLT2i initiation. To our knowledge, this is the first study using an active comparator design to examine the association between GLP-1RA use and incident ARDs. Prior work, including a retrospective cohort study using TriNetX data, compared GLP-1RA users to PS-matched controls without GLP-1RA exposure.[11] The study reported reduced risks of SLE, SSc, and RA in individuals with obesity treated with GLP-1RAs, as well as reduced SLE and SSc (but not RA) risk in GLP-1RA–treated patients with T2D. However, the absence of an active comparator and new-user design raises concerns for immortal time bias. Because the index date was defined as the time an individual first met inclusion criteria, the period between the index date and GLP-1RA initiation was "immortal" time. An ARD diagnosis in this time could not occur for members of the GLP-1RA cohort because cohort entry required remaining ARD-free until receipt of a GLP-1RA. Additionally, key confounders, including comorbidities, concomitant medications, and health care use were not accounted for in the PS estimation.

Our study addressed these methodologic limitations through a new-user design with an active comparator and application of MW to ensure balance across groups at treatment initiation. We built on work by Nassar et al[11] by calculating incidence density (accounting for time taking treatment) as opposed to incidence proportions, considering competing risk of death, including a 10-year look-back window to exclude prevalent ARD cases, and using validated ARD case definitions. Considering the age and elevated ARD risk in our cohort, the IRs we observed are consistent with previously reported estimates using rigorous case ascertainment methods.[37–41]

The absence of a protective association between GLP-1RAs and ARDs in our study may reflect differences in study population and follow-up time. Our sample was older, we did not restrict to individuals with obesity, and mean follow-up on treatment was less than two years. Evidence from Mendelian randomization studies and meta-analyses supports a causal link between elevated body mass index and increased risk of RA and psoriatic disease.[8–10,42] Because many ARDs have long preclinical phases and our mean follow-up period was relatively short (1.3–1.6 years), a Type II error is possible. A longer duration of GLP-1RA exposure may be needed to observe risk modification, particularly among patients with obesity. Nonetheless, our findings are reassuring in demonstrating no increased ARD risk associated with GLP-1RA use in individuals with T2D. If GLP-1RAs or SGLT2i triggered ARD development, we would have expected to see this within one year of exposure based on prior case reports.[12,13]

In contrast, we found that SGLT2i exposure was associated with significantly reduced risk of developing SARDs. To our knowledge, this is the first study to report a potential protective association between SGLT2i and SARD incidence. Previous observational studies have evaluated the risk of psoriasis with SGLT2i use, with two studies finding no association and one reporting an 8% increased risk compared to DPP4i.[14–16] Our findings align with two of three studies. We found no prior studies assessing the incidence of other ARDs following SGLT2i initiation.

Biologic plausibility of the association observed is suggested by emerging clinical data supporting therapeutic benefits of SGLT2i in SLE[43,44] and preclinical experiments providing mechanistic evidence of immunomodulatory effects.[45] In lupus-prone mice, empagliflozin treatment led to (1) decreased anti–double-stranded DNA levels and expression of the NLRP3 inflammasome; (2) down-regulated gene sets related to complement activation, inflammation, and oxidative damage; and (3) reduced mechanistic target of rapamycin complex 1 activity, enhancing autophagy and reducing podocyte injury.[45] SGLT2i have also been shown to modulate macrophage polarization from proinflammatory (M1) to anti-inflammatory (M2) phenotypes and reduce release of key proinflammatory cytokines (eg, interleukin-1β [IL-1β], IL-6, tumor necrosis factor) through inhibition of the NF-κB pathway and down-regulation of toll-like receptor 4.[46,47] The immunomodulatory effects of SGLT2i may explain the lower risk of SARDs observed in our study. Although we observed a 49% reduction in risk of SARDs with SGLT2i, because SARDs are relatively rare diagnoses, the absolute risk difference was small, and the number needed to treat was 3,334. Furthermore, population-level data were obtained from an ethnically diverse Canadian province with universal health care. The clinical implications of our findings should be considered with these factors in mind, and further research is needed to explore possible preventive effects in different populations, including in high-risk groups with greater potential for benefit, and across settings with different patient characteristics and prescribing patterns.

This study has limitations. Short follow-up duration may have limited our ability to detect reductions in risk, especially for conditions such as RA or psoriatic disease. Additionally, the relatively

low number of axSpA cases, likely due to the older age of our cohort, limited our ability to assess this outcome and resulted in imprecise estimates. Despite careful adjustment using MW to balance baseline covariates across treatment groups, residual and unmeasured confounding remains possible. For instance, obesity might have been underreported because ICD codes for obesity have low sensitivity.[48] Any bias from this would likely be conservative, given the greater likelihood of GLP-1RA or SGLT2i prescription in patients with obesity. We did not have access to data on body mass index. Whether excess adiposity modifies the effects of GLP-1RAs and SGLT2i on ARD risk is an important question that requires further investigation. Smoking status, a risk factor for several ARDs, was not reliably captured in our data. We used chronic obstructive pulmonary disease (COPD) as a proxy, and COPD status did not significantly differ between groups before or after adjustment. Before MW, COPD was numerically most common in new users of GLP-1RAs. Whether smoking was also more common in this group is unknown. This, like underreporting obesity, would bias against detecting a protective effect of GLP-1RAs if one truly existed.

ARD misclassification cannot be fully excluded, though we relied on strict, validated case definitions with high specificity and moderate-to-high positive predictive value.[29–33] We assumed that a person who filled a prescription took the medication that was dispensed, and this may overestimate true exposure duration in some patients. However, we suspect that most overestimates would be small because we defined exposure duration using the course of the medication dispensed rather than the course prescribed. This is a strength, as repeatedly filling prescriptions and paying dispensing fees increases the likelihood that a person is taking the medication that was prescribed to them.

It is also possible that DPP4i reduce ARD risk relative to no treatment, which could bias estimates toward the null hypothesis. Some observational studies, conducted before the widespread use of GLP-1RAs and SGLT2i in T2D, have reported protective associations between DPP4i use and various ARDs.[18,19,49] However, these findings have not been consistently replicated.[17,20,50] Given the conflicting evidence, we adopted a conservative interpretation of a neutral effect of DPP4i.

Strengths of our study include the new-user design, with an active comparator, as well as the use of MW to account for differences in baseline characteristics affecting ARD risk and modeling to account for exposure time and competing risk of death. Together, these reduce risk of bias. The population-based sample increases generalizability, and the results of prespecified sensitivity analyses enhance the robustness of our findings.

In conclusion, among adults with T2D, treatment with GLP-1RAs or SGLT2i was not associated with increased risk of developing ARDs relative to DPP4i. Notably, SGLT2i use was associated with a nearly 50% lower risk of SARDs relative to DPP4i use. These findings support the safety of these agents with respect to ARD development. Additional studies are needed to reproduce these findings and further explore potential disease-modifying mechanisms of SGLT2i in SARDs.

ACKNOWLEDGMENTS

ChatGPT models 4o and o3 (OpenAI) were used only to check grammar and suggest minor changes to improve readability. The authors take full responsibility for the content of the published article.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr Aviña-Zubieta confirms that all authors have provided the final approval of the version to be published and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. International Diabetes Federation. IDF diabetes atlas. 2021. Accessed March 25, 2025. https://www.ncbi.nlm.nih.gov/books/NBK581934/
2. Carney G, Kim JD, O'Sullivan C, et al. Treatment pattern trends of medications for type 2 diabetes in British Columbia, Canada. BMJ Open Diabetes Res Care 2022;10(6):e002995. https://doi.org/10.1136/bmjdrc-2022-002995
3. Badve SV, Bilal A, Lee MMY, et al. Effects of GLP-1 receptor agonists on kidney and cardiovascular disease outcomes: a meta-analysis of randomised controlled trials. Lancet Diabetes Endocrinol 2025;13(1):15–28. https://doi.org/10.1016/S2213-8587(24)00271-7
4. Zinman B, Wanner C, Lachin JM, et al; EMPA-REG OUTCOME Investigators. Empagliflozin, cardiovascular outcomes, and mortality in type 2 diabetes. N Engl J Med 2015;373(22):2117–2128. https://doi.org/10.1056/NEJMoa1504720
5. Lincoff AM, Brown-Frandsen K, Colhoun HM, et al; SELECT Trial Investigators. Semaglutide and cardiovascular outcomes in obesity without diabetes. N Engl J Med 2023;389(24):2221–2232. https://doi.org/10.1056/NEJMoa2307563
6. Karacabeyli D, Lacaille D. Glucagon-like peptide-1 receptor agonists in arthritis: current insights and future directions. Nat Rev Rheumatol 2025;21(11):671–683. https://doi.org/10.1038/s41584-025-01302-0
7. Karacabeyli D, Lacaille D. Effects of weight-lowering medications in people with obesity and rheumatic diseases. Curr Treatm Opt Rheumatol 2025;11(1):12. https://doi.org/10.1007/s40674-025-00232-y
8. Xie W, Huang H, Deng X, et al. Modifiable lifestyle and environmental factors associated with onset of psoriatic arthritis in patients with psoriasis: a systematic review and meta-analysis of observational studies. J Am Acad Dermatol 2021;84(3):701–711. https://doi.org/10.1016/j.jaad.2020.08.060
9. Ohno T, Aune D, Heath AK. Adiposity and the risk of rheumatoid arthritis: a systematic review and meta-analysis of cohort studies. Sci Rep 2020;10(1):16006. https://doi.org/10.1038/s41598-020-71676-6
10. Karlsson T, Hadizadeh F, Rask-Andersen M, et al. Body mass index and the risk of rheumatic disease: linear and nonlinear Mendelian randomization analyses. Arthritis Rheumatol 2023;75(11):2027–2035. https://doi.org/10.1002/art.42613

11. Nassar M, Nassar O, Abosheaishaa H, et al. Comparative outcomes of systemic diseases in people with type 2 diabetes, or obesity alone treated with and without GLP-1 receptor agonists: a retrospective cohort study from the Global Collaborative Network. J Endocrinol Invest 2025;48(2):483–497. https://doi.org/10.1007/s40618-024-02466-4
12. Kikkawa K, Hoshi H, Isoda A, et al. Long-acting glucagon-like peptide-1 receptor agonist-induced rheumatoid arthritis in a patient with type 2 diabetes mellitus. Dubai Diabetes Endocrinol J 2021;27(3):114–117. https://doi.org/10.1159/000519008
13. Ambrosio ML, Monami M, Sati L, et al. GLP-1 receptor agonist-induced polyarthritis: a case report. Acta Diabetol 2014;51(4):673–674.
14. Ma SH, Wu CY, Lyu YS, et al. Association between sodium-glucose co-transporter 2 inhibitors and risk of psoriasis in patients with diabetes mellitus: a nationwide population-based cohort study. Clin Exp Dermatol 2022;47(12):2242–2250. https://doi.org/10.1111/ced.15385
15. Shen CH, Yang YS, Lee ITL, et al. Sodium-glucose cotransporter-2 inhibitors in diabetes and inflammatory skin diseases. Br J Dermatol 2024;191(3):467–468. https://doi.org/10.1093/bjd/ljae229
16. Ma KSK, Tsai SYC, Holt A, et al. Effects of sodium-glucose cotransporter-2 inhibitors on inflammatory skin diseases in patients with type 2 diabetes. J Am Acad Dermatol 2024;91(5):934–936. https://doi.org/10.1016/j.jaad.2024.04.079
17. Zemedikun DT, Gokhale K, Chandan JS, et al. Type 2 diabetes mellitus, glycaemic control, associated therapies and risk of rheumatoid arthritis: a retrospective cohort study. Rheumatology (Oxford) 2021;60(12):5567–5575. https://doi.org/10.1093/rheumatology/keab148
18. Seong JM, Yee J, Gwak HS. Dipeptidyl peptidase-4 inhibitors lower the risk of autoimmune disease in patients with type 2 diabetes mellitus: a nationwide population-based cohort study. Br J Clin Pharmacol 2019;85(8):1719–1727. https://doi.org/10.1111/bcp.13955
19. Chen YC, Chen TH, Sun CC, et al. Dipeptidyl peptidase-4 inhibitors and the risks of autoimmune diseases in type 2 diabetes mellitus patients in Taiwan: a nationwide population-based cohort study. Acta Diabetol 2020;57(10):1181–1192. https://doi.org/10.1007/s00592-020-01533-5
20. Kridin K, Amber K, Khamaisi M, et al. Is there an association between dipeptidyl peptidase-4 inhibitors and autoimmune disease? A population-based study. Immunol Res 2018;66(3):425–430. https://doi.org/10.1007/s12026-018-9005-8
21. Shah BR, Bajaj HS, Butalia S, et al; Diabetes Canada Clinical Practice Guidelines Expert Working Group. Pharmacologic glycemic management of type 2 diabetes in adults—2024 update. Can J Diabetes 2024;48(7):415–424. https://doi.org/10.1016/j.jcjd.2024.08.002
22. Government of British Columbia. Population estimates. 2025. Updated February 9, 2025. Accessed March 25, 2025. https://www2.gov.bc.ca/gov/content/data/statistics/people-population-community/population/population-estimates
23. Discharge Abstracts Database. Hospital separations. Population Data BC. 2025. Accessed February 19, 2025. https://www.popdata.bc.ca/data/health/dad
24. Medical Services Plan. 2025. Accessed February 19, 2025. https://www.popdata.bc.ca/data/health/msp
25. National Ambulatory Care Reporting System. Population Data BC. 2025. Accessed February 19, 2025. https://www.popdata.bc.ca/data
26. Vital Events Deaths. Population Data BC. Accessed February 19, 2025. https://www.popdata.bc.ca/data/demographic/vs_deaths
27. PharmaNet. Population Data BC. 2025. Accessed February 19, 2025. https://www.popdata.bc.ca/data/health/pharmanet
28. Leong A, Dasgupta K, Bernatsky S, et al. Systematic review and meta-analysis of validation studies on a diabetes case definition from health administrative records. PLoS One 2013;8(10):e75256. https://doi.org/10.1371/journal.pone.0075256
29. Chang J, Rogers P, Lacaille D. Can American College of Rheumatology criteria for rheumatoid arthritis be assessed using self-report data?—Comparison of self-reported data with chart review [abstract]. Arthritis Rheum 2011;63(suppl 10):S49.
30. Asgari MM, Wu JJ, Gelfand JM, et al. Validity of diagnostic codes and prevalence of psoriasis and psoriatic arthritis in a managed care population, 1996-2009. Pharmacoepidemiol Drug Saf 2013;22(8):842–849. https://doi.org/10.1002/pds.3447
31. Aviña-Zubieta JA, Chan J, De Vera M, et al. Risk of venous thromboembolism in ankylosing spondylitis: a general population-based study. Ann Rheum Dis 2019;78(4):480–485. https://doi.org/10.1136/annrheumdis-2018-214388
32. Arkema EV, Jönsen A, Rönnblom L, et al. Case definitions in Swedish register data to identify systemic lupus erythematosus. BMJ Open 2016;6(1):e007769. https://doi.org/10.1136/bmjopen-2015-007769
33. Bernatsky S, Linehan T, Hanly JG. The accuracy of administrative data diagnoses of systemic autoimmune rheumatic diseases. J Rheumatol 2011;38(8):1612–1616. https://doi.org/10.3899/jrheum.101149
34. Romano PS, Roos LL, Jollis JG. Adapting a clinical comorbidity index for use with ICD-9-CM administrative data: differing perspectives. J Clin Epidemiol 1993;46(10):1075–1079. https://doi.org/10.1016/0895-4356(93)90103-8
35. Yoshida K, Hernández-Díaz S, Solomon DH, et al. Matching weights to simultaneously compare three treatment groups: comparison to three-way matching. Epidemiology 2017;28(3):387–395. https://doi.org/10.1097/EDE.0000000000000627
36. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc 1999;94(446):496–509. https://doi.org/10.1080/01621459.1999.10474144
37. Myasoedova E, Crowson CS, Kremers HM, et al. Is the incidence of rheumatoid arthritis rising?: results from Olmsted County, Minnesota, 1955-2007. Arthritis Rheum 2010;62(6):1576–1582. https://doi.org/10.1002/art.27425
38. Icen M, Crowson CS, McEvoy MT, et al. Trends in incidence of adult-onset psoriasis over three decades: a population-based study. J Am Acad Dermatol 2009;60(3):394–401. https://doi.org/10.1016/j.jaad.2008.10.062
39. Wright KA, Crowson CS, Michet CJ, et al. Time trends in incidence, clinical features, and cardiovascular disease in ankylosing spondylitis over three decades: a population-based study. Arthritis Care Res (Hoboken) 2015;67(6):836–841. https://doi.org/10.1002/acr.22512
40. Pons-Estel GJ, Alarcón GS, Scofield L, et al. Understanding the epidemiology and progression of systemic lupus erythematosus. Semin Arthritis Rheum 2010;39(4):257–268. https://doi.org/10.1016/j.semarthrit.2008.10.007
41. Garvey TD, Koster MJ, Crowson CS, et al. Incidence, survival, and diagnostic trends in GCA across seven decades in a North American population-based cohort. Semin Arthritis Rheum 2021;51(6):1193–1199. https://doi.org/10.1016/j.semarthrit.2021.09.006
42. Budu-Aggrey A, Brumpton B, Tyrrell J, et al. Evidence of a causal relationship between body mass index and psoriasis: a Mendelian randomization study. PLoS Med 2019;16(1):e1002739. https://doi.org/10.1371/journal.pmed.1002739
43. Pope J, Karacabeyli D, Aviña-Zubieta JA. Target trial emulations of sodium-glucose cotransporter 2 inhibitors in systemic lupus erythematosus. Arthritis Rheumatol 2025;77(4):390–392. https://doi.org/10.1002/art.43130
44. Ma KS, Lo JE, Kyttaris VC, et al. Efficacy and safety of sodium-glucose co-transporter 2 inhibitors for the primary prevention of cardiovascular, renal events and safety outcomes in patients with

systemic lupus erythematosus and comorbid type 2 diabetes: a population-based target trial emulation. Arthritis Rheumatol 2025; 77(4):414–422. https://doi.org/10.1002/art.43037

45. Zhao XY, Li SS, He YX, et al. SGLT2 inhibitors alleviated podocyte damage in lupus nephritis by decreasing inflammation and enhancing autophagy. Ann Rheum Dis 2023;82(10):1328–1340. https://doi.org/10.1136/ard-2023-224242
46. Abdollahi E, Keyhanfar F, Delbandi AA, et al. Dapagliflozin exerts anti-inflammatory effects via inhibition of LPS-induced TLR-4 overexpression and NF-κB activation in human endothelial cells and differentiated macrophages. Eur J Pharmacol 2022;918:174715. https://doi.org/10.1016/j.ejphar.2021.174715
47. Feijóo-Bandín S, Aragón-Herrera A, Otero-Santiago M, et al. Role of sodium-glucose co-transporter 2 inhibitors in the regulation of inflammatory processes in animal models. Int J Mol Sci 2022;23(10): 5634. https://doi.org/10.3390/ijms23105634
48. Suissa K, Schneeweiss S, Lin KJ, et al. Validation of obesity-related diagnosis codes in claims data. Diabetes Obes Metab 2021;23(12): 2623–2631. https://doi.org/10.1111/dom.14512
49. Kim SC, Schneeweiss S, Glynn RJ, et al. Dipeptidyl peptidase-4 inhibitors in type 2 diabetes may reduce the risk of autoimmune diseases: a population-based cohort study. Ann Rheum Dis 2015;74(11):1968–1975. https://doi.org/10.1136/annrheumdis-2014-205216
50. Douros A, Abrahami D, Yin H, et al. Use of dipeptidyl peptidase-4 inhibitors and new-onset rheumatoid arthritis in patients with type 2 diabetes. Epidemiology 2018;29(6):904–912. https://doi.org/10.1097/EDE.0000000000000891

Arthritis & Rheumatology
Vol. 78, No. 3, March 2026, pp 665–671
DOI 10.1002/art.43403

American College of Rheumatology
Empowering Rheumatology Professionals

BRIEF REPORT

Glucagon-Like Peptide-1 Receptor Agonist Use and the Risk of Adverse Cardiac and Kidney Outcomes Among Patients With Systemic Lupus Erythematosus and Lupus Nephritis

April Jorge,[1,2] Aakash V. Patel,[1,2] Baijun Zhou,[1] Lingxiao Zhang,[1] and Hyon Choi[1,2]

Objective. Glucagon-like peptide-1 receptor agonists (GLP-1RAs) have cardioprotective and kidney-protective benefits among patients with type 2 diabetes (T2D). We sought to determine whether GLP-1RA use improves cardiovascular (CV) and kidney outcomes among patients with systemic lupus erythematosus (SLE) and lupus nephritis (LN).

Methods. We emulated a pragmatic target trial to evaluate the impact of GLP-1RA versus comparator hypoglycemic agents, dipeptidyl peptidase 4 inhibitors (DPP4i), on CV and kidney outcomes among patients with SLE and T2D using a large, US multicenter electronic health record database. We used propensity score overlap weighting to emulate randomization among treatment groups. Outcomes included major adverse CV events, venous thromboembolism (VTE), kidney disease progression (estimated glomerular filtration rate decline ≥ 30% or new onset end-stage kidney disease), and all-cause mortality. We used Cox regression to compare hazard ratios (HR) based on the weighted populations. In a secondary analysis, we only included patients with LN.

Results. There were 910 and 1,004 initiators of GLP-1RA and DPP4i, respectively, including 267 and 324 patients with LN, respectively. Baseline covariates were balanced after propensity score overlap weighting. The risks of major adverse cardiac events (MACEs; hazard ratio [HR] 0.66, 95% confidence interval [CI] 0.48–0.91), VTE (HR 0.49, 95% CI 0.24–0.97), kidney disease progression (HR 0.77, 95% CI 0.60–0.98), and all-cause mortality (HR 0.26, 95% CI 0.10–0.68) were lower with GLP-1RA versus DPP4i use. GLP-1RA use was similarly associated with lower risks of MACE and kidney disease progression among patients with LN.

Conclusion. We found lower risks of adverse CV and kidney outcomes and mortality with GLP-1RA use compared with DPP4i use among patients with lupus and T2D.

INTRODUCTION

Systemic lupus erythematosus (SLE) is a heterogenous autoimmune disease that disproportionately impacts women of childbearing age and racial and ethnic minority groups, and it is associated with serious morbidity and premature mortality.[1] Lupus nephritis (LN) is a severe manifestation of SLE that occurs in up to 50% of patients, placing them at significantly increased risk for chronic kidney disease (CKD) and cardiometabolic disease.[1,2] Despite recent advances in treatment, patients with LN have a 10% to 30% risk of developing end-stage kidney disease (ESKD). Furthermore, atherosclerotic cardiovascular (CV) disease is the most common cause of death among patients with SLE, who have over a two-fold higher incidence of myocardial infarction and cerebrovascular accidents than those without SLE. These adverse outcomes are driven by a combination of uncontrolled lupus disease activity and chronic tissue damage including proteinuric nephropathy and treatment-related

Dr Jorge's work was supported by the National Institutes of Arthritis and Musculoskeletal and Skin Diseases (NIAMS), NIH (grant K23-AR-079040), and Rheumatology Research Foundation K Supplement. Dr Choi's work was supported by the NIAMS, NIH (grants P50-AR-060772 and R01-AR-065944).

[1]Division of Rheumatology, Allergy, and Immunology, Department of Medicine, Massachusetts General Hospital, Boston; [2]Harvard Medical School, Boston, Massachusetts.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/art.43403).

Author disclosures and graphical abstract are available at https://onlinelibrary.wiley.com/doi/10.1002/art.43403.

Address correspondence via email to April Jorge, MD, at amjorge@mgh.harvard.edu.

Submitted for publication March 30, 2025; accepted in revised form September 10, 2025.

glucocorticoid toxicity, which exacerbates conventional CV risk factors. Additional strategies are needed to mitigate the profound morbidity associated with CV disease and kidney disease progression.

Glucagon-like peptide-1 receptor agonists (GLP-1RAs) are a class of antiobesity and hypoglycemic agents that have been found to prevent major adverse cardiac events (MACEs) in patients with overweight or obesity and type 2 diabetes (T2D) and to reduce CKD progression in patients with diabetic nephropathy.[3–5] The emerging evidence of nephroprotective benefits has demonstrated proteinuria reduction and a reduction in estimated glomerular filtration rate (eGFR) decline with GLP-1RA use.[5,6] We postulated that GLP-1RAs may have similar cardioprotective and nephroprotective effects among patients with SLE and LN.

We used a large, multicenter US electronic health records (EHRs) database to emulate a pragmatic target trial[7] to evaluate whether GLP-1RAs improve kidney and CV outcomes in patients with SLE and T2D, compared with dipeptidyl peptidase 4 inhibitors (DPP4i), a comparator class of hypoglycemic agents. DPP4i have been found to have a neutral impact on the risk of CV events and have not demonstrated nephroprotective benefits.[8]

PATIENTS AND METHODS

Data source and study design. We identified patients ≥18 years of age who met administrative definition criteria for SLE and T2D in TriNetX, a federated EHR database consisting of patients from academic medical centers, community hospitals, and outpatient clinics in the United States. This database includes demographics, biometrics, medications, laboratory tests, procedures, clinical encounters, diagnostic and procedure codes, and vital status, as previously described.[9] Patients were identified as having SLE based on the presence of at least two corresponding *International Classification of Diseases Ninth Revision* (*ICD-9*) or *International Classification of Diseases Tenth Revision* (*ICD-10*) codes (*ICD-9* 710.9 or *ICD-10* M32.* excluding M32.0) at least 30 days, but no more than two years, apart. Patients with SLE were further characterized as having LN if they had one or more *ICD-10* code for LN (M32.14 or M32.15) or two or more nephritis-related codes (*ICD-9* 580–586 and 791.0; *ICD-10* N00, N03, N04, N05, N17, N18, and R80.9). We specified and emulated the protocol of a pragmatic target trial to compare the impact of GLP-1RA and DPP4i use on CV and kidney outcomes among patients with SLE and T2D and among patients with LN and T2D (Supplementary Table 1).[7] This study was approved by the Mass General Brigham Institutional Review Board, and informed consent was waived.

Eligibility criteria and treatment assignment. We identified patients with SLE who initiated a GLP-1RA or DPP4i between October 2006, when the first DPP4i was approved, and August 2021, the end of available follow-up data. GLP-1RAs included exenatide, albiglutide, dulaglutide, liraglutide, and semaglutide, and DPP4is included sitagliptin, saxagliptin, linagliptin, and alogliptin. Of note, the GLP-1RA albiglutide was discontinued in 2017, and tirzepatide, a glucose-dependent insulinotropic polypeptide and GLP-1RA, was not included because it was approved by the US Food and Drug Administration in 2022. During this study period, GLP-1RAs or DPP4is were only approved by the US Food and Drug Administration for the treatment of T2D, and we excluded patients without a diagnosis of T2D. Patients were excluded from the study if they used the comparator medication or a sodium-glucose cotransporter-2 inhibitor in the year before the index date of treatment initiation, had a diagnosis of ESKD or type 1 diabetes before the index date, or had missing covariate data, except for hemoglobin A1C (HgA1C), which was imputed.

We classified individuals into treatment groups, initiation of GLP-1RA or DPP4i, according to the strategy that their data were compatible with at baseline. We used propensity score overlap weighting to emulate randomization of treatment assignment and balance potential confounders.

Covariates. Covariates were assessed in the 12 months before the index date and included demographics; geographic region; comorbidities including CKD, heart failure, CV disease, tobacco use, obesity, and the Charlson comorbidity index[10]; medications for SLE and T2D; and health care use, including outpatient visits and emergency room or inpatient encounters. We also assessed the presence of LN, SLE severity assessed using a severity index modified from an administrative algorithm based on *ICD* codes and medication use,[9,11] diabetic complications, and HgA1C. We used multiple imputations for missing values of HgA1C.

Outcomes. Our primary outcomes included MACE and kidney disease progression. MACE was defined as a composite of ischemic stroke, myocardial infarction, heart failure, and CV death. MACE were identified by *ICD* codes occurring during inpatient encounters. The definition of CV death was adapted from a validated claims-based algorithm and included death in the month of an inpatient encounter for CV disease as well as out of hospital death without a diagnosis of cancer in the 365 days before and including the date of death and without diagnoses of infection or trauma in the 45 days before and including the date of death.[12] Kidney disease progression was defined as a decrease in eGFR by >30% or new onset ESKD. We also assessed the individual components of MACE, venous thromboembolism (VTE, including pulmonary embolism or deep vein thrombosis), all-cause death, and a control outcome of genital infection, which was expected to be null.

Statistical analysis. Descriptive statistics of baseline characteristics between the two treatment groups were performed

before and after propensity score overlap weighting and compared using standardized mean differences. We conducted a primary as treated analysis in which patients were observed from the index date of treatment initiation until the outcome of interest, disenrollment from the database, death, discontinuation of assigned treatment, initiation of the comparator treatment, or the end of the study period (August 2021). We conducted a secondary intention-to-treat analysis, in which patients were

Table 1. Baseline characteristics of GLP-1RA and DPP4i initiators with SLE and diabetes*

	Before propensity score overlap weighting			After propensity score overlap weighting		
	GLP-1RA	DPP4i	Std. Diff.	GLP-1RA	DPP4i	Std. Diff.
N	910	1,004				
Age, mean (SD)	53.0 (11.3)	58.2 (12.0)	0.45	55.0	55.0	<0.01
Female, n (%)	841 (92.4)	918 (91.4)	0.04	92.3	92.3	<0.01
Race and ethnicity, n (%)			0.12			<0.01
Asian	6 (0.7)	19 (1.9)		0.9	0.9	
Hispanic	71 (7.8)	104 (10.4)		9.0	9.0	
Black	332 (36.5)	335 (33.4)		34.7	34.7	
Other	63 (6.9)	67 (6.7)		7.2	7.2	
White, non-Hispanic	438 (48.1)	479 (47.7)		48.2	48.2	
Geographic region, n (%)			0.08			<0.01
East	202 (22.2)	200 (19.9)		20.7	20.7	
Midwest	168 (18.5)	159 (15.8)		17.0	17.0	
South	461 (50.7)	546 (54.4)		53.5	53.5	
West	79 (8.7)	99 (9.9)		8.9	8.9	
Lupus nephritis, n (%)	267 (29.3)	324 (32.3)	0.06	29.3	29.3	<0.01
SLE severity index, n (%)			0.03			<0.01
Mild	317 (34.8)	343 (34.2)		35.8	35.8	
Moderate	375 (41.2)	415 (41.3)		40.4	40.4	
Severe	218 (24.0)	246 (24.5)		23.9	23.9	
Charlson comorbidity index, mean (SD)	1.5 (1.4)	1.8 (1.8)	0.18	1.6	1.6	<0.01
Comorbidities, n (%)						
CKD stage ≥3	293 (32.2)	366 (36.5)	0.09	32.7	32.7	<0.01
Heart failure	118 (13.0)	181 (18.0)	0.14	14.9	14.9	<0.01
CVD	212 (23.3)	238 (23.7)	0.01	23.1	23.1	<0.01
Obesity	416 (45.7)	275 (27.4)	0.39	35.4	35.4	<0.01
Diabetic complications						
Diabetic nephropathy	134 (14.7)	154 (15.3)	0.02	13.7	13.7	<0.01
Diabetic retinopathy	90 (9.9)	61 (6.1)	0.14	7.4	7.4	<0.01
Diabetic neuropathy	246 (27.0)	199 (19.8)	0.17	21.8	21.8	<0.01
Diabetic vascular disease	61 (6.7)	53 (5.3)	0.06	5.1	5.1	<0.01
Tobacco use, n (%)	246 (27.0)	246 (24.5)	0.06	24.8	24.8	<0.01
SLE medication use, n (%)						
Glucocorticoids	468 (51.4)	498 (49.6)	0.04	49.4	49.4	<0.01
Hydroxychloroquine	366 (40.2)	351 (35.0)	0.11	38.3	38.3	<0.01
Azathioprine	65 (7.1)	60 (6.0)	0.05	6.1	6.1	<0.01
Methotrexate	85 (9.3)	80 (8.0)	0.05	8.6	8.6	<0.01
Mycophenolate	93 (10.2)	112 (11.2)	0.0303	10.6	10.6	<0.01
Belimumab	32 (3.5)	14 (1.4)	0.1374	2.1	2.1	<0.01
Other immunosuppressant	80 (8.8)	79 (7.9)	0.0334	8.6	8.6	<0.01
Other medications, n (%)						<0.01
ACEi or ARB	366 (40.2)	407 (40.5)	0.0065	38.4	38.4	<0.01
Insulin use	356 (39.1)	324 (32.3)	0.1433	32.9	32.9	<0.01
Metformin use	356 (39.1)	368 (36.7)	0.0509	35.2	35.2	<0.01
Sulfonylurea use	109 (12.0)	170 (16.9)	0.1412	13.0	13.0	<0.01
Other hypoglycemic medication	28 (3.1)	40 (4.0)	0.0492	3.0	3.0	<0.01
eGFR, mean (SD)	81.5 (27.0)	74.7 (28.6)	0.2416	79.0	79.0	<0.01
Health care use						
Outpatient visits, median (IQR)	9 (2–17)	7 (1–17)	0.1010	7 (2–17)	7 (2–17)	<0.01
ER/inpatient visits, n (%)	332 (36.5)	415 (41.3)	0.0996	37.8	37.8	<0.01
Hemoglobin A1c (%)	7.8 (2.5)	7.7 (3.0)	0.0310	7.6	7.6	<0.01

* ACEi, angiotensin-converting enzyme inhibitor; ARB, angiotensin receptor blocker; CKD, chronic kidney disease; CVD, cardiovascular disease; DPP4i, dipeptidyl peptidase 4 inhibitors; eGFR, estimated glomerular filtration rate; ER, emergency room; GLP-1RA, glucagon-like peptide-1 receptor agonist; IQR, interquartile range; SLE, systemic lupus erythematosus; Std. Diff., standardized difference.

observed until the earliest of the outcome, death, or end of the study period.

We conducted survival analyses with Cox proportional hazard regression to compare incidence rates for each outcome and assess hazard ratios with 95% confidence intervals (CIs) for the weighted populations. We also performed a secondary analysis restricted to patients with LN at baseline. We had full data access, and all analyses were conducted using suite of analytics software 9.4 (SAS Institute Inc.).

RESULTS

Baseline characteristics. There were 910 initiators of GLP-1RA and 1,004 initiators of DPP4i with SLE and T2D, of which 267 and 324 patients had LN, respectively (Table 1). After propensity score overlap weighting, the treatment groups were balanced across all covariates. Patients were mostly women (92%), with a mean age of 55 years. Nearly one-half of patients were non-Hispanic White, and 35% were Black. Nearly one-half used glucocorticoids, and 38% used an angiotensin-converting enzyme inhibitor or angiotensin receptor blocker. Additionally, 33% of patients had CKD stage ≥ 3, 15% had heart failure, and 35% were obese. The average HgA1C was 7.6%.

Outcomes in all patients with SLE. In the primary per-protocol analysis, GLP-1RA initiators with SLE had lower risks of MACE (adjusted hazard ratio [aHR] 0.66, 95% CI 0.48–0.91) and kidney disease progression (aHR 0.77, 95% CI 0.60–0.98) compared with initiators of DPP4i (Table 2). There was no significant difference in the aHR for individual components of myocardial infarction, ischemic stroke, or heart failure hospitalization, but the incidence of myocardial infarction and heart failure tended to be lower with GLP-1RA use. The risk of VTE was lower with GLP-1RA than DPP4i initiation (aHR 0.49, 95% CI 0.24–0.97), as was the risk of all-cause death during the study period. (aHR 0.26, 95% CI 0.10–0.68). There was no difference in the risk of the control outcome of genital infection among the treatment groups, as expected.

Outcomes in patients with LN. In the secondary analysis of patients with LN, GLP-1RA initiation compared with DPP4i

Table 2. Cardiovascular and kidney outcomes associated with GLP-1RA versus DPP4i use in systemic lupus erythematosus*

	Events, n		Follow-up time, y		Incidence rate (per 1,000 person years)		
	GLP-1RA	DPP4i	GLP-1RA	DPP4i	GLP-1RA	DPP4i	Hazard ratio (95% CI)
Per-protocol							
Cardiovascular outcomes							
MACE	32	42	1.3	1.2	61.1	92.9	0.66 (0.48–0.91)
Myocardial infarction	3	6	1.4	1.3	5.6	12.8	0.43 (0.16–1.14)
Stroke	8	6	1.4	1.3	13.9	12.5	1.08 (0.51–2.31)
Heart failure	23	26	1.4	1.2	44.1	55.7	0.78 (0.53–1.15)
VTE	7	13	1.4	1.2	12.6	26.5	0.49 (0.24–0.97)
Kidney outcome							
eGFR decline by ≥30% or new ESKD	58	68	1.3	1.1	117.5	158.1	0.77 (0.60–0.98)
All-cause death	3	11	1.4	1.3	5.6	21.6	0.26 (0.10–0.68)
Control outcome							
Genital infection	27	24	1.3	1.2	52.2	52.4	1.02 (0.69–1.50)
Intention-to-treat							
Cardiovascular outcomes							
MACE	54	63	2.7	2.5	51.3	64.7	0.80 (0.62–1.03)
Myocardial infarction	10	11	3.0	2.8	8.5	10.3	0.81 (0.44–1.47)
Stroke	13	13	3.0	2.8	11.1	11.6	0.99 (0.57–1.72)
Heart failure	40	41	2.8	2.7	36.9	40.0	0.93 (0.69–1.26)
VTE	10	19	3.0	2.8	8.9	17.5	0.53 (0.30–0.94)
Kidney outcome							
eGFR decline by ≥30% or new ESKD	92	108	2.5	2.2	95.2	124.3	0.79 (0.65–0.96)
All-cause death	10	22	3.0	2.9	8.4	19.5	0.42 (0.25–0.73)
Control outcome							
Genital infection	46	44	2.7	2.6	42.8	43.8	1.01 (0.75–1.36)

* MACE includes myocardial infarction, ischemic stroke, or heart failure hospitalization. CI, confidence interval; DPP4i, dipeptidyl peptidase 4 inhibitors; eGFR, estimated glomerular filtration rate; ESKD, end-stage kidney disease; GLP-1RA, glucagon-like peptide-1 receptor agonist; MACE, major adverse cardiac events; VTE, venous thromboembolism.

use was similarly associated with lower risks of MACE (aHR 0.64, 95% CI 0.41–0.98) and kidney disease progression (aHR 0.70, 95% CI 0.49–1.00; Table 3). The effect sizes for VTE and all-cause death did not reach statistical significance in the LN subgroup.

In the intention-to-treat analyses, the direction of effect of GLP-1RA use was similar across outcomes with some attenuation, with lower risks of kidney disease progression and all-cause death with GLP-1RA initiation compared with DPP4i initiation for all patients with SLE (progression of kidney disease aHR 0.79, 95% CI 0.65–0.96; all-cause death aHR 0.42, 95% CI 0.25–0.73) and patients with LN (progression of kidney disease aHR 0.63, 95% CI 0.47–0.85; all-cause death aHR 0.32, 95% CI 0.13–0.79]).

DISCUSSION

This emulated target trial using observational data demonstrated that GLP-1RA use was associated with a lower risk of MACE, kidney disease progression, VTE, and all-cause death, compared with DPP4i use among patients with SLE and comorbid T2D. Similar benefits were also demonstrated among patients with LN. This is the first observational study, to our knowledge, demonstrating the potential cardioprotective, nephroprotective, and mortality benefits of GLP-1RAs for patients with SLE and LN, who have been largely excluded from participating in randomized clinical trials for GLP-1RAs.[3,5]

It is widely established that GLP-1RAs reduce CV risk and attenuate progression of kidney disease among those with T2D[3,5,13,14]; however, recent studies suggest GLP-1RAs also confer similar benefits to other populations with elevated risks of CV events and/or progressive CKD.[4,6] In the Semaglutide Effects on Cardiovascular Outcomes in People with Overweight or Obesity trial, which included patients with CV disease, a body mass index of ≥27, and no prior diagnosis of diabetes, the use of the GLP-1RA semaglutide, compared with placebo, reduced the incidence of death from CV disease, nonfatal myocardial infarction, or nonfatal stroke.[4] Many patients with SLE have traditional CV risk factors, in addition to the elevated CV risk related to SLE itself, so many patients will now have indications for GLP-1RA use and may have considerable potential benefit. Although our emulated trial could not isolate the effect of GLP-1RA on patients with SLE and without T2D, we followed substantial reductions in the risk of MACE for patients with both conditions. Additionally, the

Table 3. Cardiovascular and kidney outcomes associated with GLP-1RA versus DPP4i use in lupus nephritis*

Outcomes	Events, n		Follow-up time, y		Incidence rate (per 1,000 person years)		Hazard ratio (95% CI)
	GLP-1RA	DPP4i	GLP-1RA	DPP4i	GLP-1RA	DPP4i	
Per-protocol							
Cardiovascular outcomes							
MACE	18	23	1.2	1.0	122.2	193.0	0.64 (0.41–0.98)
Myocardial infarction	2	5	1.4	1.2	11.0	32.7	0.34 (0.10–1.09)
Stroke	2	4	1.4	1.2	14.7	25.3	0.58 (0.18–1.85)
Heart failure	15	16	1.2	1.1	98.5	124.9	0.78 (0.48–1.28)
VTE	4	5	1.3	1.2	24.9	38.3	0.62 (0.25–1.56)
Kidney outcome							
eGFR decline by ≥30% or new ESKD	28	35	1.1	0.9	213.3	325.2	0.70 (0.49–1.00)
All-cause death	3	5	1.4	1.2	16.2	36.5	0.46 (0.16–1.37)
Control outcome							
Genital infection	9	8	1.3	1.1	56.6	62.9	0.90 (0.48–1.68)
Intention-to-treat							
Cardiovascular outcomes							
MACE	25	30	2.1	1.9	97.7	132.7	0.75 (0.52–1.09)
Myocardial infarction	3	8	2.5	2.3	11.1	29.0	0.39 (0.17–0.91)
Stroke	4	5	2.5	2.3	12.9	18.8	0.70 (0.28–1.73)
Heart failure	20	23	2.2	2.0	79.2	93.7	0.85 (0.56–1.29)
VTE	5	7	2.4	2.3	16.3	25.5	0.66 (0.29–1.51)
Kidney outcomes							
eGFR decline by ≥30% or new ESKD	38	52	1.9	1.5	171.8	285.0	0.63 (0.47–0.85)
All-cause death	4	11	2.5	2.4	12.2	39.3	0.32 (0.13–0.79)
Control outcome							
Genital infection	10	12	2.3	2.2	36.6	48.8	0.78 (0.44–1.39)

* MACE outcome includes combined myocardial infarction, ischemic stroke, or heart failure hospitalization. Heart failure outcome includes heart failure hospitalization. CI, confidence interval; DPP4i, dipeptidyl peptidase 4 inhibitor; eGFR, estimated glomerular filtration rate; ESKD, end-stage kidney disease; GLP-1RA, glucagon-like peptide-1 receptor agonist; MACE, major adverse cardiac events; VTE, venous thromboembolism.

kidney-protective effects of GLP-1RA are likely to extend beyond treatment of patients with diabetic nephropathy. SeMaglutide and Albuminuria Reduction Trial in Obese Individuals Without Diabetes recently demonstrated that weekly administration of semaglutide led to appreciable reduction in albuminuria in a cohort of patients with nondiabetic CKD, including 25 patients with chronic glomerulonephritis.[6] Persistent proteinuria is a major risk factor for progressive CKD in LN; therefore, reducing proteinuria is a potential mechanism of benefit for GLP-1RA for patients with LN. We found a reduction in eGFR decline or new onset ESKD in patients with LN and T2D. Future studies are needed to determine the impact of GLP-1RA on CV events, proteinuria, and kidney disease progression among patients with SLE without T2D. Furthermore, future studies should assess the impact GLP-1RA use for patients undergoing standard immunosuppression treatment for LN.

Although research is ongoing to further identify specific mechanisms contributing to the cardio- and kidney-protective properties of GLP1-RAs aside from glucose-lowering and hemodynamic effects, studies in murine models and humans suggest these agents potentially possess immunomodulatory properties.[15] GLP1-RAs have been shown in murine models to regulate lymphocyte proliferation and maintenance of regulatory T cells, and studies in humans have demonstrated that GLP-1RAs decrease production of inflammatory cytokines and oxidative stress biomarkers.[15,16] It remains unclear to what extent these mechanisms are mediated through weight reduction or separate, direct processes. However, these posited immune mechanisms of GLP-1RAs warrant further investigation in patients with SLE and LN.

This study is strengthened by the employment of the emulated target trial framework to assess causal associations from observational data, with the specific use of propensity score overlap weighting to mitigate confounding. Limitations of our study include the relatively small sample size and the potential for misclassification that is inherent to observational EHR data. This may include misclassification of medication exposure because we lacked pharmacy dispensing records and could not assess medication adherence to prescribed treatment. Lastly, we note that the inability to investigate the efficacy of GLP-1RA among patients with SLE and without T2D serves as another limitation.

To conclude, our findings suggest that GLP-1RAs confer substantial cardioprotective and kidney-protective benefits to patients with SLE and LN with comorbid T2D, who are inherently predisposed to a markedly disproportionate risk of atherosclerotic CV disease and CKD. These support a potential role for GLP1-RAs in the management and treatment of patients with SLE and LN, although additional studies are needed to assess the safety of GLP-1RAs in this population. Furthermore, patients with SLE should be included in future randomized trials involving GLP-1RAs, and future studies should evaluate the efficacy of GLP-1RA use among patients with SLE and LN without concomitant diabetes. Potential barriers to accessing these medications, including cost, should also be addressed to avoid exacerbating health care disparities. Lastly, basic science and translational studies are needed to help provide mechanistic insight for these findings, which may elucidate novel indications and benefits.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr Jorge confirms that all authors have provided the final approval of the version to be published, and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. Mahajan A, Amelio J, Gairy K, et al. Systemic lupus erythematosus, lupus nephritis and end-stage renal disease: a pragmatic review mapping disease severity and progression. Lupus 2020;29(9):1011–1020.
2. Wong CY, Ma BMY, Zhang D, et al. Cardiovascular risk factors and complications in patients with systemic lupus erythematosus with and without nephritis: a systematic review and meta-analysis. Lupus Sci Med 2024;11(1):e001152.
3. Sattar N, Lee MMY, Kristensen SL, et al. Cardiovascular, mortality, and kidney outcomes with GLP-1 receptor agonists in patients with type 2 diabetes: a systematic review and meta-analysis of randomised trials. Lancet Diabetes Endocrinol 2021;9(10):653–662.
4. Lincoff AM, Brown-Frandsen K, Colhoun HM, et al. Semaglutide and cardiovascular outcomes in obesity without diabetes. N Engl J Med 2023;389(24):2221–2232.
5. Perkovic V, Tuttle KR, Rossing P, et al. Effects of semaglutide on chronic kidney disease in patients with type 2 diabetes. N Engl J Med 2024;391(2):109–121.
6. Apperloo EM, Gorriz JL, Soler MJ, et al. Semaglutide in patients with overweight or obesity and chronic kidney disease without diabetes: a randomized double-blind placebo-controlled clinical trial. Nat Med 2025;31(1):278–285.
7. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. Am J Epidemiol 2016;183(8):758–764.
8. Giugliano D, Longo M, Signoriello S, et al. The effect of DPP-4 inhibitors, GLP-1 receptor agonists and SGLT-2 inhibitors on cardiorenal outcomes: a network meta-analysis of 23 CVOTs. Cardiovasc Diabetol 2022;21(1):42.
9. Materne E, Choi H, Zhou B, et al. Comparative risks of infection with belimumab versus oral immunosuppressants in patients with non-renal systemic lupus erythematosus. Arthritis Rheumatol 2023;75(11):1994–2002.
10. Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40(5):373–383.
11. Garris C, Jhingran P, Bass D, et al. Healthcare utilization and cost of systemic lupus erythematosus in a US managed care health plan. J Med Econ 2013;16(5):667–677.
12. Lix LM, Sobhan S, St-Jean A, et al. Validity of an algorithm to identify cardiovascular deaths from administrative health records: a

multi-database population-based cohort study. BMC Health Serv Res 2021;21(1):758.

13. Rossing P, Bain SC, Bosch-Traberg H, et al. Effect of semaglutide on major adverse cardiovascular events by baseline kidney parameters in participants with type 2 diabetes and at high risk of cardiovascular disease: SUSTAIN 6 and PIONEER 6 post hoc pooled analysis. Cardiovasc Diabetol 2023; 22(1):220.
14. Perkovic V, Tuttle KR, Rossing P, et al. Effects of semaglutide on chronic kidney disease in patients with type 2 diabetes. N Engl J Med 2024;391(2):109–121.
15. Drucker DJ. The cardiovascular biology of glucagon-like peptide-1. Cell Metab 2016;24(1):15–30.
16. Mehdi SF, Pusapati S, Anwar MS, et al. Glucagon-like peptide-1: a multi-faceted anti-inflammatory agent. Front Immunol 2023;14:1148209.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis & Rheumatology
Vol. 78, No. 3, March 2026, pp 672–683
DOI 10.1002/art.43395

Podocyte Interleukin-23 Receptor Signaling in the Pathogenesis of Lupus Nephritis

Rong Fu,[1] Afroditi Boulougoura,[1] Shuilian Yu,[1] Hao Li,[1] Wenliang Pan,[1] Yushiro Endo,[1] Rhea Bhargava,[1] Abhiyan Satyam,[1] Vivek Kasinath,[2] Jarrat Jordan,[3] Nandan Padmanabha,[4] Maria G. Tsokos,[1] Reza Abdi,[2] and George C. Tsokos[1]

Objective. Up-regulation of interleukin-23 (IL-23) in the serum and kidneys of patients with lupus nephritis (LN) has been demonstrated, but its effect on podocytes remains unknown. We hypothesized that IL-23 contributes to podocyte injury and that targeted deletion of IL-23 receptor (IL-23R) in podocytes of lupus-prone mice can prevent the development of glomerulonephritis.

Methods. Kidney biopsies were immunostained for IL-23R. In vitro experiments were conducted using a human podocyte cell line and primary murine podocytes. Human podocytes stimulated with IL-23 underwent bulk RNA sequencing. The expression of IL-23R and structure and motility of podocytes were assessed. Podocytes isolated from B6 wild-type mice injected with a minicircle (MC) encoding IL-23 were studied. To assess the role of IL-23R in the development of nephritis, we generated MRL/*lpr* mice deficient in podocyte-specific *Il23r* who were lupus prone.

Results. IL-23R was highly expressed in the glomeruli of patients with LN. IL-23R expression was also up-regulated in human podocytes and primary podocytes isolated from B6 mice after IL-23 stimulation. Human podocytes stimulated with IL-23 showed decreased expression of synaptopodin and remodeling of the actin cytoskeleton. Mice who were administered with IL-23 MC mice exhibited a significant increase in the expression of IL-23R and phosphorylated STAT3 in podocytes. Finally, MRL/*lpr*.Podo-Cre^{+} *Il23r*$^{fl/fl}$ mice showed decreased clinical and histologic features of LN.

Conclusion. IL-23R expression is increased in podocytes from mice and humans with systemic lupus erythematosus. IL-23 signaling disrupts the cytoskeleton in podocytes and increases their mobility, leading to the development of glomerulonephritis. Podocyte-specific deletion of *Il23r* in lupus-prone mice abrogates the development of LN.

INTRODUCTION

The deposition of immune complexes (ICs), activation of complement, and infiltration of the kidney by cells of the adaptive and innate immune systems have long been considered responsible for the development of kidney damage in lupus nephritis (LN).[1] However, recent findings have also highlighted the contribution of kidney-resident immune cells and immune molecules expressed by kidney parenchymal cells to disease pathogenesis.[2] Podocytes are highly specialized, terminally differentiated epithelial cells in the glomeruli of kidneys, and they are vulnerable to a variety of stimuli, resulting in a series of morphologic and physiologic changes such as cytoskeletal rearrangement, slit diaphragm protein dysregulation, mitochondria dysfunction, and increased motility.[3] The injury and/or dysfunction of podocytes play a critical role in the pathogenesis of proteinuria in the vast majority of glomerulonephritis (GN), including anti–glomerular basement membrane GN and lupus GN (III and IV).[4] Dynamic regulation of the podocyte actin cytoskeleton is vital to the normal kidney filter function, and mutations lead to the rearrangement of the actin cytoskeleton and subsequent proteinuria.[5] Because podocytes have limited ability to repair and/or regenerate, the extent of podocyte injury is considered as a major prognostic determinant in end-stage renal disease.[3] Indeed, the actin cytoskeleton of

Dr G. C. Tsokos's work was supported by the NIH (grant P01-AI-179405) and Janssen (grant 430070383).

[1]Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts; [2]Department of Medicine, Brigham and Women's Hospital, Boston, Massachusetts; [3]Janssen Pharmaceuticals, Spring House, Pennsylvania; [4]Department of Pathology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts.

Drs Fu, Boulougoura, Yu, and Li contributed equally to this work.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/art.43395).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/art.43395.

Address correspondence via email to George C. Tsokos, MD, at gtsokos@bidmc.harvard.edu.

Submitted for publication June 13, 2025; accepted in revised form August 29, 2025.

podocytes is a direct target of the antiproteinuric effect of cyclosporine and rituximab.[6,7] Any intervention that can stabilize the podocyte actin cytoskeleton or promote podocyte recovery and/or repair should have therapeutic potential in proteinuric states.

Interleukin-23 (IL-23) is a heterodimeric cytokine with a wide range of effects on the immune response.[8,9] The pivotal role of IL-23/IL-23 receptor (IL-23R)/STAT3 signaling axis has been studied in systemic lupus erythematosus (SLE), with IL-23R primarily expressed on various hematopoietic cells, including lymphocytes and macrophages.[10,11] However, recent evidence suggests that IL-23 may also act directly on kidney-resident cells.[12] We have previously demonstrated that global IL-23R deficiency prevents the development of LN in both C57BL/6 (B6) *lpr/lpr* and MRL/*lpr* mice.[13,14] Moreover, renal pathology was limited in the mice with nephritic IL-23p19$^{-/-}$.[15,16] Another interesting finding from our group is that the anti-IL-23 antibody ameliorated albuminuria in mice who were lupus prone without affecting the production of anti–double-stranded DNA (dsDNA) antibodies,[17] which gave us a hint that IL-23 could directly act on glomeruli independently of systemic inflammation. The signal transducer and activator of STAT3 is the major downstream effector of the IL-23 signaling pathway, which promotes directional cell migration and reducing the formation of actin stress.[18,19] Podocyte STAT3 activation is critical for the development of crescentic GN.[20] Meanwhile, inhibition of the STAT3 pathway partially diminishes the suppression of synaptopodin in HIV or Nef-infected podocytes and nephrotoxic serum-induced crescentic nephritis.[20–22] Here, we report that IL-23 acts directly on podocytes, up-regulates IL-23R expression on their surface, promotes STAT3 phosphorylation, and induces podocyte injury characterized by enhanced cell motility and disrupted cytoskeletal dynamics. Most importantly, targeted *Il23r* deletion in podocytes of lupus-prone mice ameliorates the development of GN and preserves podocyte structure.

METHODS

Podocyte culture. Immortalized human podocytes[23] were cultured as previously described.[24] Podocytes were grown in RPMI 1640 medium supplemented with 10% fetal bovine serum (FBS) and 1% insulin–transferrin–selenium at 33°C and differentiated into mature podocytes after transfer to 37°C for 7 to 10 days due to the presence of the temperature-sensitive SV40-T gene. Mature podocytes were stimulated with phosphate buffered saline (PBS) or 10 ng/mL recombinant human IL-23 (1290-IL; R&D) for the indicated times. Five micromoles of phosphorylated STAT3 (pSTAT3) inhibitor (Stattic, HY-13818; MedChemExpress) was added one hour before stimulation. Primary murine podocytes were isolated from eight-week-old B6 mice as described.[25,26] Briefly, mice were perfused with Dynabead M450 beads (14011; Thermo Fisher), kidneys were collected and digested, and glomeruli were enriched using a magnet. The isolated glomeruli were cultured in RPMI 1640 medium supplemented with 10% FBS for four to seven days and stimulated with PBS or 10 ng/mL recombinant mouse IL-23 (11269-ML; R&D) for 24 hours.

Mice. MRL/*lpr* and B6. *Nphs2*$^{cre+}$ (strain 008205) mice were obtained from the Jackson Laboratory. B6. *Il23r*$^{fl/fl}$ mice were generated as described previously.[12,27] Podocyte-specific *Il23r* conditional knockout (CKO) mice on an MRL/*lpr* background were generated using the Cre–loxP (locus of X-over P1) recombination system. Briefly, B6. *Il23r*$^{fl/fl}$ and B6. *Nphs2*$^{cre+}$ mice were backcrossed onto the MRL/*lpr* background for 11 generations to generate MRL/*lpr.Il23*$^{fl/fl}$ and MRL/*lpr.Nphs2*$^{cre+}$ mice. MRL/*lpr.Il23*$^{fl/fl}$ mice were then crossed with MRL/*lpr.Nphs2*$^{cre+}$ mice to generate MRL/*lpr.Nphs2*$^{cre+}$*Il23*$^{fl/fl}$ (MRL/*lpr.Podo-cre*$^{+}$*Il23r* $^{fl/fl}$, CKO) mice. Female CKO and control mice were used in three independent experiments and were euthanized between 18 and 23 weeks of age. Sera, urine, and kidneys were collected. Mice were maintained under specific pathogen-free conditions at Beth Israel Deaconess Medical Center (BIDMC). All animal procedures were approved by the Institutional Animal Care and Use Committee of BIDMC.

Human kidney tissues. All kidney samples were obtained from the Department of Pathology, BIDMC, Harvard Medical School (Institutional Review Board number 2006P000298). The studied cases included renal tissue from 11 patients with LN (International Society of Nephrology/Renal Pathology Society classes III–V) and uninvolved renal tissue from three individuals who underwent nephrectomies for primary renal neoplasms, who were used as controls.

Immunofluorescence. Human or murine kidney frozen sections (5 μm) were fixed in ice-cold acetone for 15 minutes, blocked with 5% bovine serum albumin (BSA) for one hour, and stained overnight with primary antibodies, including anti-IL-23R (NB600-1147; Novus), anti-pSTAT3 (612542; BD Biosciences), anti-synaptopodin Alexa Fluor 647 (sc-515842; Santa Cruz), and antinephrin (GP-N2; Progen). Next, sections were incubated for one hour with Alexa Fluor 488– or 555–conjugated secondary antibodies. Nuclei were counterstained with DAPI or Hoechst. Finally, sections were mounted using antifade mounting medium. For immunofluorescence staining of cultured podocytes, cells grown on coverslips were fixed with 4% paraformaldehyde for 15 minutes at room temperature, permeabilized with 0.1% Triton X-100 for 10 minutes, and blocked with 5% BSA for one hour. Cells were then incubated with primary antibodies including Alexa Fluor 488–conjugated or Alexa Fluor 594–conjugated phalloidin (A12379; Thermo Fisher), anti-IL-23R (NB600-1147; Novus), Alexa Fluor 647–conjugated anti-synaptopodin antibody (sc-515842; Santa Cruz), and anti-pSTAT3 (612542; BD Biosciences). After washing, cells were incubated with appropriate

fluorescence-conjugated secondary antibodies and mounted with antifade mounting medium containing DAPI. Images were acquired using a Zeiss LSM 880 confocal microscope or Keyence BZ-X fluorescence microscope. Fluorescence intensities were measured using ImageJ software.

Histopathology analysis. Kidney sections were fixed in 10% neutral buffered formalin and embedded in paraffin. Four-micrometer sections were cut and stained with hematoxylin and eosin. Glomeruli were evaluated microscopically in a blinded manner and scored as previously described.[28] In addition, renal cortical tissue from two CKO mice and one MRL/*lpr* control mouse was processed for electron microscopy and examined ultrastructurally as previously described.[29]

Preparation and in vivo administration of IL-23 minicircle DNA constructs. IL-23 minicircle (MC) DNA constructs were generated as previously described.[7] Briefly, a single isolated colony from a fresh plate was cultured for eight hours in Luria–Bertani broth with ampicillin and expanded in terrific brothfor 17 additional hours. Subsequently, the culture medium was replaced with fresh Luria–Bertani broth containing 1% L-arabinose. After the addition of one-half volume of fresh low-salt Luria–Bertani broth (pH 8.0) containing 1% L-arabinose, the incubation temperature was increased to 37°C, and the incubation continued for two additional hours. Episomal DNA MCs were prepared from bacteria using QIAGEN Endo free Mega preparation plasmid purification kits. Hydrodynamic delivery of 8 μg IL-23 MC DNA per mouse was performed via tail vein injection in a total volume of approximately 10% of the mouse body weight within a period of five to seven seconds in all experiments. Mice injected with an equal volume of PBS served as controls (n = 4 per group). Eight-week-old wild-type B6 mice were used, and kidney samples were collected 48 hours after injection.

Enzyme-linked immunosorbent assay. The concentration of IL-23 in the sera of IL-23 MC–treated and control mice was measured using a commercially available enzyme-linked immunosorbent assay (ELISA) kit (M2300; R&D) according to the manufacturer's instructions. Anti-dsDNA levels in the sera of MRL/*lpr* CKO and control mice were measured as described previously.[12]

Western blot. Podocytes were lysed with radioimmunoprecipitation assay buffer containing phosphatase and protease inhibitors at 4°C for 30 minutes. Lysates were centrifuged, supernatants were collected, and equal amounts of protein from each supernatant were loaded onto NuPAGE 4% to 12% Bis-Tris gels (NP0336; Thermo Fisher). Proteins were transferred to a PVDF membrane, which was subsequently blocked for one hour and incubated overnight at 4°C with rabbit anti-IL-23R (NB600-1147; Novus), rabbit anti-Jak2 (3230T; Cell Signaling Technology), rabbit anti-Tyk2 (14193; Cell Signaling Technology), rabbit anti-STAT3 (ab68153; Abcam), rabbit anti-pSTAT3 (612542; BD Biosciences), mouse anti-synaptopodin (sc-515842; Santa Cruz), mouse anti-nephrin (sc-376522; Santa Cruz), and rabbit anti-GAPDH (97166; Cell Signaling Technology) antibodies. The membrane was washed with Tris buffered saline–Tween and then incubated with horseradish peroxidase-conjugated anti-rabbit or anti-mouse secondary antibodies. Signal detection was performed using Clarity Western ECL Substrate (Bio-Rad). Band intensities were quantified using ImageJ and normalized to GAPDH.

Wound-healing assay. Podocytes were seeded in six-well plates and cultured overnight. Each well was scratched with a sterile 200-μL pipette tip and washed with PBS. Podocytes were then stimulated with PBS, 10 ng/mL IL-23, or 10 ng/mL IL-23 plus pSTAT3 inhibitor. The wound closure was observed after eight hours by phase-contrast microscopy. Each condition was examined in triplicate.

Kidney function analysis. Spot urine samples were collected. Urinary proteins were determined with BCA kit (Thermo Fisher), and creatinine levels were quantified using ELISA kit (R&D Systems). Urine protein-to-creatinine ratios (UPCRs) were calculated to assess severity of proteinuria.

Transcriptome analysis (bulk RNA sequencing of human podocytes). Total RNA was isolated from podocytes stimulated with IL-23 or PBS for two hours using the QIAGEN RNeasy Micro Kit (QIAGEN). RNA quality and concentration were assessed using NanoDrop (Thermo Fisher Scientific). Sequencing libraries were constructed with the KAPA HyperPrep Kit (Roche) and assessed on Tapestation (High sensitivity D1000 Screen Tape; Agilent). Sequencing was performed using the Illumina NovaSeq 6000 platform. Differentially expressed genes were identified by Limma. Pathway enrichment analysis was performed using ClusterProfiler.

Statistics. All results were shown as the mean ± SEM. Statistical significance was determined by *t*-test (two-tailed) for two groups or by one-way analysis of variance with Bonferroni's multiple-comparison test for three or more groups. *P* values <0.05 were considered significant.

RESULTS

Increased expression of IL-23R in glomeruli of patients and mice with LN. The expression of IL23R on podocytes has not been established. We therefore examined the expression of IL-23R in podocytes using biopsy sections from patients with LN. The clinical information of the patients is presented in Supplementary Table 1. Confocal immunofluorescence images showed increased IL-23R expression in the glomeruli of patients with LN compared to controls (Figure 1A and B).

The expression levels of IL-23R correlated inversely with serum complement 3 (C3) levels (Figure 1C). There was no significant correlation with serum C4 levels or UPCR (Figure 1D and E). Next, we examined the expression of IL-23R in kidney podocytes from MRL/*lpr* mice before (4 and 12 weeks) and after (20 weeks) the establishment of LN. Similar to findings in human LN, we followed increased expression of IL-23R in podocytes from the 12- and 20-week-old mice, but not the 4-week-old mice (Figure 1F and G). The expression levels of IL-23R correlated positively with UPCR, which suggests that the function of podocytes is compromised (Figure 1H).

Up-regulation of IL-23R in podocytes by IL-23 in vitro and in vivo. Previously, we demonstrated that IL-23 alone can induce renal inflammation by suppressing the production of arginine by tubular epithelial cells (TECs).[12] However, it remains unknown whether the phenotype of podocytes changes, including its expression of IL-23R in response to IL-23 directly. To address this question, human podocytes were cultured in vitro,[24] and after proliferation and differentiation, they were stimulated with IL-23 or PBS for 24 hours. The immunofluorescence images (Figure 2A) and Western blot results (Figure 2B) revealed that stimulation with IL-23 increased the expression of IL-23R in podocytes. IL-23 also elevated Jak2 expression, but not Tyk2 (Figure 2B). Next, primary podocytes were isolated from naive B6 mouse kidneys (Figure 2C) and stimulated with IL-23. Although the expression of IL-23R in murine podocytes is low under normal conditions, it increased significantly after 24 hours of IL-23 stimulation by immunofluorescence (Figure 2D). To validate this observation in vivo, an MC complementary DNA construct encoding IL-23p19 and IL-12/23p40 (IL-23 MC) was administered intravenously to eight-week-old wild-type B6 mice, resulting in highly elevated and sustained IL-23 levels in the serum of animals who were treated (Figure 2E). Immunofluorescence images confirmed the induction of IL-23R on synaptopodin$^+$ podocytes after the administration of IL-23 MC (Figure 2F).

IL-23 and the structure and function of human podocytes. To further characterize phenotypic changes in podocytes responding to IL-23, bulk RNA sequencing was performed on human podocytes stimulated with IL-23 or PBS. Pathway enrichment analysis revealed pathways related to synapse/cell adhesion assembly and organization and cell motility and migration as the top pathways altered by IL-23 stimulation (Figure 3A). Podocytes have a complex cellular architecture maintained by a well-organized actin cytoskeleton. To determine whether *IL-23 disrupts the structure of actin fibers leading to* cytoskeletal damage *in podocytes, we* stained IL-23-treated podocytes with phalloidin, a selective F-actin marker. Immunofluorescent imaging revealed that IL-23 stimulation led to F-actin disassembly, indicating cytoskeletal disruption (Figure 3B). Next, we asked whether IL-23 down-regulates the expression of nephrin and synaptopodin in podocytes, which could explain their increased mobility. Western blot analysis revealed a reduction in synaptopodin, but not nephrin, 24 hours after IL-23 stimulation (Figure 3C). To evaluate whether IL-23 promotes podocyte motility, a wound-healing assay was conducted on cultured human podocytes stimulated with IL-23 (10 ng/mL) for 24 hours. IL-23 significantly increased podocyte migration into the wound area, indicating enhanced motility (Figure 3D).

Induction of rapid STAT3 activation in podocytes by IL-23. In T cells, it has been shown that *IL-23 signals through STAT3; inhibition of STAT3 impairs IL-23-driven Th17 polarization*.[30–32] Having demonstrated that IL-23R can be induced on podocytes, we explored the IL-23 signaling pathway in podocytes. As expected, immunofluorescence analysis confirmed increased expression of pSTAT3 in synaptopodin$^+$ podocytes in the kidneys of B6 mice 48 hours after IL-23 MC administration (Figure 4A). Consistently, Western blot (Figure 4B) and immunofluorescence staining (Figure 4C) demonstrated that stimulation with IL-23 induced phosphorylation and nuclear localization of STAT3 in cultured human podocytes, which is indicative of STAT3 activation. Stattic, a specific STAT3 inhibitor, reverted IL-23-induced phosphorylation of STAT3. Moreover, Stattic reversed IL-23-mediated F-actin disassembly and reduced podocyte hypermotility by F-actin staining and wound-healing assay (Figure 4D and E). Collectively, our data indicate that IL-23 activates STAT3 in podocytes, promoting cytoskeletal disorganization and increased motility.

Amelioration of LN by podocyte-specific deletion of *Il23r* in MRL/*lpr* mice. Having demonstrated that IL-23 up-regulates IL-23R in podocytes and disrupts their structure and motility, and given that serum IL-23 levels are elevated in patients with SLE,[33] we investigated the importance of IL-23R in the pathogenesis of LN in vivo. To this end, we generated MRL/*lpr* CKO mice who lacked *Il23r* only in podocytes (MRL/*lpr.Podo-cre*$^+$*Il23r*$^{fl/fl}$; Figure 5A) and studied them in parallel with MRL/*lpr. Il23*$^{fl/f}$, or wild-type MRL/*lpr* control, mice between 18 and 23 weeks of age. CKO mice had less proteinuria (Figure 5B) and demonstrated significantly reduced glomerular pathology compared to controls (Figure 5C). Notably, electron microscopy showed decreased IC deposition and preserved podocyte structure with reduced foot process effacement (Figure 5D). Interestingly, the levels of serum dsDNA antibodies were comparable between CKO and control mice (Figure 5E). Collectively, these findings demonstrate that podocyte-specific *Il23r* deletion ameliorates LN without affecting systemic autoimmunity.

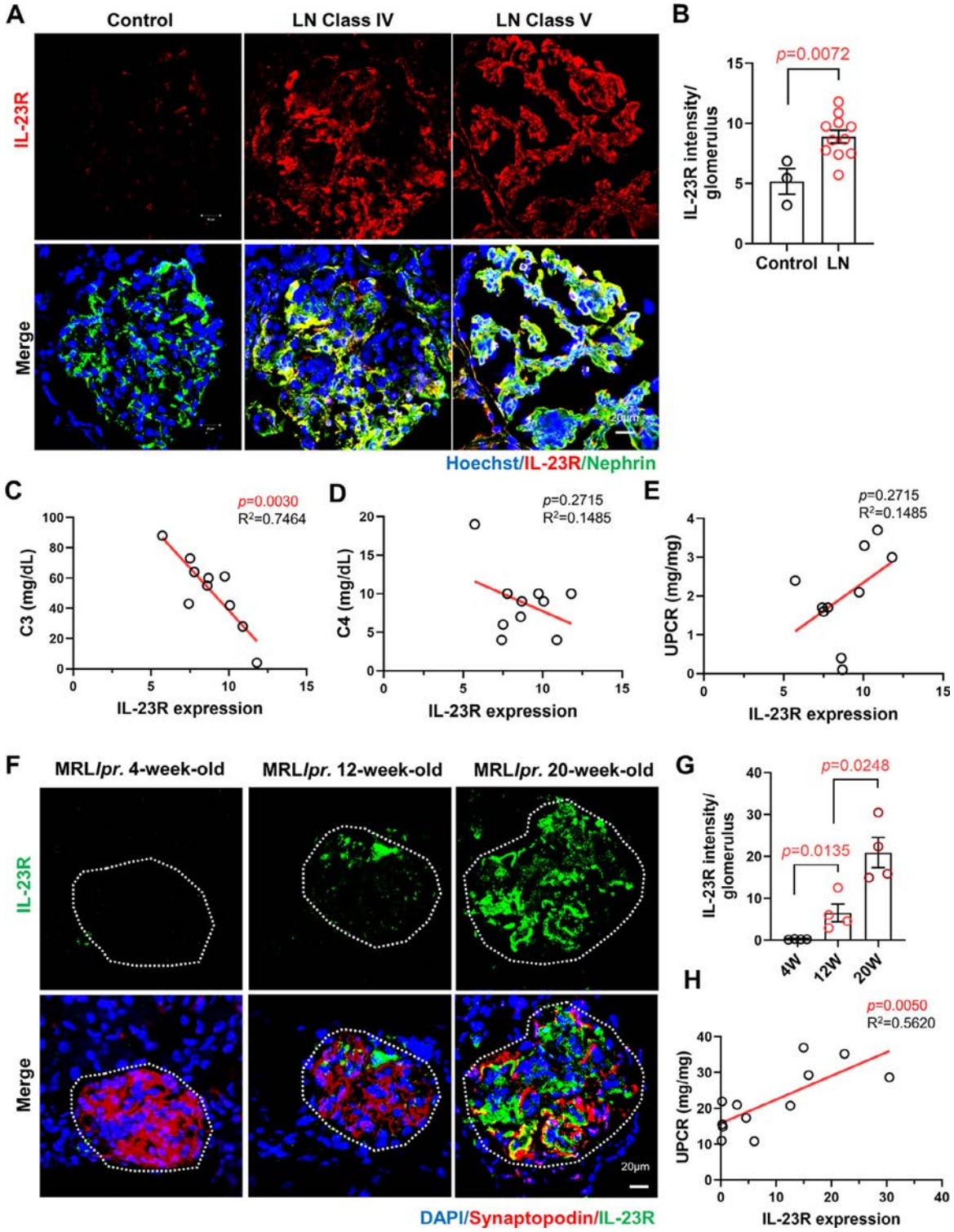


Figure 1. Increased expression of interleukin-23 receptor (IL-23R) in the glomeruli of patients and mice with lupus nephritis (LN). (A) Representative images of human kidney biopsies from patients with LN and controls, stained for IL-23R (red), nephrin (green), and Hoechst (blue). Scale, 20 μm. (B) Quantification plot of IL-23R staining intensity in glomeruli from patients with LN (n = 11) and controls (n = 3). Correlation coefficients between glomerular IL-23R staining intensity and (C) serum complement 3 (C3) levels, (D) serum C4 levels, and (E) urine protein-to-creatinine ratio (UPCR) in patients with LN. Data represent the mean ± SEM. (F) Representative images of kidney sections from 4-week-old (4W), 12W, and 20W female MRL/*lpr* mice stained for IL-23R (green), synaptopodin (red), and DAPI (blue). The white dotted outline delineates the glomerulus. Scale, 20 μm. (G) Quantification plot of IL-23R expression in the glomeruli of MRL/*lpr* mice at different ages (n = 4 per group). (H) Correlation coefficients between glomerular IL-23R staining intensity and UPCR.

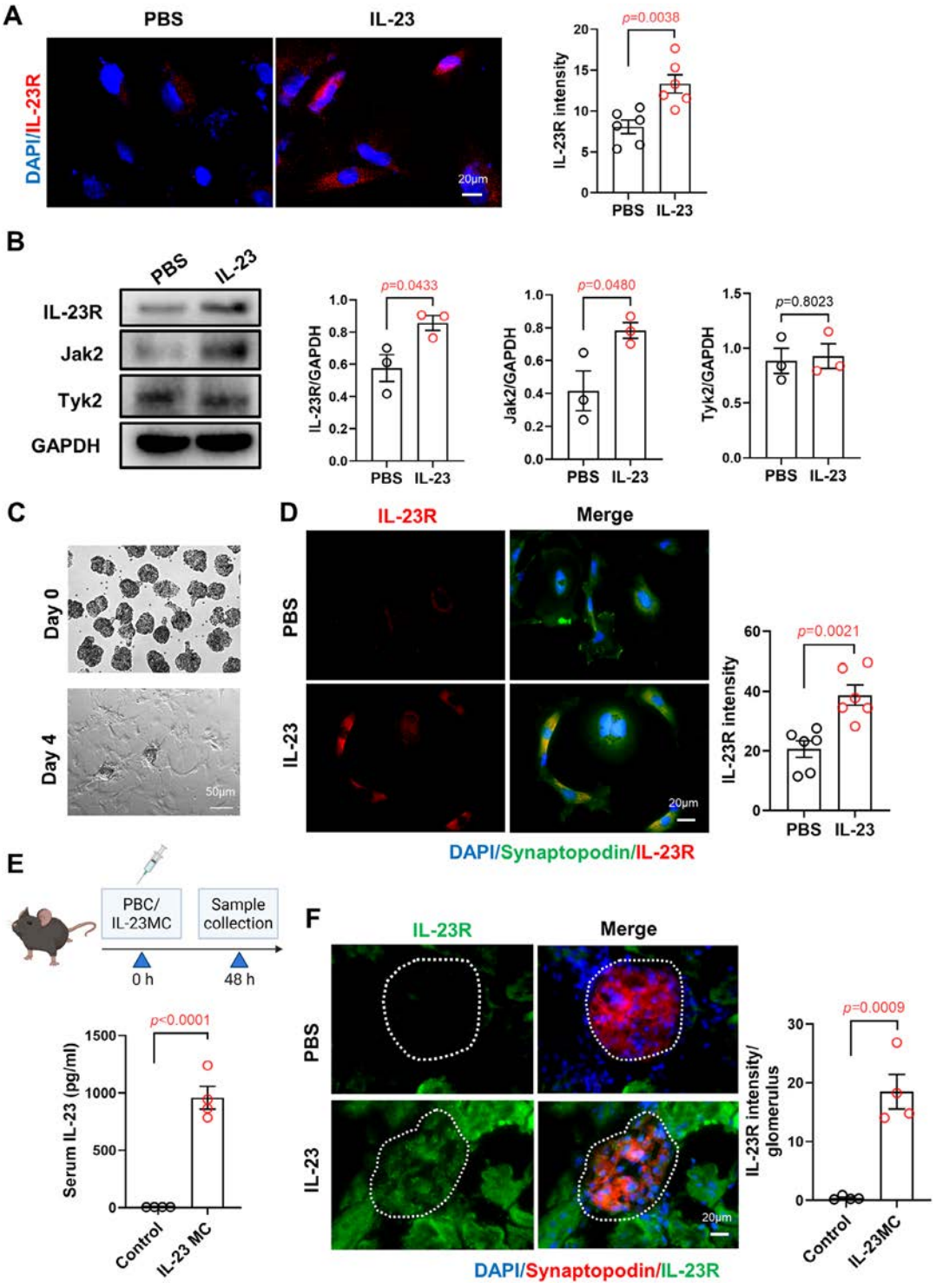
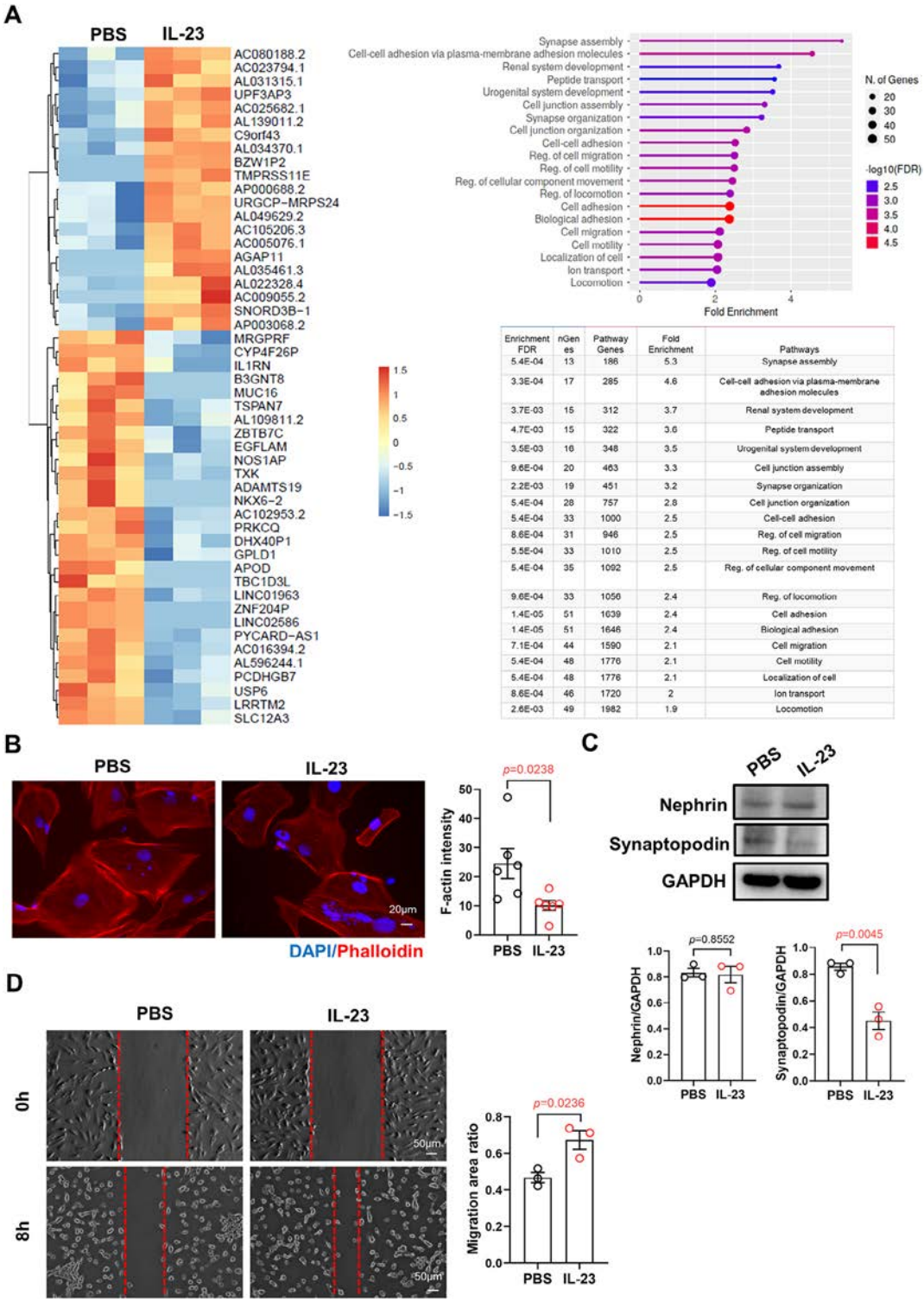


Figure 2. Interleukin-23 receptor (IL-23R) expression in human podocytes after stimulation with IL-23. Human immortalized podocytes were stimulated with phosphate buffered saline (PBS) or recombinant human IL-23 (10 ng/mL) for 24 hours (24 h). (A) Representative immunofluorescence images of IL-23 (red) and quantification plot of staining intensity. Nuclei were counterstained with DAPI (blue). Scale, 20 μm. (B) Representative Western blot and densitometric quantification plot of IL-23R, Jak2, and Tyk2 expression. (C) Primary podocytes were isolated from B6 mice and cultured in vitro for four days. Scale, 50 μm. (D) Representative immunofluorescence images and quantification plot of IL-23R staining intensity in mouse primary podocytes stimulated with PBS or recombinant mouse IL-23 (10 ng/mL) for 24 h. IL-23R (red), synaptopodin (green), and DAPI (blue) are shown. Scale, 20 μm. IL-23 minicircle (MC) was administered to wild-type B6 mice at eight weeks of age via hydrodynamic tail vein injection. (E) Serum IL-23 levels measured by enzyme-linked immunosorbent assay. (F) Representative immunofluorescence images and quantification plot of IL-23R expression in glomeruli from mice injected with IL-23 MC or PBS (n = 4 per group). IL-23R (green), synaptopodin (red), and DAPI (blue) are shown. The white dotted outline delineates the glomerulus. Scale, 20 μm. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.43395/abstract.



Enrichment FDR	nGenes	Pathway Genes	Fold Enrichment	Pathways
5.4E-04	13	186	5.3	Synapse assembly
3.3E-04	17	285	4.6	Cell-cell adhesion via plasma-membrane adhesion molecules
3.7E-03	15	312	3.7	Renal system development
4.7E-03	15	322	3.6	Peptide transport
3.5E-03	16	348	3.5	Urogenital system development
9.6E-04	20	463	3.3	Cell junction assembly
2.2E-03	19	451	3.2	Synapse organization
5.4E-04	28	757	2.8	Cell junction organization
5.4E-04	33	1000	2.5	Cell-cell adhesion
8.6E-04	31	946	2.5	Reg. of cell migration
5.5E-04	33	1010	2.5	Reg. of cell motility
5.4E-04	35	1092	2.5	Reg. of cellular component movement
9.6E-04	33	1056	2.4	Reg. of locomotion
1.4E-05	51	1639	2.4	Cell adhesion
1.4E-05	51	1646	2.4	Biological adhesion
7.1E-04	44	1590	2.1	Cell migration
5.4E-04	48	1776	2.1	Cell motility
5.4E-04	48	1776	2.1	Localization of cell
8.6E-04	46	1720	2	Ion transport
2.6E-03	49	1982	1.9	Locomotion

Figure 3. Interleukin-23 (IL-23) compromises the structure and function of human podocytes. (A) Heatmap of differentially expressed transcripts and enriched pathway analysis of bulk RNA sequencing data from human podocytes stimulated with recombinant human IL-23 or phosphate buffered saline (PBS) for two hours. (B–D) Human immortalized podocytes were stimulated with recombinant human IL-23 (10 ng/mL) for 24 or 8 hours. (B) Representative phalloidin staining and quantification plot of mean actin intensity. Scale, 20 μm. (C) Western blot and quantification plots of nephrin and synaptopodin. (D) Representative images and quantification plots of wound-healing assay showing the area of migration in human podocytes stimulated with PBS or recombinant human IL-23 (10 ng/mL) for eight hours after wounding. Red dashed lines indicate the wound edges. Scale, 50 μm. FDR, false discovery rate; N., number; Reg., regulation. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.43395/abstract.

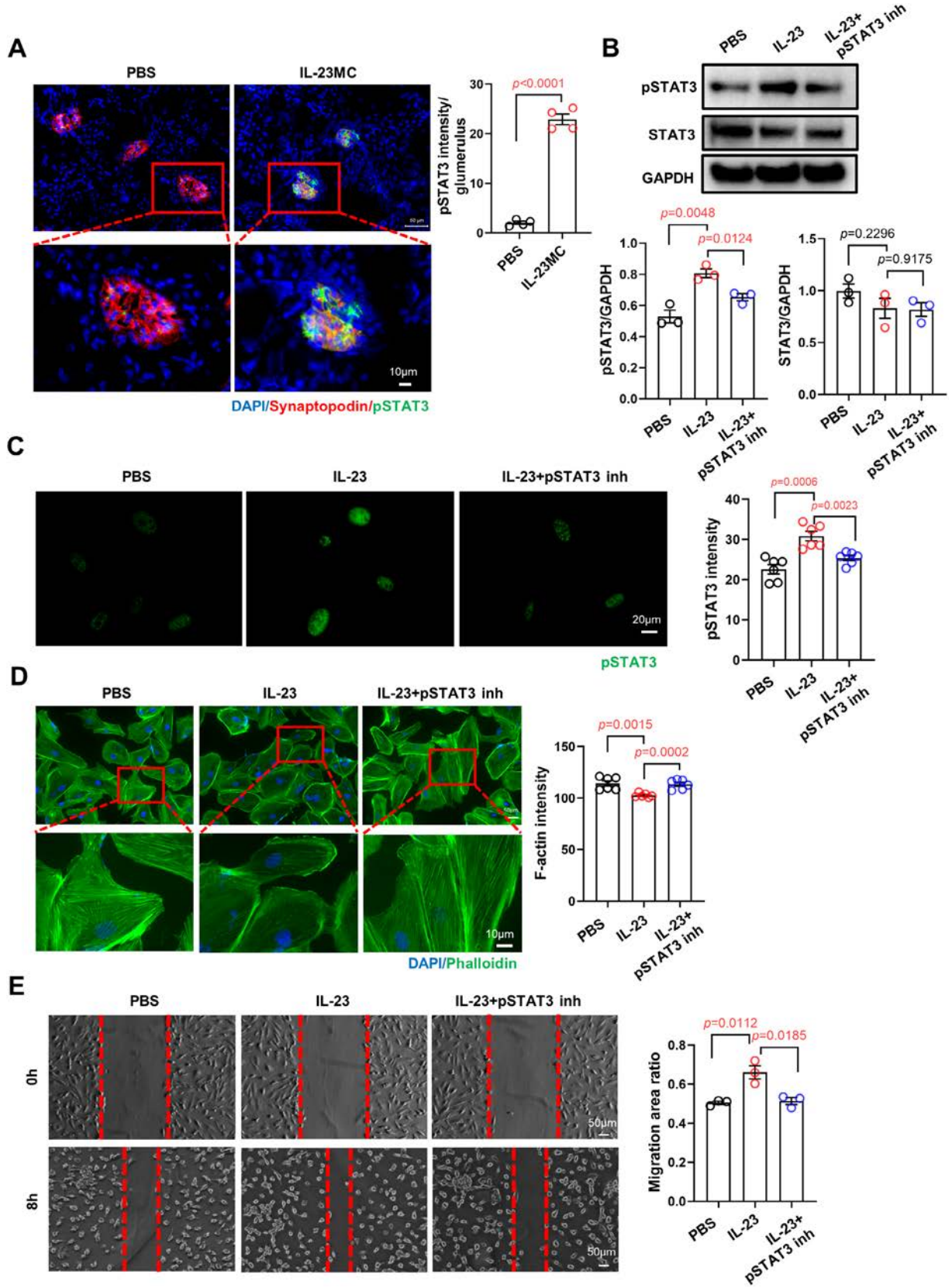


Figure 4. Legend on next page.

DISCUSSION

The expression of immune molecules on podocytes and their role in their demise have been poorly studied.[2] In this study, we demonstrate that IL-23R expression is increased in podocytes from patients with LN and lupus-prone mice and that IL-23 acts directly on podocytes, causing injury. IL-23 firstly induces expression of IL-23R on the surface membrane of podocytes and subsequently activates its downstream signaling partner, pSTAT3, leading to enhanced motility and actin cytoskeleton rearrangement. More importantly, we demonstrate that podocyte-specific deficiency of *Il23r* in lupus-prone mice ameliorates GN.

Podocytes have a complex cellular architecture that is maintained by a well-organized actin cytoskeleton.[34,35] Actin dynamics in podocytes are regulated by various signaling receptors and integrators.[36] Synaptopodin is a cytoskeletal protein that interacts with α-actinin to bundle and elongate actin filaments, and its degradation results in the rearrangement or disruption of the actin cytoskeleton.[37,38] It is essential for the function and structure of podocytes and is down-regulated in podocytes in GN.[6,39] Here, we show that IL-23 disrupts actin fibers and leads to cytoskeletal damage in podocytes by suppressing synaptopodin expression. Podocytes are central to glomerular filtration and remain static and quiescent under physiologic conditions but become motile following injury.[29,40,41] The precise molecular mechanisms by which IL-23 influences the cytoskeletal dynamics in podocytes remain incompletely understood. Although our data suggest IL-23/IL-23R signaling is involved in podocyte actin remodeling, potentially via downstream pathways such as STAT3, further investigations are needed to delineate the involvement of other regulatory components. For instance, small GTPases of the Rho family (eg, RhoA, Rac1, and Cdc42), which are key regulators of cytoskeletal organization,[42] may play a role in mediating IL-23-induced structural changes. Future studies using specific inhibitors or genetic models targeting these small GTPases would help to clarify this mechanism and strengthen the therapeutic potential of IL-23-targeted strategies. Our observations provide rationale for further studies to explore whether pharmacological inhibition of the IL-23/IL-23R pathway can preserve podocyte structure and avert GN.

A central pathologic feature of GN is podocyte foot process effacement, a dynamic process that involves redistribution of the actin cytoskeleton.[40] Of interest is our observation that *Il23r* deficiency in mice who were lupus prone abrogated IC deposition and, more importantly, decreased podocyte foot process effacement, which is critical for maintaining glomerular filtration barrier integrity. Interestingly, despite high levels of circulating anti-dsDNA antibodies, glomerular IC deposition was markedly reduced in mice deficient in IL-23R. This observation suggests that IL-23 signaling in podocytes may play a crucial role in regulating the local environment necessary for IC deposition. One potential mechanism is that IL-23R deficiency may alter the activation status of podocytes, leading to reduced expression of adhesion molecules or receptors involved in IC trapping and clearance. Additionally, IL-23 signaling can modulate the local production of cytokines or chemokines by podocytes, thereby shaping immune cell recruitment or inflammatory cascades within the glomerulus. Another possibility is that the absence of IL-23 signaling influences the glomerular microenvironment in a manner that promotes more efficient clearance of deposited ICs, either through changes in mesangial cell behavior, matrix composition, or podocyte-matrix interactions. These hypotheses highlight an underexplored role of IL-23 in glomerular IC dynamics and warrant further mechanistic investigation. Recently, we demonstrated that genetic deficiency of CaMK4 in the podocytes of lupus-prone mice suppresses the deposition of ICs in the glomeruli, but it increases their deposition in the liver, suggesting that if ICs are not deposited in the kidney, they find their way to other organs.[43]

Recent studies have highlighted the immune phenotype of kidney-resident cells, demonstrating that selective deletion of important molecules in podocytes mitigates autoimmune kidney disease.[2,12,29,43,44] These findings led to further investigation of the interactions between immune cells and kidney-resident cells in patients with LN. Recent data of single-cell spatial transcriptomic profiling from childhood-onset LN showed activation of resident glomerular cell subsets leading to the secretion of proinflammatory chemokines, cytokines, and fibrotic molecules.[45] In our current study, we observed no significant changes in periglomerular or peritubular immune cell infiltration upon podocyte-specific Il23r deletion. Our previous study[12] demonstrated that IL-23 directly activates renal TECs, leading to enhanced glycolysis and CaMK4 up-regulation, which suppresses ARG1 expression and promotes local inflammation. TEC-specific deletion of Il23r or Camk4 attenuated renal inflammation, highlighting the proinflammatory role of IL-23 signaling in TECs and the distinct, cell

Figure 4. Interleukin-23 (IL-23) induces rapid STAT3 activation and nuclear translocation in podocytes. (A) Representative immunofluorescence images and quantification plot of the expression of phosphorylated STAT3 (pSTAT3) in the glomeruli from mice injected IL-23 minicircle (MC) or phosphate buffered saline (PBS; n = 4 per group). Synaptopodin (red), pSTAT3 (green), and DAPI (blue) are shown. Scale, 50 μm. (B–E) Human immortalized podocytes were stimulated with PBS or recombinant human IL-23 (10 ng/mL) or IL-23 plus pSTAT3 inhibitor. (B) Western blots and quantification plots of pSTAT3 and STAT3 after stimulation for one hour (1h). (C) Representative immunofluorescence images and quantification plot of intranuclear pSTAT3 expression (green) after stimulation for 1h. Scale, 20 μm. (D) Representative phalloidin staining images and quantification plot of mean actin intensity after stimulation for 24h. Scale, 10 μm. (E) Representative images of wound-healing assay and quantification plot of the area of podocyte migration after stimulation for 8h. Red dashed lines indicate the wound edges. Scale, 50 μm. inh, inhibitor.

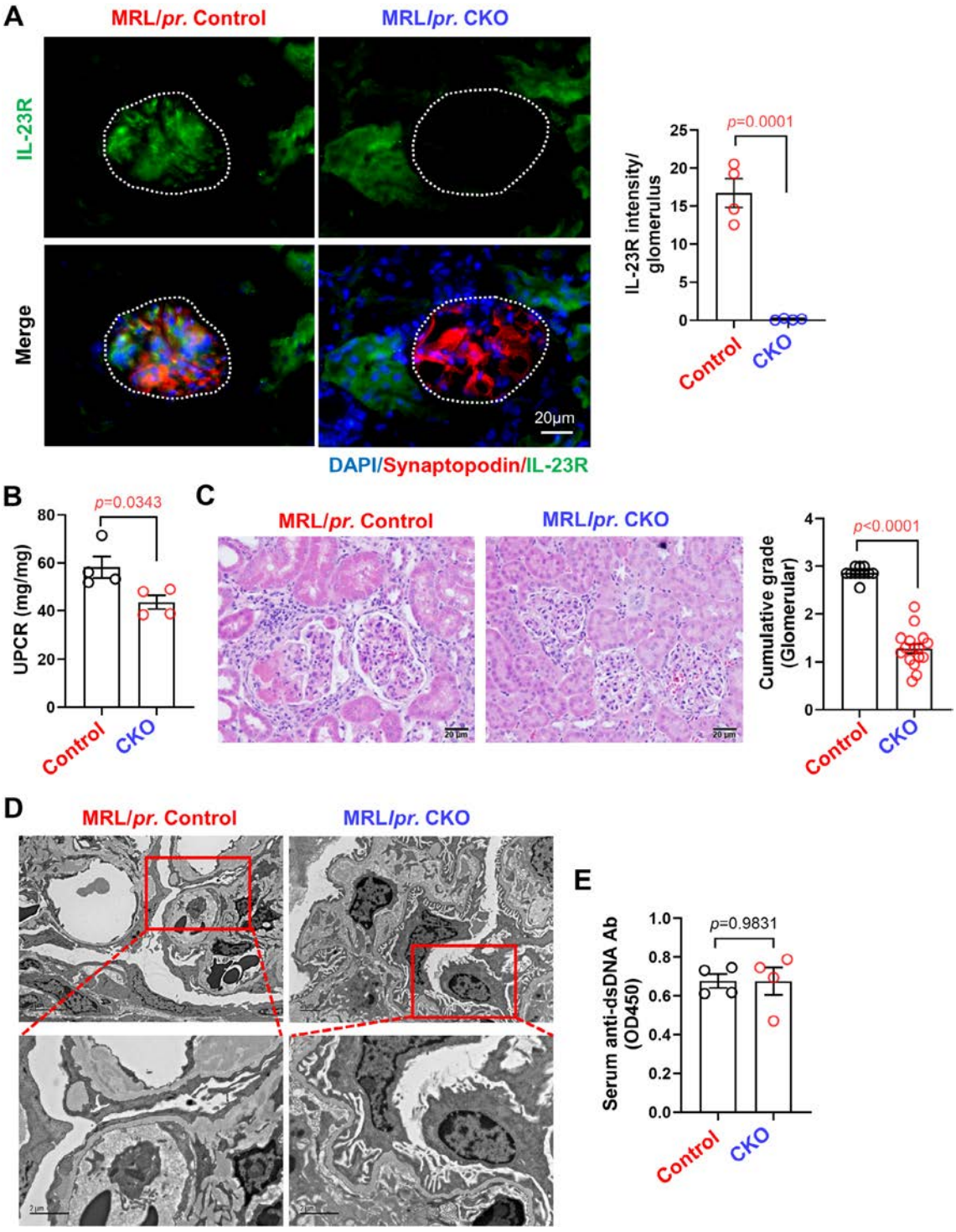


Figure 5. Podocyte-specific deletion of *Il23r* in mice who were lupus prone ameliorates lupus nephritis. MRL/*lpr* mice deficient in podocyte-specific *Il23r* (conditional knockout [CKO]) were euthanized at the ages of 18 to 23 weeks. Age-matched MRL/*lpr Il23*$^{fl/fl}$ or wild-type MRL/lpr mice were used as controls. (A) Representative immunofluorescence images and quantification of the expression of interleukin-23 receptor (IL-23R) in glomeruli from CKO and control mice (n = 4 per group). The white dotted outline delineates the glomerulus. Scale, 20 μm. (B) Urine was collected, and the urine protein-to-creatinine ratio (UPCR) was measured in both groups of mice (n = 4 per group; three independent experiments). (C) Representative images of hematoxylin and eosin staining of kidney sections and cumulative glomerular scores of two independent cohorts of mice aged 18 to 20 weeks and 22 to 23 weeks (n = 11 control mice and n = 16 CKO mice). Scale, 20 μm. (D) Representative electron micrographs of glomeruli from CKO and control mice (scale, 2 μm), with enlarged views of the boxed areas shown below. (E) Serum anti–double-stranded DNA (dsDNA) levels in CKO and control mice (n= 4 per group; three independent experiments). Ab, antibody; OD450, optic density 450 nm. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.43395/abstract.

type–specific effects of IL-23 in kidney pathology. In contrast, the current study focused on podocyte-specific Il23r deletion, which revealed a distinct role for IL-23R in modulating glomerular injury and IC deposition. Taken together, IL-23 signaling exerts cell type–specific effects in the kidney: in TECs, it contributes to interstitial inflammation, whereas in podocytes, it modulates glomerular immune responses and structural integrity.

Immunosuppressive agents are the cornerstone in the treatment of GN but are associated with a high risk of side effects, especially infections.[46–48] Our study presents evidence that corroborates a previous report from our laboratory that podocyte injury and GN may occur independently of systemic inflammation.[29] Thus, targeting IL-23 with specific inhibitors could be of value in the treatment of GN and may represent a safer alternative, with fewer side effects compared to the commonly used cytotoxic drugs.[49] Podocyte-specific delivery of drugs or biologics should offer therapeutic approaches devoid of systemic side effects.

Clinical trials targeting the IL-12/IL-23 axis have failed to show clinical benefit in patients with SLE. Guselkumab, an IL-23p19-subunit inhibitor, did not demonstrate superiority compared to standard-of-care therapy in patients of a phase 2 study (ORCHID-LN) with active LN, although the trial was terminated due to enrollment challenges.[50] A phase 3 clinical study (LOTUS) evaluating the efficacy of ustekinumab, an IL-12/23p40-subunit inhibitor, in patients with active SLE was also discontinued due to lack of expected efficacy following planned futility analysis, despite the promising data of its preceding phase 2 study.[51,52] One possible explanation for the failed clinical trials is that IL-23 is not the central pathogenic driver in all patients with SLE or LN. Our data indicate that IL-23R up-regulation is central for the development of podocyte and kidney injury, implying that patients with SLE who do not up-regulate IL-23R in podocytes may not develop LN. Because it is quite possible that LN may develop through multiple molecular pathways, only patients with elevated IL-23R expression may be the ideal candidates for therapies modulating the IL-23/IL23R axis. Given the molecular heterogeneity of LN, future studies should aim to stratify patients based on podocyte IL-23R expression to identify those most likely to benefit from targeted therapy. New therapeutic agents in development directly targeting IL-23R may also improve the ability to more effectively antagonize this pathway and potentially reveal novel IL-23R-dependent biology.[53]

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr G. C. Tsokos confirms that all authors have provided the final approval of the version to be published and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

ROLE OF THE STUDY SPONSOR

Janssen had no role in the study design or in the collection, analysis, or interpretation of the data, the writing of the manuscript, or the decision to submit the manuscript for publication. Publication of this article was not contingent upon approval by Janssen.

REFERENCES

1. Tsokos GC. Systemic lupus erythematosus. N Engl J Med 2011; 365(22):2110–2121.
2. Tsokos GC, Boulougoura A, Kasinath V, et al. The immunoregulatory roles of non-haematopoietic cells in the kidney. Nat Rev Nephrol 2024;20(4):206–217. Erratum in: *Nat Rev Nephrol* 2024;20:70.
3. Mathieson PW. The podocyte as a target for therapies--new and old. Nat Rev Nephrol 2011;8(1):52–56.
4. Jennette JC. Rapidly progressive crescentic glomerulonephritis. Kidney Int 2003;63(3):1164–1177.
5. Faul C, Asanuma K, Yanagida-Asanuma E, et al. Actin up: regulation of podocyte structure and function by components of the actin cytoskeleton. Trends Cell Biol 2007;17(9):428–437.
6. Faul C, Donnelly M, Merscher-Gomez S, et al. The actin cytoskeleton of kidney podocytes is a direct target of the antiproteinuric effect of cyclosporine A. Nat Med 2008;14(9):931–938.
7. Fornoni A, Sageshima J, Wei C, et al. Rituximab targets podocytes in recurrent focal segmental glomerulosclerosis. Sci Transl Med 2011; 3(85):85ra46.
8. Gaffen SL, Jain R, Garg AV, et al. The IL-23-IL-17 immune axis: from mechanisms to therapeutic testing. Nat Rev Immunol 2014;14(9): 585–600.
9. Teng MW, Bowman EP, McElwee JJ, et al. IL-12 and IL-23 cytokines: from discovery to targeted therapies for immune-mediated inflammatory diseases. Nat Med 2015;21(7):719–729.
10. Li H, Tsokos GC. IL-23/IL-17 axis in inflammatory rheumatic diseases. Clin Rev Allergy Immunol 2021;60(1):31–45.
11. Zhang Z, Kyttaris VC, Tsokos GC. The role of IL-23/IL-17 axis in lupus nephritis. J Immunol 2009;183(5):3160–3169.
12. Li H, Tsokos MG, Bhargava R, et al. IL-23 reshapes kidney resident cell metabolism and promotes local kidney inflammation. J Clin Invest 2021;131(12):e142428.
13. Dai H, He F, Tsokos GC, et al. IL-23 limits the production of IL-2 and promotes autoimmunity in lupus. J Immunol 2017;199(3):903–910.
14. Kyttaris VC, Zhang Z, Kuchroo VK, et al. Cutting edge: IL-23 receptor deficiency prevents the development of lupus nephritis in C57BL/6-lpr/lpr mice. J Immunol 2010;184(9):4605–4609.
15. Paust HJ, Turner JE, Steinmetz OM, et al. The IL-23/Th17 axis contributes to renal injury in experimental glomerulonephritis. J Am Soc Nephrol 2009;20(5):969–979.
16. Ooi JD, Phoon RK, Holdsworth SR, et al. IL-23, not IL-12, directs autoimmunity to the Goodpasture antigen. J Am Soc Nephrol 2009; 20(5):980–989.
17. Kyttaris VC, Kampagianni O, Tsokos GC. Treatment with anti-interleukin 23 antibody ameliorates disease in lupus-prone mice. Biomed Res Int 2013;2013:861028.
18. Teng TS, Lin B, Manser E, et al. Stat3 promotes directional cell migration by regulating Rac1 activity via its activator betaPIX. J Cell Sci 2009;122(Pt 22):4150–4159.
19. Parrini MC, Camonis J. Cell motility: the necessity of Rac1 GDP/GTP flux. Commun Integr Biol 2011;4(6):772–774.

20. Dai Y, Gu L, Yuan W, et al. Podocyte-specific deletion of signal transducer and activator of transcription 3 attenuates nephrotoxic serum-induced glomerulonephritis. Kidney Int 2013;84(5):950–961.
21. He JC, Husain M, Sunamoto M, et al. Nef stimulates proliferation of glomerular podocytes through activation of Src-dependent Stat3 and MAPK1,2 pathways. J Clin Invest 2004;114(5):643–651.
22. Feng X, Lu TC, Chuang PY, et al. Reduction of Stat3 activity attenuates HIV-induced kidney injury. J Am Soc Nephrol 2009;20(10): 2138–2146.
23. Saleem MA, O'Hare MJ, Reiser J, et al. A conditionally immortalized human podocyte cell line demonstrating nephrin and podocin expression. J Am Soc Nephrol 2002;13(3):630–638.
24. Satyam A, Tsokos MG, Tresback JS, et al. Cell derived extracellular matrix-rich biomimetic substrate supports podocyte proliferation, differentiation and maintenance of native phenotype. Adv Funct Mater 2020;30(44):1908752.
25. Takemoto M, Asker N, Gerhardt H, et al. A new method for large scale isolation of kidney glomeruli from mice. Am J Pathol 2002;161(3): 799–805.
26. Sung SJ, Ge Y, Dai C, et al. Dependence of glomerulonephritis induction on novel intraglomerular alternatively activated bone marrow-derived macrophages and Mac-1 and PD-L1 in lupus-prone NZM2328 mice. J Immunol 2017;198(7):2589–2601.
27. Aden K, Rehman A, Falk-Paulsen M, et al. Epithelial IL-23R signaling licenses protective IL-22 responses in intestinal inflammation. Cell Rep 2016;16(8):2208–2218.
28. Kikawada E, Lenda DM, Kelley VR. IL-12 deficiency in MRL-Fas(lpr) mice delays nephritis and intrarenal IFN-gamma expression, and diminishes systemic pathology. J Immunol 2003;170(7):3915–3925.
29. Maeda K, Otomo K, Yoshida N, et al. CaMK4 compromises podocyte function in autoimmune and nonautoimmune kidney disease. J Clin Invest 2018;128(8):3445–3459.
30. Poholek CH, Raphael I, Wu D, et al. Noncanonical STAT3 activity sustains pathogenic Th17 proliferation and cytokine response to antigen. J Exp Med 2020;217(10):e20191761.
31. McGeachy MJ, Chen Y, Tato CM, et al. The interleukin 23 receptor is essential for the terminal differentiation of interleukin 17-producing effector T helper cells in vivo. Nat Immunol 2009;10(3):314–324.
32. Chen Z, Laurence A, Kanno Y, et al. Selective regulatory function of Socs3 in the formation of IL-17-secreting T cells. Proc Natl Acad Sci U S A 2006;103(21):8137–8142.
33. Vukelic M, Laloo A, Kyttaris VC. Interleukin 23 is elevated in the serum of patients with SLE. Lupus 2020;29(14):1943–1947.
34. Scott RP, Quaggin SE. Review series: the cell biology of renal filtration. J Cell Biol 2015;209(2):199–210.
35. Perico L, Conti S, Benigni A, et al. Podocyte-actin dynamics in health and disease. Nat Rev Nephrol 2016;12(11):692–710.
36. Blaine J, Dylewski J. Regulation of the actin cytoskeleton in podocytes. Cells 2020;9(7):1700.
37. Asanuma K, Yanagida-Asanuma E, Faul C, et al. Synaptopodin orchestrates actin organization and cell motility via regulation of RhoA signalling. Nat Cell Biol 2006;8(5):485–491.
38. Yu H, Kistler A, Faridi MH, et al. Synaptopodin limits TRPC6 podocyte surface expression and attenuates proteinuria. J Am Soc Nephrol 2016;27(11):3308–3319.
39. Tian X, Ishibe S. Targeting the podocyte cytoskeleton: from pathogenesis to therapy in proteinuric kidney disease. Nephrol Dial Transplant 2016;31(10):1577–1583.
40. Torban E, Braun F, Wanner N, et al. From podocyte biology to novel cures for glomerular disease. Kidney Int 2019;96(4):850–861.
41. Brahler S, Yu H, Suleiman H, et al. Intravital and kidney slice imaging of podocyte membrane dynamics. J Am Soc Nephrol 2016;27(11): 3285–3290.
42. Matsuda J, Asano-Matsuda K, Kitzler TM, et al. Rho GTPase regulatory proteins in podocytes. Kidney Int 2021;99(2):336–345.
43. Bhargava R, Li H, Maeda K, et al. Calcium calmodulin kinase IV deficiency in podocytes prevents the development of lupus nephritis. Clin Immunol 2025;271:110427.
44. Umeda M, Karino K, Satyam A, et al. Hypoxia promotes the expression of ADAM9 by tubular epithelial cells, which enhances transforming growth factor beta1 activation and promotes tissue fibrosis in patients with lupus nephritis. Arthritis Rheumatol 2025;77(2): 180–189.
45. Danaher P, Hasle N, Nguyen ED, et al. Childhood-onset lupus nephritis is characterized by complex interactions between kidney stroma and infiltrating immune cells. Sci Transl Med 2024;16(775):eadl1666.
46. Tian J, Luo Y, Wu H, et al. Risk of adverse events from different drugs for SLE: a systematic review and network meta-analysis. Lupus Sci Med 2018;5(1):e000253.
47. Bultink IEM, de Vries F, van Vollenhoven RF, et al. Mortality, causes of death and influence of medication use in patients with systemic lupus erythematosus vs matched controls. Rheumatology (Oxford) 2021; 60(1):207–216.
48. Kunzler ALF, Tsokos GC. Infections in patients with systemic lupus erythematosus: the contribution of primary immune defects versus treatment-induced immunosuppression. Eur J Rheumatol 2023; 10(4):148–158.
49. Ru Y, Ding X, Luo Y, et al. Adverse events associated with anti-IL-23 agents: clinical evidence and possible mechanisms. Front Immunol 2021;12:670398.
50. Anders HJ, Chan TM, Sanchez-Guerrero J, et al. Efficacy and safety of guselkumab in patients with active lupus nephritis: results from a phase 2, randomized, placebo-controlled study. Rheumatology (Oxford) 2025;64(5):2731–2740.
51. van Vollenhoven RF, Kalunian KC, Dorner T, et al. Phase 3, multicentre, randomised, placebo-controlled study evaluating the efficacy and safety of ustekinumab in patients with systemic lupus erythematosus. Ann Rheum Dis 2022;81(11):1556–1563.
52. van Vollenhoven RF, Hahn BH, Tsokos GC, et al. Efficacy and safety of ustekinumab, an IL-12 and IL-23 inhibitor, in patients with active systemic lupus erythematosus: results of a multicentre, double-blind, phase 2, randomised, controlled study. Lancet 2018;392(10155): 1330–1339.
53. Bissonnette R, Pinter A, Ferris LK, et al. An oral interleukin-23-receptor antagonist peptide for plaque psoriasis. N Engl J Med 2024;390(6):510–521.

Arthritis & Rheumatology
Vol. 78, No. 3, March 2026, pp 684–694
DOI 10.1002/art.43417

American College of Rheumatology
Empowering Rheumatology Professionals

Efficacy of Telitacicept in Childhood-Onset Systemic Lupus Erythematosus: A Prospective Multicenter Cohort Study With Inverse Probability of Treatment Weighting–Adjusted Comparison to a Historical Control Group Treated With Belimumab

Chong Luo,[1,2] Shu Su,[3] Jingyue Liu,[4] Junyu Lin,[5] Jin Song,[6] Zongwen Chen,[7] Yuanyuan Peng,[8] Lijun Jiang,[9] Meng Jia,[3] Li Wang,[1,2] Xinhui Jiang,[6] Wei Zhang,[8] Zanhua Rong,[9] Bo Zhao,[5] Yajun Wang,[4] and Xuemei Tang[1,2]

Objective. The therapeutic effects of telitacicept combined with standard of care (SoC) in childhood-onset systemic lupus erythematosus (cSLE) remain unclear. This study aims to evaluate its efficacy in comparison with belimumab combined with SoC.

Methods. We performed a prospective cohort study across seven tertiary hospitals in China, enrolling patients aged 5 to 18 years with cSLE. Primary endpoints included the proportions of patients attaining Lupus Low Disease Activity State (LLDAS) and Definition of Remission in SLE (DORIS) at 3, 6, and 12 months after treatment, complemented by relapse rates and steroid tapering metrics. Propensity score–based inverse probability of treatment weighting (IPTW) was employed to address baseline imbalances between the telitacicept group (n = 60) and the historical control group treated with belimumab (n = 67).

Results. Among 60 enrolled patients (median follow-up 12 [interquartile range 6–12] months), LLDAS and DORIS achievement rates progressively increased from 19.6% and 7.8% at 6 months to 65.7% and 37.1% at 12 months. Flare rates remained low (3 months: 1.7%; 6 months: 11.8%; 12 months: 14.3%). Notably, 74.3% of patients achieved prednisone doses ≤7.5 mg/day by 12 months (vs baseline 13.3%, $P < 0.0001$). IPTW-adjusted analyses demonstrated comparable efficacy between the telitacicept and belimumab groups across all endpoints in terms of LLDAS and DORIS attainment, steroid reduction, and immunologic normalization (all $P > 0.05$).

Conclusion. Telitacicept combined with SoC exhibited significant improvements in disease activity, steroid reduction, and immunologic markers, with therapeutic outcomes comparable to belimumab in cSLE management.

Supported by National Key R&D Program of China (2021YFC2702000) and Kuanren Talent programs of the Second Affiliated Hospital of Chongqing Medical University (202417-16).

[1]Chong Luo, PhD, Li Wang, MD, Xuemei Tang, MD: Department of Rheumatology and Immunology, Children's Hospital of Chongqing Medical University, National Clinical Research Center for Child Health and Disorders, Ministry of Education Key Laboratory of Child Development and Disorders, Chongqing, China; [2]Chong Luo, PhD, Li Wang, MD, Xuemei Tang, MD: Chongqing Key Laboratory of Child Rare Diseases in Infection and Immunity, Chongqing, China; [3]Shu Su, PhD, Meng Jia, MS: Department of Epidemiology and Biostatistics, The Second Affiliated Hospital of Chongqing Medical University, Chongqing, China; [4]Jingyue Liu, MS, Yajun Wang, MD: Department of Pediatrics, The First People's Hospital of Yunnan Province, Kunming, China; [5]Junyu Lin, MSc, Bo Zhao, MD: Department of Nephrology & Rheumatology, The Affiliated Children's Hospital of Kunming Medical University, Kunming Children's Hospital, Kunming, China; [6]Jin Song, MS, Xinhui Jiang, BS: Department of Rheumatology Immunization, Guiyang Maternal and Child Health Care Hospital, Guiyang, China; [7]Zongwen Chen, MS: Department of Comprehensive Pediatric Internal Medicine, Chongqing University Three Gorges Hospital, Chongqing, China; [8]Yuanyuan Peng, BS, Wei Zhang, PhD: Pediatric Immunology and Rheumatology Department, Chengdu Women's and Children's Central Hospital, School of Medicine, University of Electronic Science and Technology of China, Chengdu, China; [9]Lijun Jiang, MD, Zanhua Rong, MD: Department of Pediatrics, the Second Hospital of Hebei Medical University, Shijiazhuang.

Drs Luo, Su, and Liu are co-first authors and contributed equally to the work; Drs Zhao, Yajun Wang, and Tang are co-last authors and contributed equally to this work.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/art.43417).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/art.43417.

Address correspondence via email to Bo Zhao, MD, at zbkmetyy@163.com; or to Yajun Wang, MD, at wyj6371@sina.com; or to Xuemei Tang, MD, at tangxuemei2008@163.com.

Submitted for publication May 14, 2025; accepted in revised form September 22, 2025.

INTRODUCTION

Systemic lupus erythematosus (SLE) is a chronic, multisystem autoimmune disorder marked by recurrent episodes of remission and exacerbation, often resulting in progressive organ damage. Childhood-onset SLE (cSLE), defined as disease onset before 18 years of age, has a reported global prevalence of approximately 1.89 to 25.70 per 100,000 children.[1,2] An estimated 60,000 children in China are affected. Clinically, cSLE demonstrates more aggressive disease course than adult-onset SLE (aSLE), characterized by rapid disease progression, greater severity, and a higher frequency of renal and neuropsychiatric involvement.

The management of cSLE is complicated by the vulnerability of children undergoing critical phases of growth and development. Traditional treatments, such as steroids and immunosuppressants, may have adverse effects on bone development, body image, and growth potential. Prolonged glucocorticoid exposure in children further increases the risk of complications such as metabolic syndrome and cataract formation. These limitations underscore the urgent need for therapeutic strategies that are both effective and better tolerated. In recent years, biologics agents have expanded the treatment landscape, offering promising alternatives for aSLE and the potential to mitigate treatment-related toxicity in children with cSLE.[3]

Telitacicept is a novel recombinant fusion protein that comprises the extracellular domain of the B cell activating factor receptor, transmembrane activator, and CAML interactor (TACI) with an Fc fragment. TACI receptor exhibits high affinity for B lymphocyte stimulator (BLyS) and a proliferation-inducing ligand (APRIL). Telitacicept can competitively bind these two cytokines and block their interaction with B cell membrane receptors (eg, BCMA, BR3, TACI), inhibiting the development and survival of mature B cells and plasma cells, there at achieving dual-target treatment for SLE.[4] Several clinical studies have demonstrated the safety and efficacy of telitacicept in the treatment of aSLE.[5–7] Its subcutaneous route of administration offers practical advantages for children with SLE by reducing the need for hospital visits and admissions, thereby facilitating reintegration into school and life. This multicenter, prospective study aims to evaluate the efficacy of telitacicept in children with cSLE, with comparisons made against a historical cohort treated with belimumab which those centers took part in.[8]

MATERIALS AND METHODS

Study population. This multicenter, observational, prospective cohort study (MR-50-23-016168) was conducted between January 2022 and October 2024 across seven hospitals in China, including four children's hospitals and three pediatric departments of general hospital. Eligible participants were patients aged 5 to 18 years with a confirmed diagnosis of cSLE, receiving telitacicept in combination with standard of care (SoC), including glucocorticoids and immunosuppressive agents such as cyclophosphamide, mycophenolate mofetil, and tacrolimus. Telitacicept was administered subcutaneously once weekly at a dose of 80 mg for patients with a baseline bodyweight <50 kg and 160 mg for those ≥50 kg. Follow-up assessments were conducted at 3, 6, and 12 months during the study period. The study aimed to analyze the clinical data of telitacicept combined with SoC in treating cSLE to determine the efficacy of telitacicept and its impact on steroid reduction. Glucocorticoid tapering regimen in this study recommended that the prednisone dose was tapered by 2.5 to 5 mg/day every two to four weeks during the tapering phase. The tapering regimen for each patient was still be made by the treating physician. Participating physicians did not receive any financial incentives. All patient data were anonymized and kept confidential.

Inclusion and exclusion criteria. Eligible participants were pediatric patients meeting all of the following inclusion criteria: (1) aged 5 to 18 years; (2) fulfilled the diagnostic criteria for SLE based on the 1997 American College of Rheumatology (ACR) criteria and/or the 2019 EULAR/ACR classification criteria;[9,10] (3) had not achieved low disease activity with SoC alone; and (4) initiated subcutaneous telitacicept therapy as an adjunct to ongoing SoC regimen, based on rheumatologist's assessment for either disease activity control or glucocorticoid-sparing purposes. The exclusion criteria were as follows: (1) patients who had achieved Lupus Low Disease Activity State (LLDAS) or Definition of Remission in SLE (DORIS) with SoC therapy at baseline; (2) treatment duration with telitacicept of less than three months; (3) insufficient follow-up data (less than three months); or (4) prior exposure to B cell biologics within three months before enrollment.

Data collection. We collected demographic information, disease duration at enrollment, SLE-related manifestations, and organ involvement for each patient. Additionally, data on the Systemic Lupus Erythematosus Disease Activity Index 2000 (SLEDAI-2K),[11] Patient Global Assessment (PGA; scale 0–3),[12] and disease flares were collected at each visit. Medication records included detailed information on glucocorticoid regimens (converted to prednisone equivalent doses), hydroxychloroquine use, and immunosuppressant prescriptions, capturing both usage patterns and dosage adjustments. Laboratory evaluations encompassed: complete blood counts, hepatic and renal function, urinalysis or 24-hour urine protein measurements, lymphocyte subset analyses, Ig levels (IgG, IgM), complement components (C3, C4), and autoantibody profiles. Baseline was defined as the first enrollment in the cohort, and there were two to five days for further assessment before telitacicept initiation. Missing or unclear data were verified with the treating physician based on the patient's medical records, and any missing data were documented.

Ethical considerations. This study was conducted in accordance with the Helsinki Declaration, and informed consent was obtained from all participants. The study protocol was

approved by the ethics committee of the Children's Hospital of Chongqing Medical University (approval number: 2022/02). Written informed consent was obtained from the legal guardians of all participating children.

Endpoints. The primary endpoints were the proportions of patients with cSLE achieving LLDAS and DORIS at 3, 6, and 12 months after treatment with telitacicept combined with SoC. LLDAS was defined as meeting the following five criteria: (1) SLEDAI-2K ≤4, with no evidence of active disease in any major organ; (2) no new disease activity features compared to the previous assessment; (3) PGA ≤1; (4) daily prednisone equivalent dose ≤7.5 mg; and (5) standard maintenance doses of immunosuppressants and approved biologics were allowed.[13] DORIS was defined using the following criteria: (1) clinical SLEDAI-2K = 0 (not including serological markers anti–double stranded DNA (dsDNA) and complement C3); (2) PGA < 0.5; (3) daily prednisone equivalent dose ≤5 mg; and (4) standard maintenance doses of immunosuppressants and biologics were allowed.[14]

The secondary endpoints were as follows: (1) improvement in SLEDAI-2K categorized as mild (≤6), moderate (7–12), or severe (>12) disease activity; (2) PGA and SLEDAI-2K at each timepoint, along with the flare index, stratified into mild/moderate or severe flares[15]; and (3) evaluation of prednisone equivalent glucocorticoid doses, serological markers, and immunoglobulin levels. All adverse events were systematically recorded and reported throughout the study.

Historical control population treated with belimumab. Full details of the belimumab study from January 2021 to January 2024 have been reported.[8] The telitacicept group was compared with the data of the above centers (Children's Hospital of Chongqing Medical University and The First People's Hospital of Yunnan Province) from the historical belimumab study (Supplementary Figure 1). Baseline clinical characteristics of 67 patients with cSLE from the historical belimumab study are presented in Table 1. SLEDAI-2K scores of the 67 patients with cSLE from the historical belimumab group were evaluated and included in the current analysis.

Statistical analysis. Descriptive methods were used to analyze baseline measurements (age and sex) and clinical manifestations. Continuous variables with normal distribution were expressed as mean ± SD, and comparisons between groups were made using the independent samples *t*-test. Nonnormally distributed data were expressed as median (interquartile range [IQR]) and analyzed using rank-sum tests. Categorical variables were described using percentages and analyzed using the chi-square test, corrected chi-square test, or Fisher's exact test.

Missing data were imputed using random forest prediction models, with error rates calculated using the out-of-bag method. Propensity score–based inverse probability of treatment weighting (IPTW) was then applied to construct a weighted cohort of patients treated with telitacicept or belimumab. To calculate the inverse probability of treatment weights, each patient's propensity to receive telitacicept or belimumab was estimated using a logistic regression model that included baseline confounding variables. These variables were partially merged to improve the fit of

Table 1. Clinical characteristics of patients with cSLE in the telitacicept and belimumab groups at baseline*

Variables	Telitacicept	Belimumab
N	60	67
Age, mean ± SD, y	13.00 ± 2.76	11.43 ± 2.88
Female, n (%)	52 (86.7)	62 (92.5)
Age at onset, mean ± SD, y	10.81 ± 2.58	9.85 ± 3.14
Disease duration, median (IQR), m	17 (1.75–31.25)	6.25 (2–24)
Weight, median (IQR), kg	40.75 (33.38–51.13)	37.5 (26.5–44.75)
Clinical manifestations, n (%)		
Fever	11 (18.3)	19 (28.4)
Mucocutaneous	31 (51.7)	34 (50.7)
Renal[16]	35 (58.3)	37 (55.2)
Massive proteinuria	13 (21.7)	6 (9.0)
Musculoskeletal	12 (20.0)	16 (23.9)
Neurologic	7 (11.7)	12 (17.9)
Cardiopulmonary	16 (26.7)	22 (32.8)
Hematologic	22 (36.7)	25 (37.3)
Leukopenia	14 (23.3)	15 (22.4)
Thrombocytopenia	5 (8.3)	12 (17.9)
Hemolytic anemia	5 (8.3)	7 (10.4)
Elevated ALT/AST	5 (8.3)/3 (5.0)	10 (14.9)/10 (14.9)
Elevated BUN/creatine	13 (21.7)/8 (13.3)	8 (11.9)/4 (6.0)
Serologic test, n (%)		
Positive ANA	57 (95.0)	66 (98.5)
Positive anti-dsDNA	40 (66.7)	43 (64.2)
Hypocomplementemia3	47 (78.3)	57 (85.1)
Hypocomplementemia4	30 (50.0)	44 (65.7)
SLE severity, n (%)		
SLEDAI-2K, median (IQR)	12 (7, 16)	10 (6, 13)
≤6 (mild activity), n (%)	14 (23.3)	19 (28.4)
7–12 (moderate activity), n (%)	12 (20.0)	29 (43.3)
>12 (severe activity), n (%)	34 (56.7)	19 (28.4)
≥10, n (%)	38 (63.3)	39 (58.2)
PGA, median (IQR)	2 (1.5–2.5)	2 (1.5–2.5)
Treatment, n (%)		
Steroids	60 (100.0)	67 (100.0)
Dosage, mean ± SD, mg/d	31.5 ± 19.5	30.2 ± 15.3
Hydroxychloroquine	58 (96.7)	60 (89.6)
Cyclophosphamide	17 (28.3)	15 (22.4)
Mycophenolate mofetil	37 (61.7)	39 (58.2)
Tacrolimus	11 (18.3)	2 (3.0)
MTX	2 (3.3)	1 (1.5)
CSA	1 (1.6)	3 (4.5)
Telitacicept treatment, 80 mg per week/160 mg per week	44 (73.3)/16 (26.7)	–

* Renal means pediatric lupus nephritis.[16] ALT/AST, alanine transaminase/aspartate transaminase; ANA, antinuclear antibodies; BUN, blood urea nitrogen; CSA, cyclosporine; cSLE, childhood-onset SLE; dsDNA, double-stranded DNA; IQR, interquartile range; MTX, methotrexate, PGA, Patient Global Assessment; SLE, systemic lupus erythematosus; SLEDAI-2K, Systemic Lupus Erythematosus Disease Activity Index 2000.

the analytical models. Patients who received telitacicept were assigned a weight of 1/(propensity score), whereas those who received belimumab were assigned a weight of 1/(1 – propensity score). Weights were stabilized and trimmed at the 0.25th and 99.75th percentiles to reduce variability in the estimates. Balance among covariates was assessed at the standardized mean difference (SMD), with a difference of 0.1 or less considered to indicate good balance. The adjusted generalized SE inflation factor was calculated to test for multicollinearity among covariates. Covariates with residual imbalance after IPTW (SMD > 0.1) were controlled by including them as covariates in the Cox model. In the final analyses, multiple outcomes for comparing treatment effects between telitacicept and belimumab were assessed using regression analyses. The association between binary outcome variables and drug use were evaluated using Cox proportional hazards model to calculate hazard ratios and corresponding 95% confidence intervals (CIs). Continuous outcome variables were analyzed using linear mixed-effects mode to calculate the regression coefficient and corresponding 95% CI. Potential confounding factors were adjusted for in subsequent multivariate analyses to identify independent risk factors. R version 4.5.0 (R Project for Statistical Computing) software was used for statistical analysis.

RESULTS

Patient characteristics. The enrollment flow chart is presented in Supplementary Figure 2. In the prospective telitacicept study cohort, a total of 60 patients with cSLE with 195 visits were enrolled between March 2022 and October 2024. As shown in Table 1, the cohort had a mean ± SD age of 13 ± 2.76 years at enrollment, with a median disease duration of 17 months (IQR 1.75–31.25). The median follow-up duration was 12 months (IQR 6–12), and 35 patients completed 12 months of telitacicept treatment. All participants were receiving concomitant steroid therapy at baseline, with a mean ± SD prednisone equivalent dose of 30 ±19.5 mg/day.

Treatment outcomes with telitacicept combined with SoC. *Treatment efficacy.* Median SLEDAI-2K score significantly declined over time, from 12 (IQR 7–16) at baseline to 4 (IQR 2–6) at 3 months ($P < 0.001$) and 3 (IQR 0–4) at 12 months ($P < 0.001$). Proportion of patients with mild diseases activity (SLEDAI-2K ≤ 6) significantly increased from 23.3% at baseline to 82.9% at 12 months ($P < 0.001$), whereas those with severe disease activity (SLEDAI-2K > 12) significantly declined from 46.7% to 5.7% ($P < 0.001$; Figure 1A). Similarly, the median PGA declined from 2 (IQR 1.5–2.5) at baseline to 0.9 (IQR 0–1) at 6 months ($P < 0.001$) and 0 (IQR 0–0.875) at 12 months ($P < 0.001$). By 6 and 12 months, 19.6% and 65.7% of patients achieved LLDAS, whereas 7.8% and 37.1% attained DORIS, respectively (Supplementary Figure 3). During follow-up, five mild and seven severe relapses occurred. The flare risks were 1.7% (1 of 60) at 3 months, 11.8% (6 of 51) at 6 months, and 14.3% (5 of 35) at 12 months.

Improvement in clinical manifestation and immunologic markers. Of the nine patients presenting with thrombocytopenia and/or anemia at baseline, all achieved normalized platelet counts by three months. Hemoglobin levels progressively improved, surpassing 90 g/L by six months and remained stable within the 100–120 g/L range thereafter. Among the seven patients presenting with neurologic involvement, clinical manifestations comprised headache in five individuals (two of whom also experienced blurred vision), convulsions accompanied by impaired consciousness in one patient, and hand numbness in one patient. All patients achieved complete neurologic recovery within three months of treatment. At baseline, 26 patients exhibited 24-hour urine protein levels >500 mg. By 12 months, 58.8% of patients achieved a >50% reduction in proteinuria, with levels below 500 mg/day. During the 12-month follow-up period, all patients with baseline renal dysfunction demonstrated complete functional recovery, maintaining normal renal parameters throughout the 12-month follow-up period.

The proportion of patients with positive anti-dsDNA declined from 66.7% at baseline to 44.1% at 12 months ($P < 0.05$). C3 normalization rates increased from 21.7% at baseline to 54.9% by 6 months, stabilizing at 57.1% at 12 months ($P < 0.001$). C4 normalization increased from 50% at baseline to 74.5% at 6 months, further improving 80% by 12 months ($P < 0.01$). Both IgG and IgM levels declined significantly from baseline to three months ($P < 0.001$). Notably, IgG levels exhibited a modest but statistically significant rebound between 6 months and 12 months ($P < 0.05$; Supplementary Figure 4).

Steroid reduction. The median prednisone equivalent dose deceased significantly from 30.0 mg/day (IQR 13.8–45.0) at baseline to 20.0 mg/day (IQR 10.0–30.0) at 3 months ($P < 0.01$), 10.0 mg/day (IQR 5.68–17.5) at 6 months ($P < 0.001$), and 5.0 mg/day (IQR 3.75–8.75) at 12 months ($P < 0.001$). By 12 months, 74.3% of patients maintained doses ≤7.5 mg/day, a significant increase from the baseline rate of 13.3% ($P < 0.0001$; Figure 1B). Additionally, the proportion of patients achieving >50% steroid dose reduction from baseline increased significantly from 11.6% at 3 months to 62.9% at 12 months ($P < 0.0001$; Figure 1C).

Comparison of telitacicept combined with SoC and historical belimumab control. *Baseline characteristics.* Both groups received similar SoC regimens, primarily consisting of glucocorticoids and hydroxychloroquine, combined with either cyclophosphamide or mycophenolate mofetil. In the telitacicept group, the proportions were 27.6% (16 of 58) and 60.3% (35 of 58), respectively, whereas in the belimumab group, they were 21.7% (13 of 60) and 58.3% (35 of 60). Before IPTW, baseline imbalances were observed between groups for several covariates, including sex, age, age at onset, disease duration, massive proteinuria, and neurologic and cardiopulmonary involvement

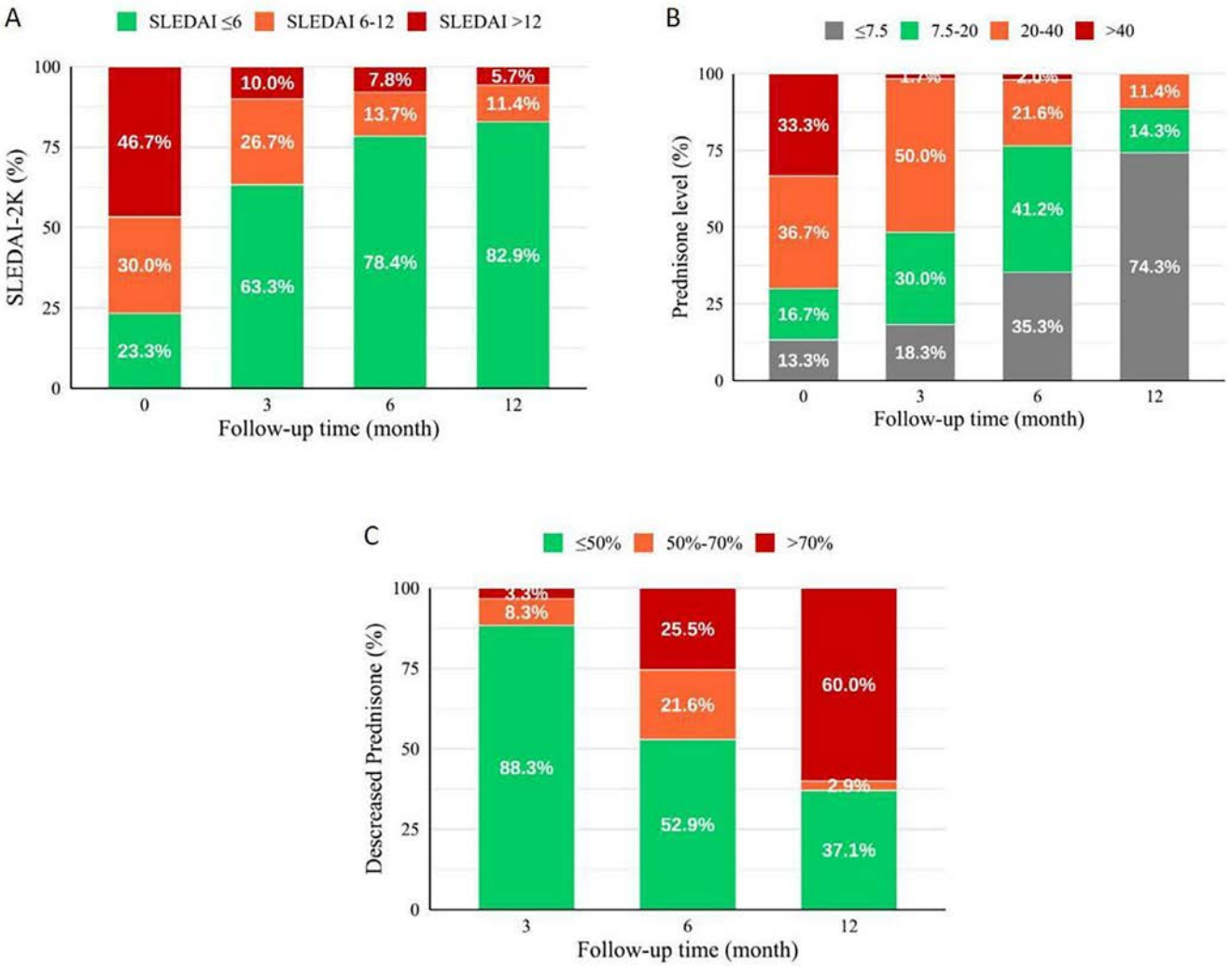


Figure 1. Proportions of patients from the telitacicept group with (A) different disease activities, (B) different daily prednisone equivalent doses, and (C) steroid reduction compared with baseline. SLEDAI-2K, Systemic Lupus Erythematosus Disease Activity Index 2000. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.43417/abstract.

(Table 2). After applying IPTW, the model demonstrated good predictive ability (area under the curve = 0.720, Supplementary Figure 5A), and six confounding covariates were well balanced with SMD <0.1 (Supplementary Figure 5B–D). However, neurologic involvement remained imbalanced (SMD = 0.121), as seen in Supplementary Figure 5C. To minimize potential residual confounding, we adjusted for this variable in our subsequent regression analysis.

Treatment efficacy after propensity score–based IPTW adjustment. After IPTW, the baseline characteristics were well balanced. Although the telitacicept group had numerically higher median SLEDAI-2K scores compared to the belimumab control group, the difference was not statistically significant ($P > 0.05$). Both groups demonstrated rapid and significant reductions in median SLEDAI-2K and PGA by three months ($P < 0.001$). However, the belimumab group exhibited lower SLEDAI-2K scores at three and six months and a lower PGA at three months compared to the telitacicept group ($P < 0.05$). By 12 months, no significant differences in SLEDAI-2K or PGA scores were observed between the two groups ($P > 0.05$; Figure 2A and B). Steroid doses significantly declined in both groups after treatment ($P < 0.01$), with no significant differences in daily doses at each time point ($P > 0.05$) (Figure 2C). Unweighted data are also presented in the chart for observation. During follow-up, eight patients in the telitacicept group and one patient in the belimumab control group successfully discontinued steroids. One patient in the telitacicept group experienced disease relapse three months after steroid cessation.

Immunologic markers after IPTW adjustment. Both groups exhibited increasing proportions of patients achieving negative anti-dsDNA and normalized C3 levels over time. There were no significant differences of negative anti-dsDNA and normalized C3 levels at each time point between the two groups ($P > 0.05$; Figure 3A and B). Serum IgG levels declined significantly from baseline to three months in both groups ($P < 0.05$), with no differences between the two groups at each time point ($P > 0.05$)

Table 2. Baseline characteristics of patients with cSLE treated with telitacicept and belimumab before and after IPTW*

	Unmatched				IPTW			
	Telitacicept group	Belimumab group	*P*	SMD	Telitacicept group	Belimumab group	*P*	SMD
n	60	67			61.9[a]	65.1[a]		
Female sex, n (%)	52 (86.7)	62 (92.5)	0.426	0.193	57.3 (92.6)	59.6 (91.6)	0.824	0.037
Age, mean ± SD, y	13.00 ± 2.76	11.43 ± 2.88	0.002	0.557	12.14 ± 2.82	11.92 ± 3.17	0.750	0.071
Age at onset, mean ± SD, y	10.81 ± 2.58	9.85 ± 3.1	0.064	0.335	10.32 ± 3.00	10.21 ± 2.78	0.851	0.039
Duration, mean ± SD, m	26.30 ± 30.34	17.15 ± 22.27	0.053	0.344	20.95 ± 24.38	20.51 ± 26.25	0.926	0.018
Mucocutaneous, n (%)	31 (51.7)	34 (50.7)	1.000	0.018	31.3 (50.6)	30.7 (47.1)	0.725	0.071
Renal, n (%)	35 (58.3)	37 (55.2)	0.862	0.063	35.2 (56.9)	37.1 (57.0)	0.990	0.002
Massive proteinuria, n (%)	13 (21.7)	6 (9.0)	0.079	0.359	9.7 (15.7)	9.4 (14.5)	0.873	0.033
Neurologic, n (%)	7 (11.7)	12 (17.9)	0.462	0.177	10.2 (16.4)	13.8 (21.1)	0.597	0.121
Cardiopulmonary, n (%)	16 (26.7)	22 (32.8)	0.573	0.135	19.9 (32.2)	20.2 (31.0)	0.903	0.025
Hematologic, n (%)	22 (36.7)	25 (37.3)	1.000	0.013	20.7 (33.4)	19.4 (29.7)	0.664	0.080
Leukopenia, n (%)	14 (23.3)	15 (22.4)	1.000	0.023	13.0 (20.9)	12.2 (18.8)	0.765	0.053

* cSLE, childhood-onset systemic lupus erythematosus; IPTW, inverse probability of treatment weighting; SMD, standardized mean difference.
[a] The counts in the weighted cohort may not sum to expected totals owing to rounding.

(Figure 3C). Baseline CD19 levels were low in both groups and further decreased after treatment ($P < 0.05$). The telitacicept group had lower CD19 levels at 3 months but slightly higher levels than the belimumab group by 12 months ($P < 0.05$; Figure 3D). The CD4/CD8 ratio remained stable in both groups throughout follow-up, consistently near the lower limit of the normal reference range, with no significant differences between the two groups ($P > 0.05$; Supplementary Figure 6).

Achievement of LLDAS and DORIS after IPTW. No significant differences were observed in the proportions of patients achieving LLDAS or DORIS between the two treatment groups at any assessed time point ($P > 0.05$; Figure 3E and F). Regarding disease flares, 12 relapses were recorded in 135 subsequent visits in the telitacicept group, compared to 9 relapses in 173 visits in the belimumab group, with no statistically significant difference ($P > 0.05$). In terms of safety, one case of fungal infection occurred in each group at three-month follow-up, leading to discontinuation of telitacicept or belimumab, respectively. No other severe adverse events were reported during the study period.

Multivariate regression analysis following IPTW adjustment. Multivariate Cox regression analysis showed no significant association between the treatment group and any clinical outcomes after adjusting for residual confounding due to neurologic involvement (all *P* values > 0.05). Using the belimumab group as the reference, the telitacicept group showed the adjusted hazard ratios of 0.898 (95% CI 0.578–1.395) for LLDAS achievement, 0.750 (95% CI 0.370–1.520) for DORIS achievement, 0.938 (95% CI 0.616–1.426) for anti-dsDNA negativity, and 0.873 (95% CI 0.662–1.152) for C3 normalization. Similarly, linear mixed-effects modeling revealed that, compared with the belimumab group, the telitacicept group showed no significant difference in in steroid dose (β = −0.149, 95% CI −0.680 to 0.378) and SLEDAI-2K (β = −0.104, 95% CI −0.370 to 0.161) over the follow-up period.

DISCUSSION

Clinical remission (CR) and LLDAS are recommended as ideal treatment targets for SLE, given their strong associations with reduced organ damage, fewer disease flares, and successful glucocorticoid tapering.[12,17] In recent years, these treat-to-targets strategies have also been proposed for cSLE.[18,19] Compared to aSLE, cSLE often exhibits more extensive multi-organ involvement and severer disease activity. Persistent disease activity and prolonged steroid use increase the risk of organ damage and growth retardation in cSLE. The concept of disease modification—emphasizing long-term outcomes—has thus gained prominence in SLE management.[20] This approach advocates for early, targeted, and sustained interventions that address fundamental drivers of disease, such as immune dysregulation and B cell hyperactivity. By modifying the natural disease course, the goal is to achieve sustained low disease activity with the lowest possible drug toxicity, ultimately preventing or delaying organ damage while improving quality of life. However, achieving treatment targets and glucocorticoids tapering in patients with cSLE under SoC remains slower and less successful than in adults.[21] The EULAR 2023 guidelines and Chinese pediatric SLE guidelines both highlight the importance of early use of biologic therapies to facilitate disease modification, particularly in children with high disease activity or refractory disease.[22,23]

Belimumab, a BLyS-specific inhibitor, acts by blocking soluble BLyS from binding to B cell receptors. In contrast, telitacicept is a novel fusion protein that targets the TACI receptor on B cells. By competitively binding to both APRIL and BLyS in circulation, telitacicept effectively suppresses the maturation and survival of B cells and plasma cells, thereby achieving dual-target inhibition

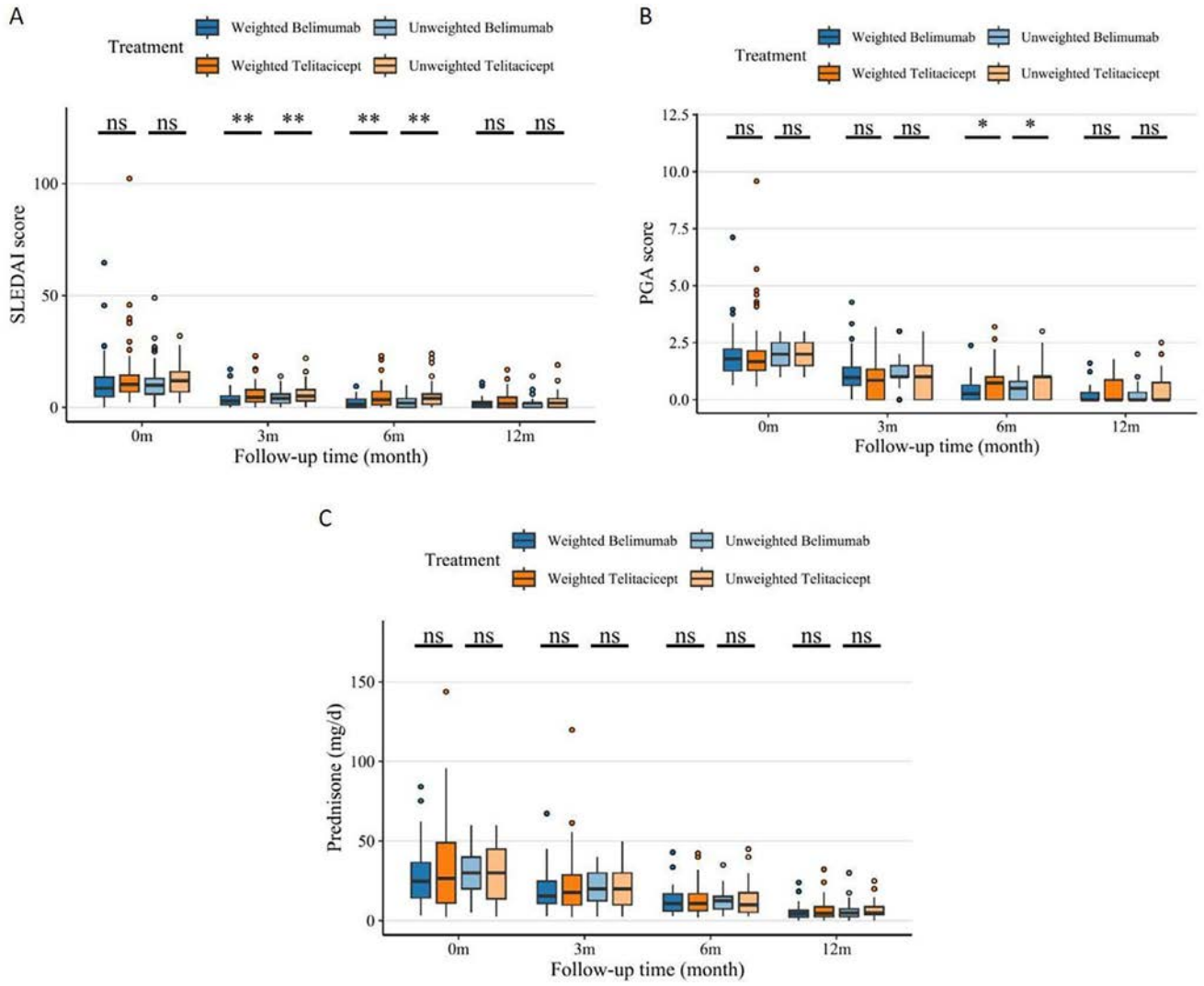


Figure 2. Treatment efficacy comparison between the belimumab group and the telitacicept group after and before adjustment using propensity score–based IPTW. (A) Weighted and unweighted SLEDAI-2K scores, (B) weighted and unweighted PGA, and (C) weighted and unweighted steroid doses. *P* values were calculated using rank-sum tests due to the nonnormal distribution of the variables. **P* < 0.05, ***P* < 0.01, ****P* < 0.001, ns: *P* > 0.05, telitacicept vs belimumab. IPTW, inverse probability of treatment weighting; ns, not significant; PGA, Patient Global Assessment; SLEDAI-2K, Systemic Lupus Erythematosus Disease Activity Index 2000. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.43417/abstract.

in SLE. Both biologics exert therapeutic effects through B cell modulation. A prior multicenter study led by our center demonstrated that belimumab, when combined with SoC, made LLDAS and steroid reduction attainable targets for cSLE.[8] This prospective real-world study on telitacicept in cSLE further reveals that telitacicept in combination with SoC not only markedly improves major organ manifestations but also enhances the attainment of LLDAS and CR in cSLE.

Previous Chinese studies demonstrated that only 20.7% and 10.4% of patients with SLE receiving SoC achieved LLDAS and CR, respectively,[24] substantially lower than those observed in our study. Although the study by Jin et al in adult SLE reported telitacicept with SoC achieved LLDAS rates of 26.39% at 3 months and 34.72% at 6 months,[25] higher than the rates in our cohort at the same time points, our cohort demonstrated superior outcomes by 12 months with a 65.7% LLDAS attainment rate versus 47.62% in adults, despite the adult cohort having lower baseline disease activity (median SLEDAI-2K < 10).

By three months, the median SLEDAI-2K scores decreased from 12 to 4, with only 10% of patients maintaining severe disease activity (SLEDAI-2K > 12). The findings corroborate previous reports of telitacicept's efficacy in hematologic and renal manifestations.[26–28] Immunologically, anti-dsDNA seroconversion rates progressively increased, whereas C3 and C4 levels normalized in 54.9% and 74.5% of patients by six months (exceeding aSLE results of 48.61% and 58.33%, respectively).[28] Consistent with other telitacicept studies, IgG and IgM levels

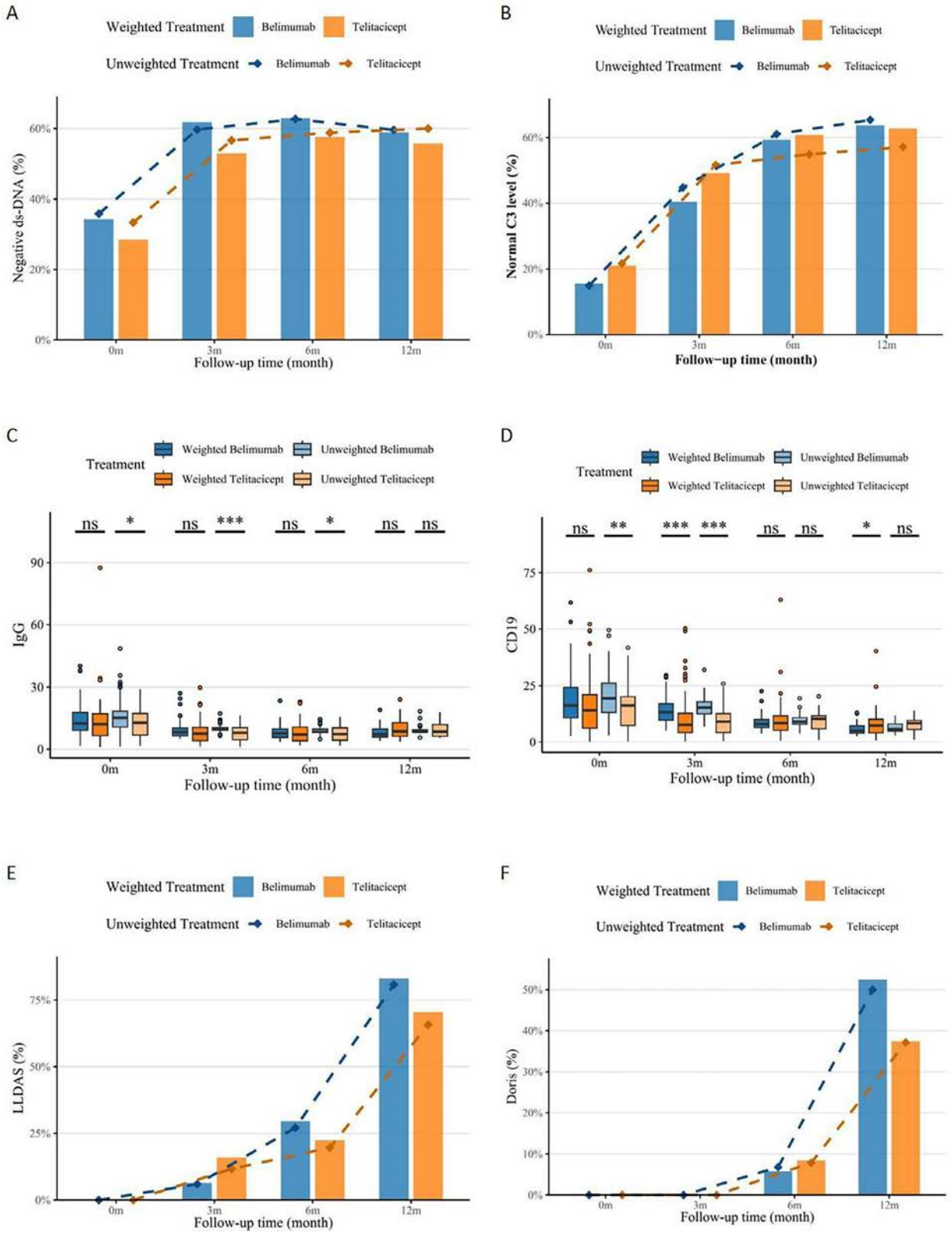


Figure 3. Comparison of immunologic markers and LLDAS/Doris between the belimumab and the telitacicept group after and before adjustment using propensity score–based IPTW. Proportions of patients with (A) negative anti-dsDNA antibodies, (B) normal C3 levels, (C) IgG levels, (D) Percentage of CD19+ cells, (E) proportions of LLDAS, and (F) proportions of Doris. *P* values were calculated using rank-sum tests due to the nonnormal distribution of the variables. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, ns: $P > 0.05$, telitacicept vs belimumab. Doris, Definition of Remission in SLE; dsDNA, double-stranded DNA; IPTW, inverse probability of treatment weighting; LLDAS, Lupus Low Disease Activity State. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.43417/abstract.

declined by 3 months and remained low through 12 months.[7,25] The safety profile remained favorable despite most patients not receiving prophylactic Ig support. The single serious infection (fungal pneumonia) occurred in a newly diagnosed patient receiving concurrent high-dose steroids (55 mg/day), hydroxychloroquine, cyclophosphamide, and telitacicept. The patient's CD19+ B cells plummeted from 11.39% to 0.97% at three months, suggesting the infection risk stemmed from cumulative immunosuppression rather than telitacicept monotherapy. Therefore, proactive measures such as Ig support or accelerated glucocorticoid tapering should be considered when initiating telitacicept alongside high-dose steroids and potent immunosuppressants.

The median baseline weight in this study was 40.75 kg, with 33.3% of patients receiving steroids >40 mg/day. The steroid tapering pattern revealed an initial increase in proportion of patients receiving 20 to 40 mg/day before subsequent reduction, demonstrating successful dose optimization over time. Notably, patients with cSLE typically require higher glucocorticoid doses than adults, posing unique challenges in achieving recommended tapering targets—≤7.5 mg/day as per EULAR guidelines and ≤5 mg/day according to the Chinese Guidelines for Childhood-Onset SLE—within six months of treatment initiation.[22] Despite combination therapy with biologics, our cohort's median steroid dose reduced from 30 mg/day at baseline to 10 mg/day at six months—a significant improvement, though still above the ideal target. These findings align with aSLE studies of patients receiving telitacicept combined with SoC, including significant reduction in SLEDAI-2K scores, effective correction of hypocomplementemia, and facilitation of steroid tapering.[6,29]

Current studies demonstrate that belimumab combined with SoC is effective and safe for cSLE, highlighting the therapeutic role of biologics in disease control and steroid tapering.[3,8] Although limited aSLE data show comparable performance between belimumab and telitacicept in achieving treatment targets (including LLDAS, PGA scores, and steroid tapering),[30,31] direct comparisons evidence in cSLE remains scarce. This study represents the largest prospective cohort study of telitacicept-treated patients with cSLE to date and is the first to compare the effectiveness of these two biologics in cSLE. To address potential confounding from nonconcurrent study periods, we employed IPTW to adjust baseline characteristics. Although the belimumab group achieved a lower median SLEDAI-2K at three and six months, baseline SLEDAI-2K in the telitacicept group was slightly higher. Both groups showed comparable LLDAS and DORIS remission at all time points, with DORIS rates exceeding those reported in a previous single-center study of belimumab (15.4%).[32] These results provide robust evidence supporting the comparable clinical effectiveness of both biologics in cSLE management when combined with SoC.

Our study demonstrated a higher rate of complete steroid discontinuation in the telitacicept group compared with belimumab-treated patients. This observation contrasts with previous reports of successful steroid withdrawal within one year in belimumab-treated cohorts.[32,33] These collective findings suggest that, with careful monitoring of disease activity and immunologic markers, steroid tapering can be accelerated in some patients.

Median CD19+ cell levels were low in both groups at baseline and further declined with B cell inhibitor therapy, reflecting the suppressive effects of glucocorticoids and immunosuppressants on B cells in addition to biologics. The CD4/CD8 ratio remained near the lower limit of normal, consistent with lupus-related CD4 depletion or CD8 expansion.[34,35] The underlying mechanisms remain unclear, and treatment with B cell inhibitors did not appear to normalize this immunologic imbalance.

Several limitations should be noted in this study. First, the LLDAS criteria were used in place of the recently developed childhood-specific LLDAS (cLLDAS) criteria, which incorporate weight-adjusted glucocorticoid thresholds (prednisone ≤0.15 mg/kg/day or ≤7.5 mg/day, whichever is lower) and are thus more appropriate for pediatric populations.[36] Although the median weight of participants (40.75 kg) approximated adult values—potentially minimizing the impact of this limitation—the cLLDAS criteria were not available at the time of study initiation in 2022, as they were published in 2023. Future studies will incorporate cLLDAS to better reflect pediatric-specific treatment targets. Second, the prospective cohort design employed historical controls rather than randomized assignment. Although IPTW adjusted for baseline characteristics, the study periods for the two cohorts were different, and residual confounding result may also arise from secular trends (particularly in diagnosis and management approaches) in addition to differences in patient characteristics. Nevertheless, this real-world comparison of a large cSLE cohort provides valuable clinical insights given the current evidence gap in pediatric lupus study. Third, the median follow-up duration of 12 months was insufficient to assess long-term treatment outcomes, including cumulative organ damage, infection risk, and impacts on growth and development. These longer-term outcomes are currently being addressed in an ongoing longitudinal cohort study.

In summary, telitacicept combined with SoC provides comprehensive therapeutic benefits for cSLE, including effective clinical symptom remission, improvement in immunologic markers, successful steroid reduction, and high rates of LLDAS achievement. These findings support the role of telitacicept as a valuable addition to the therapeutic arsenal for pediatric SLE, particularly in patients requiring rapid disease control and steroid minimization. The comparable effectiveness between these two biologics provides clinicians with expanded options for personalized treatment approaches in pediatric lupus management.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software,

investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding authors, Drs Luo and Shu confirm that all authors have provided the final approval of the version to be published and take responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. Smith EMD, Lythgoe H, Midgley A, et al. Juvenile-onset systemic lupus erythematosus: update on clinical presentation, pathophysiology and treatment options. Clin Immunol 2019;209:108274.
2. Kamphuis S, Silverman ED. Prevalence and burden of pediatric-onset systemic lupus erythematosus. Nat Rev Rheumatol 2010;6(9): 538–546.
3. Brunner HI, Abud-Mendoza C, Viola DO, et al; Paediatric Rheumatology International Trials Organisation (PRINTO) and the Pediatric Rheumatology Collaborative Study Group (PRCSG). Safety and efficacy of intravenous belimumab in children with systemic lupus erythematosus: results from a randomised, placebo-controlled trial. Ann Rheum Dis 2020;79(10):1340–1348.
4. Wu D, Li J, Xu D, et al. Telitacicept in patients with active systemic lupus erythematosus: results of a phase 2b, randomised, double-blind, placebo-controlled trial. Ann Rheum Dis 2024;83(4):475–487.
5. Liu Y, Zhu Y, Su Y, et al. Response to telitacicept in systemic lupus erythematosus patients: a real-world, single-center observational study in China. Int J Dermatol 2024;63(9):e213–e215.
6. Hu Y, Wang P, Cao X, et al. The effectiveness of telitacicept in patients with systemic lupus erythematosus: a retrospective, real-world study. Lupus 2025;34(2):133–139.
7. Ji L, Geng Y, Zhang X, et al. B cell pathway dual inhibition for systemic lupus erythematosus: a prospective single-arm cohort study of telitacicept. MedComm (2020) 2024;5(4):e515.
8. Wang L, Liang X, Cao Z, et al. Evaluation of belimumab in treatment of Chinese childhood-onset systemic lupus erythematosus: a prospective analysis from a multicentre study. Rheumatology (Oxford) 2024; 63(5):1437–1446.
9. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1997;40(9):1725.
10. Aringer M, Costenbader K, Daikh D, et al. 2019 European League Against Rheumatism/American College of Rheumatology classification criteria for systemic lupus erythematosus. Arthritis Rheumatol 2019;71(9):1400–1412.
11. Gladman DD, Ibañez D, Urowitz MB. Systemic Lupus Erythematosus Disease Activity Index 2000. J Rheumatol 2002;29(2):288–291.
12. Petri M, Hellmann D, Hochberg M. Validity and reliability of lupus activity measures in the routine clinic setting. J Rheumatol 1992;19(1): 53–59.
13. Franklyn K, Lau CS, Navarra SV, et al; Asia-Pacific Lupus Collaboration. Definition and initial validation of a lupus low disease activity state (LLDAS). Ann Rheum Dis 2016;75(9):1615–1621.
14. van Vollenhoven R, Voskuyl A, Bertsias G, et al. A framework for remission in SLE: consensus findings from a large international task force on definitions of remission in SLE (DORIS). Ann Rheum Dis 2017;76(3):554–561.
15. Petri M, Buyon J, Kim M. Classification and definition of major flares in SLE clinical trials. Lupus 1999;8(8):685–691.
16. Subspecialty Group of Immunology, The Society of Pediatrics, Chinese Medical Association. Evidence-based guidelines for diagnosis and treatment of common renal diseases in children (for trial) (VI): guidelines for the diagnosis and treatment of lupus nephritis. Article in Chinese. Zhonghua Er Ke Za Zhi 2010;48(9):687–690.
17. Kikuchi J, Hanaoka H, Saito S, et al. Lupus low disease activity state within 12 months is associated with favourable outcomes in severely active systemic lupus erythematosus. Rheumatology (Oxford) 2022; 61(9):3777–3791.
18. Cody EM, Wilson BE, Ogbu EA, et al. Usefulness of the lupus low disease activity state as a treatment target in childhood-onset SLE. Lupus Sci Med 2023;10(1):e000884.
19. Smith EMD, Tharmaratnam K, Al-Abadi E, et al. Attainment of low disease activity and remission targets reduces the risk of severe flare and new damage in childhood lupus. Rheumatology (Oxford) 2022;61(8): 3378–3389.
20. van Vollenhoven R, Askanase AD, Bomback AS, et al. Conceptual framework for defining disease modification in systemic lupus erythematosus: a call for formal criteria. Lupus Sci Med 2022;9(1): 18–34.
21. Multi-center Research Collaboration Group; Pediatric Allergy Immunology Rheumatology Branch; the Asia Pacific Association of Medicine and Bio-Immunology; Editorial Committee of Chinese Journal of Practical Pediatrics. Efficacy and safety analysis of belimumab for 28 weeks in real-world Chinese children with systemic lupus erythematosus: a retrospective multicenter pilot study. Article in Chinese. Zhongguo Shiyong Erke Zazhi 2021;36(11): 858–868.
22. Subspecialty Group of Immunology; the Society of Pediatrics; Chinese Medical Association. Chinese guidelines for the diagnosis and treatment of childhood-onset systemic lupus erythematosus. Article in Chinese. Zhonghua Er Ke Za Zhi 2021;59: 1009–1024.
23. Fanouriakis A, Kostopoulou M, Andersen J, et al. EULAR recommendations for the management of systemic lupus erythematosus: 2023 update. Ann Rheum Dis 2024;83(1):15–29.
24. Ren L, Zhao C, Zhao Y, et al. Low disease activity and remission status of systemic lupus erythematosus in a real-world study. Article in Chinese. Beijing Da Xue Xue Bao Yi Xue Ban 2024; 56(2):273–278.
25. Jin HZ, Li YJ, Wang X, et al. Efficacy and safety of telitacicept in patients with systemic lupus erythematosus: a multicentre, retrospective, real-world study. Lupus Sci Med 2023;10(2):e001074.
26. Cheng J, Peng Y, Wu Q, et al. Efficacy and safety of telitacicept therapy in systemic lupus erythematosus with hematological involvement. Clin Rheumatol 2024;43(7):2229–2236.
27. Chen Y, Lei X, Xu J, et al. Belimumab versus telitacicept in sequential treatment after rituximab for refractory lupus nephritis: a real-world multicentre study. Lupus Sci Med 2025;12(1):e001296.
28. Huang X, Lin F, Chen H. Efficacy and safety of telitacicept in patients with lupus nephritis: a single-center, real-world retrospective study. Clin Exp Nephrol 2024;28(9):902–909.
29. Fang F, Duan H, Ding S. Effectiveness and safety of telitacicept in patients with systemic lupus erythematosus: a single center, retrospective, real-world study. Clin Rheumatol 2025;44(3):1113–1122.
30. Jin HZ, Cai ML, Wang X, et al. Effectiveness and safety of belimumab and telitacicept in systemic lupus erythematosus: a real-world, retrospective, observational study. Clin Rheumatol 2025; 44(1):247–256.
31. Fan C, Yang T, Zheng S, et al. Frequency and predictors for early-achieved lupus low disease activity state in systemic lupus erythematosus patients treated with telitacicept or belimumab: a real-life, single-center observational study. Front Immunol 2024;15:1423035.
32. Wang D, Shan C, Liu J, et al. Efficacy and safety of belimumab for the treatment of refractory childhood-onset systemic lupus erythematosus: a single-center, real-world, retrospective study. Front Immunol 2022;13:1067721.

33. Collins CE, Dall'Era M, Kan H, et al. Response to belimumab among patients with systemic lupus erythematosus in clinical practice settings: 24-month results from the OBSErve study in the USA. Lupus Sci Med 2016;3(1):e000118.

34. Yang Z, Yu W, Lu Y. Circulating lymphocyte subpopulations in patients with systemic lupus erythematosus and their correlation with disease activity. Clin Exp Med 2023;23(8):4757–4763.

35. Żabińska M, Krajewska M, Kościelska-Kasprzak K, et al. CD3(+)CD8(+)CD28(-) T lymphocytes in patients with lupus nephritis. J Immunol Res 2016;1058165.

36. Smith EMD, Aggarwal A, Ainsworth J, et al; International cSLE T2T Task Force. PReS-endorsed international childhood lupus T2T task force definition of childhood lupus low disease activity state (cLLDAS). Clin Immunol 2023;250:109296.

Arthritis & Rheumatology
Vol. 78, No. 3, March 2026, pp 695–704
DOI 10.1002/art.43398

The Effect of Benralizumab and Mepolizumab on Use of Oral Glucocorticoids in Patients With Eosinophilic Granulomatosis With Polyangiitis

Parameswaran K. Nair,[1] Bernhard Hellmich,[2] Arnaud Bourdin,[3] David R. W. Jayne,[4] Florence Roufosse,[5] Nader Khalidi,[1] Lena Börjesson Sjö,[6] Ying Fan,[7] Christopher McCrae,[8] Sofia Necander,[6] Eva Rodríguez-Suárez,[9] Anat Shavit,[10] Claire Walton,[11] Peter A. Merkel,[12] and Michael E. Wechsler[13]

Objective. The phase 3 MANDARA study demonstrated noninferiority of benralizumab versus mepolizumab for remission in patients with eosinophilic granulomatosis with polyangiitis (EGPA). More benralizumab-treated patients achieved complete withdrawal of oral glucocorticoids (OGCs). The objectives of this post hoc analyses are to further elucidate the efficacy of benralizumab and mepolizumab in facilitating reductions in OGCs through investigating timings and sustainability of reducing OGCs as well as the cumulative use of OGCs.

Methods. Adults with EGPA requiring ≥7.5 mg/day OGC with or without immunosuppressive therapy were randomized to benralizumab 30 mg (n = 70) or mepolizumab 300 mg (n = 70) subcutaneously every 4 weeks for 52 weeks. Investigators tapered OGCs for patients with stable EGPA based on clinical judgment. Reductions in OGC use (to ≤4 mg/day, by ≥50%, or complete withdrawal) were considered sustained if achieved by week 40 and maintained through week 52.

Results. The 12-month cumulative dose of OGC was approximately 1,800 mg in both groups. At weeks 49 to 52, the median OGC dose was 1.16 (minimum–maximum 0.0–16.6) and 3.00 (minimum–maximum 0.0–25.7) mg/day for benralizumab and mepolizumab, respectively. Time to first or sustained OGC reduction to ≤4 mg/day or by ≥50% and accrued OGC-free duration were similar between groups. More benralizumab- than mepolizumab-treated patients completely withdrew OGCs (33 of 70 vs 20 of 70; hazard ratio [HR] for the time to the first complete withdrawal 1.84 [95% confidence interval [CI] 1.06–3.27]; unstratified log-rank test P = 0.0291) and sustained complete withdrawal (17 of 70 vs 7 of 70; HR for time to reduction 2.97 [95% CI 1.26–7.77]; unstratified log-rank test P = 0.0268). In both groups, complete OGC withdrawal was similar across patient subgroups, including by antineutrophilic cytoplasmic antibody status and immunosuppressive use.

Conclusion. Targeting the interleukin-5 receptor with benralizumab or circulating interleukin-5 with mepolizumab are effective OGC-sparing strategies in patients with EGPA.

INTRODUCTION

Eosinophilic granulomatosis with polyangiitis (EGPA) is a rare, progressive disease characterized by eosinophilic inflammation, necrotizing small- to medium-vessel vasculitis, and the presence of asthma.[1,2] Standard treatment for EGPA includes oral glucocorticoids (OGCs) and other immunosuppressive therapies that control symptoms, reduce eosinophilic inflammation,

Clinical Trial Registration: NCT04157348
This study was sponsored by AstraZeneca.
[1]McMaster University and St. Joseph's Healthcare, Hamilton, Ontario, Canada; [2]Klinik für Innere Medizin, Rheumatologie, Pneumologie, Nephrologie und Diabetologie, Medius Kliniken, Akademisches Lehrkrankenhaus der Universität Tübingen, Kirchheim unter Teck, Germany; [3]Department of Respiratory Diseases, University of Montpellier, CHU Montpellier, INSERM, CNRS, Montpellier, France; [4]Department of Medicine, University of Cambridge, Cambridge, United Kingdom; [5]Department of Internal Medicine, Hôpital Erasme, Université Libre de Bruxelles, Brussels, Belgium; [6]Late-stage Respiratory & Immunology, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden; [7]Late-stage Respiratory & Immunology, BioPharmaceuticals R&D, AstraZeneca, Gaithersburg, Maryland; [8]Translational Science & Experimental Medicine, Early Respiratory & Immunology, BioPharmaceuticals R&D, AstraZeneca, Gaithersburg, Maryland (current address: Precision Medicine, Amgen Inc., Thousand Oaks, California); [9]Translational Science and Experimental Medicine, Early Respiratory & Immunology, BioPharmaceuticals R&D, AstraZeneca, Cambridge, United Kingdom (current address: Novartis, London, United Kingdom); [10]BioPharmaceuticals Medical, AstraZeneca, Cambridge, United Kingdom (current address: Boehringer Ingelheim, Ingelheim am Rhein, Germany); [11]Late-stage Respiratory & Immunology, BioPharmaceuticals R&D, AstraZeneca, Cambridge, United Kingdom; [12]Division of Rheumatology, Department of Medicine, Division of Epidemiology, Department of Biostatistics, Epidemiology, and Informatics, University of Pennsylvania, Philadelphia, Pennsylvania; [13]Department of Medicine, National Jewish Health, Denver, Colorado.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/art.43398).

control vasculitis, and prolong survival.[3,4] However, patients with EGPA typically require long-term use of such therapies, which is associated with significant toxicity.[3,5] Glucocorticoid-related side effects are frequent, and the risk of metabolic (eg, adrenal insufficiency, glucose intolerance, and dyslipidemia) and other long-term toxicities increases even with low daily doses.[6]

The most frequent adverse events in a study focusing on patients with antineutrophilic cytoplasmic antibody (ANCA)–associated vasculitis included skin atrophy, osteoporosis, and myopathy with a cumulative dose of prednisone (or equivalent) ≥935 mg leading to an 80% increased likelihood of a clinically meaningful change in the Glucocorticoid Toxicity Index.[6] This risk may be lessened by tapering OGCs, but many patients with EGPA are unable to taper or fully discontinue OGCs because of disease relapses or adrenal insufficiency.[5,7,8] The considerable negative effects of OGC-related toxicity on health-related quality of life (HRQoL) in patients with chronic diseases has been well characterized in the literature,[5,9,10] including in a recent large, real-world study in patients with EGPA demonstrating that patients receiving ≥7.5 mg/day had worse HRQoL than patients receiving lower daily doses when assessed using two different validated questionnaires.[11] Taken together this evidence supports the clinical importance of therapies that facilitate the reduction, or elimination, of OGCs without affecting EGPA disease control.

Targeting eosinophilic inflammation is a well-recognized treatment strategy for patients with EGPA.[3,12,13] Both benralizumab and mepolizumab are subcutaneously administered biologic therapies that target eosinophils via different mechanisms of action. Benralizumab is a monoclonal antibody that targets the interleukin-5 receptor α subunit, leading to rapid eosinophil depletion via interleukin-5 blockade and antibody-dependent cellular cytotoxicity.[14] Benralizumab is an approved therapy for severe eosinophilic asthma and was recently approved for EGPA in the US and the European Union based on the results of the Phase 3 MANDARA trial.[15,16] Mepolizumab is a monoclonal antibody that targets circulating interleukin-5 and was the first approved biologic therapy for EGPA,[17] based on the results of the Phase 3 MIRRA trial.[18] Mepolizumab is also an approved therapy for severe eosinophilic asthma.[19]

The head-to-head MANDARA study demonstrated noninferiority of benralizumab to mepolizumab for achieving remission in patients with relapsing or refractory EGPA receiving one of these eosinophil-targeting therapies in addition to prednisolone (≥7.5 mg/day) with or without immunosuppressive therapy over 52 weeks.[12] Both agents were shown to allow for the reduction of OGC dosing, thereby adding data supporting the previously reported OGC-sparing effects of mepolizumab[18,20] and benralizumab[21–24] in patients with EGPA. The effect of treatment on OGC use during weeks 48 to 52 was a prespecified secondary endpoint of the MANDARA study, and it was shown that 41% of benralizumab- versus 26% of mepolizumab-treated patients were able to completely withdraw OGCs during this time period.[12] These post hoc exploratory analyses further examine the OGC-sparing effect of benralizumab and mepolizumab treatment in patients with EGPA.

PATIENTS AND METHODS

Study design. MANDARA was a Phase 3, randomized, double-blind, 52-week, noninferiority, head-to-head study (NCT04157348) that evaluated the efficacy and safety of benralizumab versus mepolizumab in adults with relapsing or refractory EGPA. The design and main results of the study have been described previously.[12]

Briefly, adults aged ≥18 years with EGPA (based on asthma and blood eosinophilia requiring stable OGCs [≥7.5 mg/day]) with or without stable immunosuppressive therapy for ≥4 weeks before randomization were recruited in the study. An Interactive Response Technology/Randomization and Trial Supply Management (IRT/RTSM) was used to centrally assign patients to randomized (1:1) treatment with benralizumab 30 mg or mepolizumab 300 mg subcutaneously every 4 weeks. Randomization was stratified by region (North America, Japan, and Western Europe); the IRT/RTSM assigned each patient a unique code (assigned sequentially in each stratum) from a list prepared by a computerized system. During the double-blind period, investigators were encouraged to taper OGCs after 4 weeks of treatment for patients with no active EGPA symptoms (Birmingham Vasculitis Activity Score [BVAS] of 0) based on clinical judgment. Currently, there are no guidelines that provide an algorithm on how to taper OGC use in EGPA.[4]

Endpoints. The analyses reported here include the time to first reduction of OGCs (≤4 mg/day, ≥50% reduction, and complete withdrawal, defined as the date of randomization to the first 4-week period meeting the reduction threshold); time to sustained reduction of OGCs (≤4 mg/day, ≥50% reduction, and complete withdrawal, defined as achieving reduction by week 40 and maintaining reduction until week 52); the average daily dose and cumulative dose of OGCs; complete OGC withdrawal during weeks 48 to 52 by patient baseline subgroups (assessed by demographic [sex, age, race, geographic region, and body mass index] and clinical characteristics at baseline [blood eosinophil count, OGC use, time since diagnosis, immunosuppressive therapy, Vasculitis Damage Index score, ANCA status, and

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/art.43398.

Address correspondence via email to Parameswaran K. Nair, MD, PhD, at parames@mcmaster.ca.

Submitted for publication February 7, 2025; accepted in revised form September 2, 2025.

number of relapses over past 2 years]); the accrued duration of OGC 0 mg/day; and the correlation of clinical outcomes (including forced expiratory volume in 1 second [FEV_1], the 6-item Asthma Control Questionnaire [ACQ-6] score, weight, fasting glucose, total cholesterol, and triglyceride levels) with accrued duration of OGC 0 mg/day.

The rationale for using the ≤4 mg/day OGC dose threshold was to align with the definition of remission used for the primary endpoint of the study, as recommended by regulatory authorities for registrational approval. A 12-week period of no active disease with OGC reduction was considered an appropriate definition of "sustained" remission, consistent with clinical decision-making. Accordingly, working backwards from week 52, week 40 was selected. Relapse assessment was conducted in a blinded manner, and was defined as active vasculitis (BVAS >0); OR active asthma symptoms and/or signs with a corresponding worsening in ACQ-6 score; OR active nasal and/or sinus disease, with a corresponding worsening in at least one of the questions on the Sino-nasal Symptoms Questionnaire AND warranting an increased total daily dose of OGC therapy to >4 mg/day; OR an increased dose or addition of immunosuppressive therapy; OR hospitalization related to worsening of EGPA.

Statistical analysis. For the purposes of analysis, all doses of OGCs were converted to their prednisone/prednisolone equivalent (in milligrams). Endpoints were analyzed using the median time to reduction threshold by the Kaplan–Meier method, *P* values from the unstratified log-rank test and hazard ratio (HR) from a Cox proportional hazards model (time to first OGC reduction and time to sustained OGC reduction); proportional odds models (average daily OGC dose during weeks 48–52 and accrued duration of OGC 0 mg/day); the marginal standardization method in logistic regression (OGC dose reduction during weeks 48–52); and Spearman rank correlation coefficient with loess (clinical outcomes). All models were adjusted for treatment group, baseline OGCs, baseline BVAS, and region. Baseline OGCs and baseline BVAS were expected to influence the endpoints; without adjustment, any imbalances between treatment groups would make it difficult to assess the relationship between treatment and endpoint. "Region" was a stratification variable in the design of the study; adjusting for stratification variables produces a more precise estimate of the treatment effect and reduces bias. For subgroup analyses, the covariate "region" was not included in the model when the subgroup was "geographical region" or "race," to avoid confounding. The endpoints in these post hoc analyses were not multiplicity-protected; therefore, all reported *P* values are nominal.

Research ethics. The trial was conducted in accordance with the ethical principles of the Declaration of Helsinki and is consistent with the International Council for Harmonisation Good Clinical Practice guidelines, the applicable regulatory requirements, and the AstraZeneca policy on bioethics. All patients provided written informed consent.[12]

RESULTS

Patient demographics and characteristics. In total, 140 patients were randomized to benralizumab (n = 70) or mepolizumab (n = 70), with a mean ± SD age of 52.3 ± 14.1 years; 60.0% were women. Baseline demographics were comparable between the two groups, as previously described.[12]

Patient clinical characteristics and disease manifestations at baseline are reported for those receiving an OGC dose of ≥12 mg/day (benralizumab n = 18 and mepolizumab n = 14) and < 12 mg/day at baseline (benralizumab n = 52 and mepolizumab n = 56), as some disease manifestations of EGPA, such as cardiomyopathy, alveolar hemorrhage, or glomerulonephritis, may require higher doses of OGCs (Supplementary Table S1). Patient demographics and clinical characteristics according to whether complete withdrawal of OGCs was achieved (benralizumab n = 29 and mepolizumab n = 18) or not achieved (benralizumab n = 41 and mepolizumab n = 52) at weeks 48 to 52 are shown in Supplementary Table S2. Patients with relapsing disease, male patients, and those receiving OGCs ≥12 mg/day were less likely to achieve complete withdrawal of OGCs. It should be noted that at baseline a higher proportion of male versus female patients were receiving OGC ≥12 mg/day (38% vs 13% of patients, respectively).

At baseline, the median OGC daily dose was 10 (minimum–maximum 5–30) mg/day and 10 (minimum–maximum 7.5–40) mg/day in benralizumab and mepolizumab recipients, respectively (Supplementary Table S3); whereas at weeks 49 to 52, it was 1.16 (minimum–maximum 0.0–16.6) mg/day versus 3.00 (minimum–maximum 0.0–25.7) mg/day, respectively.

Time to the first reduction of OGCs. The median time to the first reduction of OGC dosing to ≤4 mg/day was similar between benralizumab (median 24.14; 95% confidence interval [CI] 20.14–28.14 weeks) and mepolizumab (median 28.14; 95% CI 24.14–32.14 weeks), with an HR of 1.26 (95% CI 0.87–1.84; unstratified log-rank test *P* = 0.3954) (Figure 1A and Supplementary Table S4). The median time to the first ≥50% reduction of OGC was also similar (unstratified log-rank test *P* = 0.1962) between benralizumab (median 20.14; 95% CI 16.14–24.14 weeks) and mepolizumab (median 20.14; 95% CI 20.14–24.14 weeks), with an HR of 1.28 (95% CI 0.89–1.83) (Figure 1B and Supplementary Table S4). The median time to the first complete withdrawal of OGCs could not be estimated as <50% of patients achieved this outcome (33 of 70 [47.1%] in the benralizumab group vs 20 of 70 [28.6%] in the mepolizumab group). The HR for the time to the first complete withdrawal of OGCs was 1.84 (95% CI 1.06–3.27), with more patients in the benralizumab than the mepolizumab group able to completely withdraw OGCs

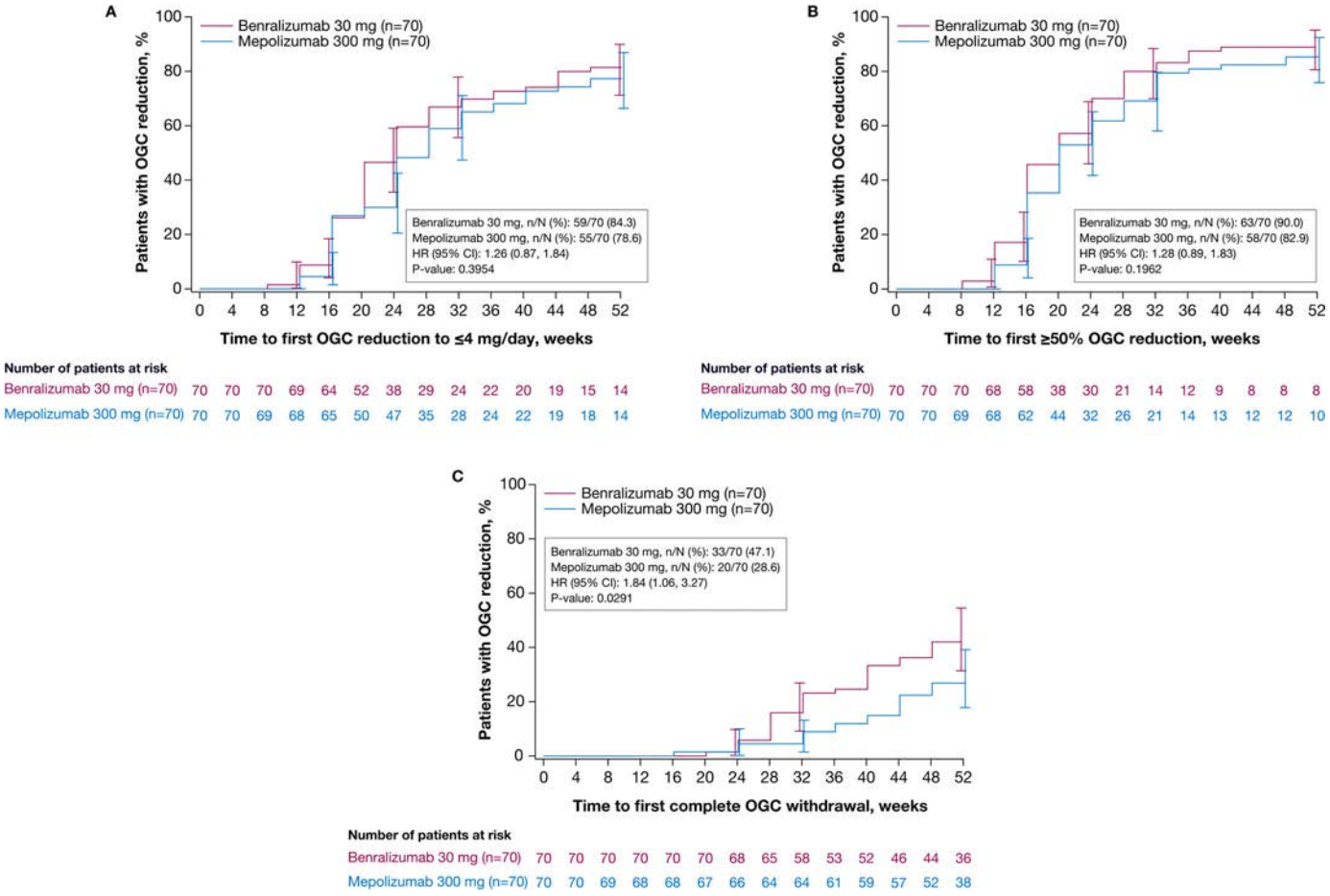


Figure 1. Time to first (A) OGC reduction to ≤4 mg/day, (B) ≥50% OGC reduction, and (C) complete OGC withdrawal. Time to first reduction was defined as the date of randomization to the first 4-week period meeting the OGC reduction threshold. The HR (benralizumab vs mepolizumab) and 95% CI were estimated using a Cox regression model with the Efron method to control for ties. The covariates in the model included treatment arm, baseline dose of prednisone, baseline Birmingham Vasculitis Activity Score, and region. An HR of >1 favored benralizumab. *P* values were derived from an unstratified log-rank test. CI, confidence interval; HR, hazard ratio; OGC, oral glucocorticoid.

during the double-blind phase (unstratified log-rank test $P = 0.0291$) (Figure 1C and Supplementary Table S4).

Time to sustained reduction of OGCs. The median time to the first sustained reduction of OGC to ≤4 mg/day was similar between groups (unstratified log-rank test $P = 0.6255$): 36.14 (95% CI 24.14 to not estimable) weeks for the benralizumab group and 32.14 (95% CI 28.14 to not estimable) weeks for the mepolizumab group, with an HR of 0.97 (95% CI 0.62–1.52) (Figure 2A and Supplementary Table S4). The median time to the first sustained ≥50% reduction of OGCs was also similar (unstratified log-rank test $P = 0.4683$) between benralizumab (median 24.14; 95% CI 24.14–32.14 weeks) and mepolizumab (median 28.14; 95% CI 20.14–32.14 weeks), with an HR of 1.14 (95% CI 0.77–1.69) (Figure 2B and Supplementary Table S4). The median time to sustained complete withdrawal of OGCs was not estimable as <50% of patients achieved this (17 of 70 [24.3%] in the benralizumab group vs 7 of 70 [10.0%] in the mepolizumab group). The HR for the time to sustained complete withdrawal of OGCs was 2.97 (95% CI 1.26–7.77), with more patients in the benralizumab than the mepolizumab group able to sustain complete withdrawal of OGCs (unstratified log-rank test $P = 0.0268$) (Figure 2C and Supplementary Table S4).

Average daily dose and cumulative use of OGCs. A heatmap of patient-level, 4-week average daily dose of OGC indicated that benralizumab-treated patients tapered faster, and more patients were able to eliminate and maintain withdrawal from OGCs during the study, than mepolizumab-treated patients (Figure 3; a color vision-deficiency friendly version is shown in Supplementary Figure S1). Relapses during the double-blind period are denoted by "R" on the heatmap.

During the 52-week double-blind period, the median cumulative use of OGC was 1,813 (minimum–maximum 486–5,674) mg vs 1,846 (minimum–maximum 573–7,999) mg for benralizumab and mepolizumab recipients, respectively. Eighteen (25.7%)

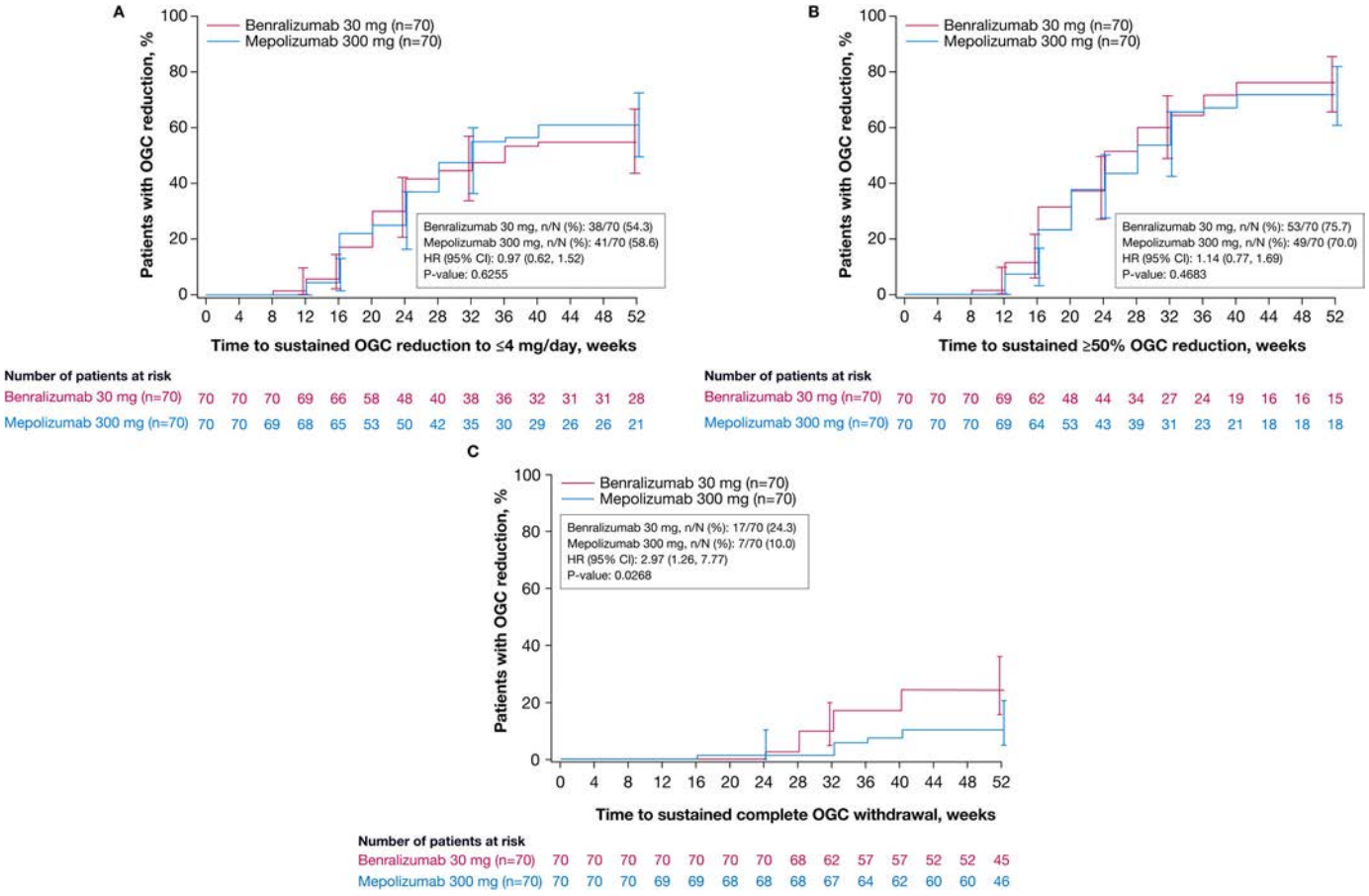


Figure 2. Time to sustained (A) OGC reduction to ≤4 mg/day, (B) ≥50% OGC reduction, and (C) complete OGC withdrawal. Time to sustained reduction was defined as achieving OGC ≤4 mg/day, ≥50% OGC reduction, or complete OGC withdrawal by week 40 and maintaining reduction through to week 52. The HR (benralizumab vs mepolizumab) and 95% CI were estimated using a Cox regression model with the Efron method to control for ties. The covariates in the model included treatment arm, baseline dose of prednisone, baseline Birmingham Vasculitis Activity Score, and region. An HR of >1 favored benralizumab. *P* values were derived from an unstratified log-rank test. CI, confidence interval; HR, hazard ratio; OGC, oral glucocorticoid.

benralizumab and 11 (15.7%) mepolizumab recipients had a cumulative dose of <1,000 mg (Supplementary Table S5).

OGC withdrawal across subgroups. Subgroup analyses of the proportion of patients in both treatment groups achieving complete withdrawal of OGCs during the last 4 weeks of the double-blind period are shown in Figure 4. More patients who were receiving a baseline OGC dose <12 mg/day achieved complete OGC withdrawal (26 of 52 [50.0%] benralizumab and 15 of 56 [26.8%] mepolizumab patients; difference 22.27 [95% CI 4.56–39.99]) compared with patients receiving a baseline OGC dose ≥12 mg/day (3 of 18 [16.7%] benralizumab and 3 of 14 [21.4%] mepolizumab patients; difference: −2.85 [95% CI −28.97 to 23.26]). ANCA status (at baseline/historic) did not impact patients' ability to withdraw OGCs. Of those who were receiving non-OGC immunosuppressive therapy at baseline, 10 of 26 (38.5%) benralizumab and 7 of 24 (29.2%) mepolizumab recipients were able to withdraw OGCs (difference 8.00; 95% CI −16.72 to 32.72). Findings were comparable for benralizumab and mepolizumab, regardless of the type of immunosuppressive therapy (methotrexate, azathioprine, or mycophenolate). In patients who were not using non-OGC immunosuppressive drugs at baseline, 19 of 44 (43.2%) benralizumab and 11 of 46 (23.9%) mepolizumab recipients were able to withdraw OGCs (difference 19.94; 95% CI 1.03–38.86).

Accrued OGC-free duration. There was no statistically significant difference in total accrued duration of OGC-free status between treatment groups (odds ratio 1.62; 95% CI 0.86–3.07; $P = 0.1389$) (Supplementary Figure S2 and Supplementary Table S6).

Correlation of accrued OGC-free and clinical outcomes duration. Correlations of accrued duration of OGC dose of 0 mg/day and asthma control (ACQ-6), lung function (as assessed by FEV_1), weight, fasting glucose, cholesterol, and

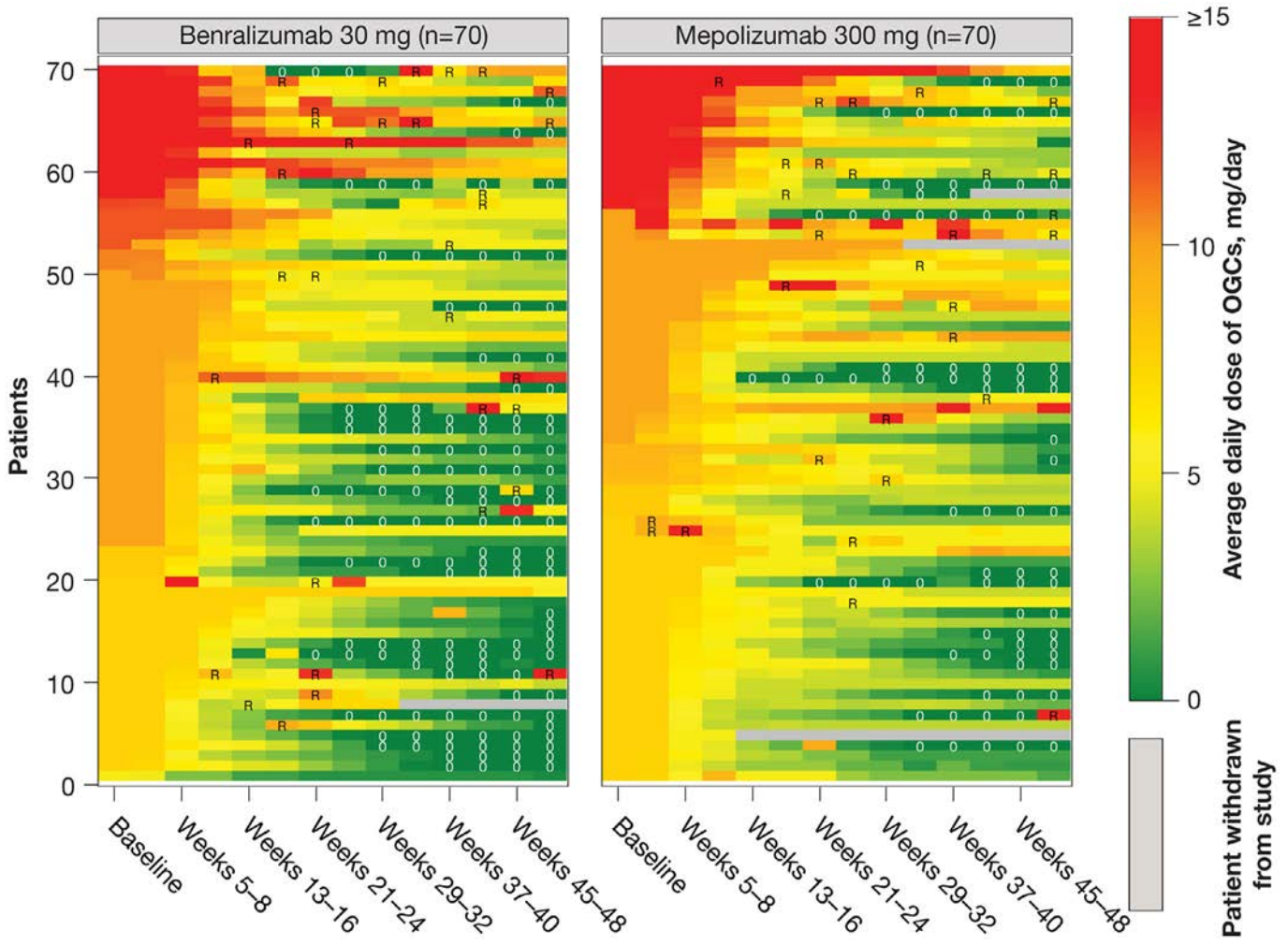


Figure 3. Heatmap of the average daily dose of OGCs throughout the double-blind period. Each row represents a single patient's OGC use over time during the study. Decreasing OGC doses are represented by a gradient from red to orange and yellow to green, with dark green and "0" denoting 0 mg/day OGC. OGC, oral glucocorticoid; R, relapse.

triglycerides were assessed. There was no loss of asthma control and no deterioration in lung function in either treatment group despite complete withdrawal of OGCs (Supplementary Figure S3A and B). A weak positive correlation between accrued duration of complete withdrawal of OGCs and (improved) lung function was found in the benralizumab-treated group (Spearman rank correlation 0.280; $P = 0.0250$) but not in the mepolizumab-treated group (Spearman rank correlation –0.004; $P = 0.9743$).

Weight, fasting glucose, cholesterol, and triglycerides were investigated to determine if there was an effect of OGC tapering, as these parameters can be adversely affected by OGC use. A weak negative correlation between accrued duration of complete OGC withdrawal and weight was observed in the mepolizumab-treated group (Spearman rank correlation –0.271; $P = 0.0264$) but not in the benralizumab-treated group (Spearman rank correlation –0.195; $P = 0.1130$) (Supplementary Figure S4A). There was no correlation between accrued duration of complete withdrawal of OGCs and fasting glucose level (Spearman rank correlation for benralizumab –0.079; $P = 0.5308$; and mepolizumab –0.061; $P = 0.6305$; Supplementary Figure S4B), total cholesterol (Spearman rank correlation for benralizumab –0.142; $P = 0.2539$; and mepolizumab –0.112; $P = 0.3734$; Supplementary Figure S4C), or triglyceride levels (Spearman rank correlation for benralizumab 0.034; $P = 0.7901$; and mepolizumab –0.226; $P = 0.0977$; Supplementary Figure S4D).

DISCUSSION

Standard of care for EGPA often requires long-term use of OGCs. OGCs have nonspecific, anti-inflammatory effects, inhibiting pro-inflammatory cell recruitment and cytokines. Long-term use of OGCs is associated with toxicity and a high risk of significant adverse effects even with low daily doses.[3,6,25,26] Analyses of responses from patient interviews suggest that stopping or

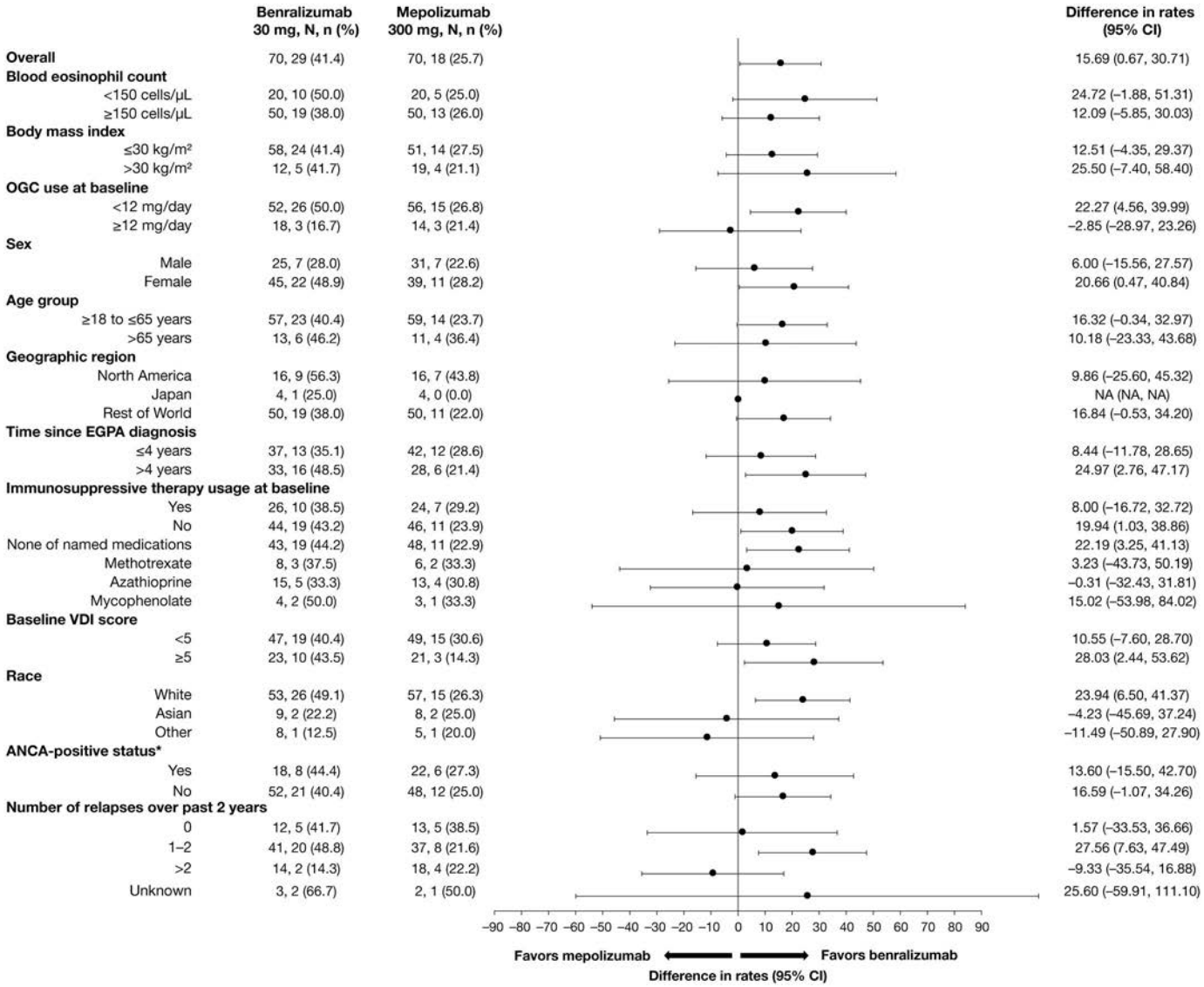


Figure 4. A subgroup analysis of the proportion of patients achieving complete OGC withdrawal during weeks 48 through 52. Rates and the difference (benralizumab – mepolizumab) were estimated using the marginal standardization method in a logistic regression model. The covariates in the model included treatment arm, subgroup, and interaction between treatment arm and the subgroup (treatment arm × subgroup). The *P* value for each interaction term was estimated from logistic regression without using the marginal standardization method. ANCA, antineutrophilic cytoplasmic antibody; CI, confidence interval; EGPA, eosinophilic granulomatosis with polyangiitis; NA, not applicable; OGC, oral glucocorticoid; VDI, Vasculitis Damage Index.

reducing the dose of OGCs is the number one priority for patients with EGPA and for clinicians.[27] Guidelines recommend adding conventional immunosuppressives to taper OGCs, although there is currently no specific guidance on how to taper OGC use in EGPA.[3,4] Furthermore, relapses (mainly airway-related manifestations) are common during OGC tapering.[28]

Eosinophilic inflammation is a key feature of EGPA. Therefore, agents that specifically target eosinophils as a primary source of inflammation can have OGC-sparing effects in EGPA.[12,18,20–22] This post hoc analysis of the MANDARA study of patients with EGPA, in which treatment with either benralizumab or mepolizumab was associated with the ability to reduce OGC use, extends understanding of the OGC-sparing effects of such treatments.

Large reductions were observed in the median daily dose of OGCs to a similar extent in both treatment groups, demonstrating the efficacy of targeting eosinophils in EGPA. The reductions from baseline to week 52 were much larger than in the placebo-controlled MIRRA study that evaluated efficacy of mepolizumab versus placebo in patients with EGPA.[18] In contrast to the MIRRA study, MANDARA was a head-to-head trial, and all patients had active systemic therapy while tapering OGC. This is likely to have facilitated decisions to taper OGC more quickly, as the perceived risk of harm to patients was lower, and investigators were aware

of the benefits of eosinophil-targeting therapy following US Food and Drug Administration approval of mepolizumab for treatment of EGPA.[17]

Studies in both vasculitis[6] and asthma[29] have shown that cumulative exposure to OGC was significantly correlated with OGC-related toxicity, which underscores the importance of OGC sparing in patients with EGPA. The median cumulative dose of OGCs during the double-blind period of the MANDARA study was similar in both treatment groups, approximately 1,800 mg. Long-term data from the open-label extension of MANDARA will provide additional insights on whether patients are able to further taper and/or sustain reductions in OGCs with longer treatment duration, specifically with respect to decreasing below the 935 mg threshold reportedly linked to a risk of OGC-related toxicity.[6] It is notable that the cumulative dose from MANDARA was lower than in other studies with standard-of-care regimens for vasculitis,[30,31] although these other studies included patients with severe disease requiring higher OGC doses at baseline, which may account for this difference.

The time to the first reduction of OGCs, the time to sustained reduction of OGCs, and accrued duration of OGC-free status were similar between the benralizumab and mepolizumab groups in the present analysis. During weeks 48 to 52, however, benralizumab-treated patients were more likely to fully eliminate the use of OGCs and achieved a greater depletion of blood eosinophils.[12] These results may be because of differences in the mechanism of action between mepolizumab and benralizumab; whereas mepolizumab targets interleukin-5 only, benralizumab's antibody-dependent cellular cytotoxicity mechanism leads to the rapid depletion of blood eosinophils to near-zero levels.[14,17,32,33] Although this mechanism reduces eosinophilic inflammation, a direct link between eosinophil counts in the blood or tissues and OGC use has yet to be elucidated.

The ability to achieve complete OGC withdrawal was not affected by baseline/historic ANCA status or concomitant non-OGC immunosuppressive use, suggesting that patients who are ANCA positive or who are considered to have more severe or difficult to control disease and are treated with non-OGC immunosuppressive therapy can still benefit from the addition of eosinophil-targeting therapies. The length of accrued duration of complete OGC withdrawal was not associated with deterioration in lung function or a loss of asthma control (with a low relapse rate in both treatment groups), suggesting that OGCs may be tapered in patients with EGPA receiving eosinophil-targeting therapies without affecting these parameters. Taken together, the primary results from both the MIRRA and MANDARA studies, in addition to the post hoc analyses presented here, should provide clinicians with reassurance that targeting interleukin-5/interleukin-5 receptors can not only control disease but also facilitate OGC reductions or discontinuation without any deleterious effect on treatment efficacy.

These analyses carry some limitations. First, because patients starting at a baseline OGC dose of <12 mg/day were more able to completely withdraw OGCs than those starting at higher dosages, it is possible that, with a longer study duration, some patients receiving higher doses might have been able to reduce their use of OGC further. Second, the possibility of reducing the dose/discontinuing concomitant non-OGC immunosuppressive therapies was not specified in the study protocol for the double-blind period and was not assessed herein. Third, following advice from regulatory authorities to align the MANDARA protocol with that of the MIRRA study and thereby allow clinicians leeway in prescribing, the tapering of OGCs was not standardized during this study. Although a recommended schedule was distributed with the study protocol, ultimately the rate of OGC tapering was a matter of clinical judgment by the investigator and standard practice. Additionally, data on OGC-related toxicity (the items on the Glucocorticoid Toxicity Index) were not assessed specifically as study endpoints, and as such any differences between groups as a result of treatment is not known. Fourth, adrenal function was not evaluated robustly in this study; some analyses from asthma studies have shown that this may play a role in preventing some patients from further reducing OGC use.[34] Fifth, because an evaluation of airway eosinophilia was not possible with the MANDARA data, the present analysis cannot assess to what extent this affected OGC tapering in this study. Finally, all the data described here were post hoc exploratory analyses in a small clinical trial patient population, without active severe manifestations, and none of the analyses were multiplicity-protected. The clinical implications of these findings require further investigation in real-world settings.

In summary, the MANDARA study showed that a significant reduction in OGC exposure was possible with use of the eosinophil-targeting therapies benralizumab and mepolizumab while maintaining disease control, suggesting these biologics may help reduce the burden and toxicity of prolonged treatment with OGCs. Although updated guidance on treatment strategies and guidelines for management of EGPA is awaited, these data support that complete withdrawal of OGCs is possible and should become a new treatment goal for patients with EGPA.

ACKNOWLEDGMENTS

The authors thank the patients and their caregivers and site investigators and site staff who participated in this study. The authors would like to thank Emmanuelle Maho for statistical programming support, critical review, and input. Medical writing support was provided by Caroline Ridley and Anna Mett of Springer Health+, Springer Healthcare Ltd, UK, which was funded by AstraZeneca in accordance with Good Publication Practice guidelines.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr Nair confirms that all authors have provided the final approval

of the version to be published and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

ROLE OF THE STUDY SPONSOR

AstraZeneca had no role in the study design or in the collection, analysis, or interpretation of the data; the writing of the manuscript; or the decision to submit the manuscript for publication. Publication of this article was not contingent upon approval by AstraZeneca.

ADDITIONAL DISCLOSURES

Lena Börjesson Sjö, Ying Fan, Christopher McCrae, Sofia Necander, Eva Rodríguez-Suárez, Anat Shavit, and Claire Walton are/were employees of AstraZeneca.

Data availability statement

Data underlying the findings described in this manuscript may be obtained in accordance with AstraZeneca's data sharing policy described at https://astrazenecagrouptrials.pharmacm.com/ST/Submission/Disclosure. Data for studies directly listed on Vivli can be requested through Vivli at www.vivli.org. Data for studies not listed on Vivli could be requested through Vivli at https://vivli.org/members/enquiries-about-studies-not-listed-on-the-vivli-platform/. The AstraZeneca Vivli member page is also available, and it outlines further details at https://vivli.org/ourmember/astrazeneca/.

REFERENCES

1. Fijolek J, Radzikowska E. Eosinophilic granulomatosis with polyangiitis: advances in pathogenesis, diagnosis, and treatment. Front Med (Lausanne) 2023;10:1145257.
2. Grayson PC, Ponte C, Suppiah R, et al; DCVAS Study Group. 2022 American College of Rheumatology/European Alliance of Associations for Rheumatology classification criteria for eosinophilic granulomatosis with polyangiitis. Arthritis Rheumatol 2022;74(3):386–392.
3. Chung SA, Langford CA, Maz M, et al. 2021 American College of Rheumatology/Vasculitis Foundation guideline for the management of antineutrophil cytoplasmic antibody-associated vasculitis. Arthritis Rheumatol 2021;73(8):1366–1383.
4. Hellmich B, Sanchez-Alamo B, Schirmer JH, et al. EULAR recommendations for the management of ANCA-associated vasculitis: 2022 update. Ann Rheum Dis 2024;83(1):30–47.
5. Daugherty J, Lin X, Baxter R, et al. The impact of long-term systemic glucocorticoid use in severe asthma: a UK retrospective cohort analysis. J Asthma 2018;55(6):651–658.
6. Scherbacher PJ, Hellmich B, Feng YS, et al. Prospective study of complications and sequelae of glucocorticoid therapy in ANCA-associated vasculitis. RMD Open 2024;10(1):e003956.
7. Comarmond C, Pagnoux C, Khellaf M, et al; French Vasculitis Study Group. Eosinophilic granulomatosis with polyangiitis (Churg-Strauss): clinical characteristics and long-term followup of the 383 patients enrolled in the French Vasculitis Study Group cohort. Arthritis Rheum 2013;65(1):270–281.
8. Strehl C, Bijlsma JW, de Wit M, et al. Defining conditions where long-term glucocorticoid treatment has an acceptably low level of harm to facilitate implementation of existing recommendations: viewpoints from an EULAR task force. Ann Rheum Dis 2016;75(6):952–957.
9. Cheah JTL, Robson JC, Black RJ, et al. The patient's perspective of the adverse effects of glucocorticoid use: a systematic review of quantitative and qualitative studies. From an OMERACT working group. Semin Arthritis Rheum 2020;50(5):996–1005.
10. Saketkoo LA, Russell AM, Jensen K, et al. Health-related quality of life (HRQoL) in sarcoidosis: diagnosis, management, and health outcomes. Diagnostics (Basel) 2021;11(6):1089.
11. Spiera R, Dolin P, Shavit A, et al. The impact of eosinophilic granulomatosis with polyangiitis on the health-related quality of life of patients and their ability to work: evidence from a real-world survey in clinical practice [abstract]. Arthritis Rheumatol 2024;76(9). https://acrabstracts.org/abstract/the-impact-of-eosinophilic-granulomatosis-with-polyangiitis-on-the-health-related-quality-of-life-of-patients-and-their-ability-to-work-evidence-from-a-real-world-survey-in-clinical-practice/
12. Wechsler ME, Nair P, Terrier B, et al; MANDARA Study Group. Benralizumab versus mepolizumab for eosinophilic granulomatosis with polyangiitis. N Engl J Med 2024;390(10):911–921.
13. Hellmich B, Wechsler M, Merkel P, et al. Effect of benralizumab versus mepolizumab on reduction in oral glucocorticoid use in patients with eosinophilic granulomatosis with polyangiitis (OP0188). Ann Rheum Dis 2024;83(Suppl 1):186–187.
14. Kolbeck R, Kozhich A, Koike M, et al. MEDI-563, a humanized anti-IL-5 receptor alpha mAb with enhanced antibody-dependent cell-mediated cytotoxicity function. J Allergy Clin Immunol 2010;125(6):1344–1353.e2.
15. Fasenra. Prescribing information. AstraZeneca; 2017. Accessed November 14, 2025. https://www.accessdata.fda.gov/drugsatfda_docs/label/2017/761070s000lbl.pdf
16. European Medicines Agency. Fasenra. Updated February 27, 2025. Accessed November 14, 2025. https://www.ema.europa.eu/en/medicines/human/EPAR/fasenra
17. PR Newswire. FDA approves first drug for eosinophilic granulomatosis with polyangiitis, a rare disease formerly known as the Churg-Strauss Syndrome. Published December 12, 2017. Accessed November 14, 2025. https://www.prnewswire.com/news-releases/fda-approves-first-drug-for-eosinophilic-granulomatosis-with-polyangiitis-a-rare-disease-formerly-known-as-the-churg-strauss-syndrome-300570130.html
18. Wechsler ME, Akuthota P, Jayne D, et al; EGPA Mepolizumab Study Team. Mepolizumab or placebo for eosinophilic granulomatosis with polyangiitis. N Engl J Med 2017;376(20):1921–1932.
19. Nucala. Prescribing information. GlaxoSmithKline; 2023. Accessed November 14, 2025. https://www.accessdata.fda.gov/drugsatfda_docs/label/2023/125526Orig1s021,761122Orig1s011Corrected_lbl.pdf
20. Jayne DRW, Terrier B, Hellmich B, et al. Mepolizumab has clinical benefits including oral corticosteroid sparing irrespective of baseline EGPA characteristics. ERJ Open Res 2024;10(1):00509–2023.
21. Guntur VP, Manka LA, Denson JL, et al. Benralizumab as a steroid-sparing treatment option in eosinophilic granulomatosis with polyangiitis. J Allergy Clin Immunol Pract 2021;9(3):1186–1193.e1.
22. Nanzer AM, Maynard-Paquette AC, Alam V, et al. Long-term effectiveness of benralizumab in eosinophilic granulomatosis with polyangiitis. J Allergy Clin Immunol Pract 2024;12(3):724–732.
23. Bettiol A, Urban ML, Padoan R, et al; European EGPA Study Group. Benralizumab for eosinophilic granulomatosis with polyangiitis: a retrospective, multicentre, cohort study. Lancet Rheumatol 2023;5(12):e707–e715.

24. Cottu A, Groh M, Desaintjean C, et al; French Vasculitis Study Group. Benralizumab for eosinophilic granulomatosis with polyangiitis. Ann Rheum Dis 2023;82(12):1580–1586.
25. Schleimer RP, Bochner BS. The effects of glucocorticoids on human eosinophils. J Allergy Clin Immunol 1994;94(6 Pt 2):1202–1213.
26. Walsh GM, Sexton DW, Blaylock MG. Corticosteroids, eosinophils and bronchial epithelial cells: new insights into the resolution of inflammation in asthma. J Endocrinol 2003;178(1):37–43.
27. Merkel PA, Hellmich B, Pagnoux C, et al. How symptoms and treatments impact the experiences of patients with eosinophilic granulomatosis with polyangiitis participating in a clinical trial (POS0858). Ann Rheum Dis 2024;83(Suppl 1):1071–1072.
28. Merkel PA, Jayne D, Specks U, et al. Characteristics of relapses in patients with eosinophilic granulomatosis with polyangiitis [abstract]. Arthritis Rheumatol 2024;76(suppl 9). https://acrabstracts.org/abstract/characteristics-of-relapses-in-patients-with-eosinophilic-granulomatosis-with-polyangiitis/
29. Price DB, Trudo F, Voorham J, et al. Adverse outcomes from initiation of systemic corticosteroids for asthma: long-term observational study. J Asthma Allergy 2018;11:193–204.
30. Jayne DRW, Merkel PA, Schall TJ, et al; ADVOCATE Study Group. Avacopan for the treatment of ANCA-associated vasculitis. N Engl J Med 2021;384(7):599–609.
31. Stone JH, Tuckwell K, Dimonaco S, et al. Trial of tocilizumab in giant-cell arteritis. N Engl J Med 2017;377(4):317–328.
32. Bleecker ER, FitzGerald JM, Chanez P, et al; SIROCCO study investigators. Efficacy and safety of benralizumab for patients with severe asthma uncontrolled with high-dosage inhaled corticosteroids and long-acting β2-agonists (SIROCCO): a randomised, multicentre, placebo-controlled phase 3 trial. Lancet 2016;388(10056):2115–2127.
33. FitzGerald JM, Bleecker ER, Nair P, et al; CALIMA study investigators. Benralizumab, an anti-interleukin-5 receptor α monoclonal antibody, as add-on treatment for patients with severe, uncontrolled, eosinophilic asthma (CALIMA): a randomised, double-blind, placebo-controlled phase 3 trial. Lancet 2016;388(10056):2128–2141.
34. Menzies-Gow A, Gurnell M, Heaney LG, et al. Oral corticosteroid elimination via a personalised reduction algorithm in adults with severe, eosinophilic asthma treated with benralizumab (PONENTE): a multicentre, open-label, single-arm study. Lancet Respir Med 2022;10(1):47–58.

Arthritis & Rheumatology
Vol. 78, No. 3, March 2026, pp 705–712
DOI 10.1002/art.43339

American College of Rheumatology
Empowering Rheumatology Professionals

Distinct Cancer Risk Profiles in Patients With Systemic Sclerosis With Autoantibody Stratification

Arjun Mahajan,[1] Maria Vazquez-Machado,[2] Nikki Zangenah,[3] Jeffrey A. Sparks,[1] and Avery H. LaChance[1]

Objective. Patients with systemic sclerosis (SSc) face increased cancer risk compared to the general population, yet current evidence on specific cancer patterns and their relationship to autoantibody status remain poorly characterized. This study seeks to evaluate cancer risk patterns in patients with SSc and investigate associations between specific autoantibodies and cancer development.

Methods. This multicenter cohort study analyzed five-year cancer incidences in 66,637 adults with SSc versus matched controls with seborrheic keratosis using electronic medical records from 128 health care organizations (from 2014 to 2024). Patients were stratified by autoantibody status (RNA polymerase III, anticentromere, or anti-Scl-70) when available. Primary outcomes included five-year incidences of hematologic and solid-organ cancers, with hazard ratios (HRs) calculated via Cox proportional hazards regression.

Results. Patients with SSc demonstrated elevated five-year all-type cancer risk (HR 1.17, 95% confidence interval [CI] 1.11–1.23). Hematologic cancer risk was significantly increased (HR 1.68, 95% CI 1.50–1.88), particularly for multiple myeloma (HR 2.13, 95% CI 1.61–2.81) and myelodysplastic syndromes (HR 2.03, 95% CI 1.49–2.77). For solid-organ cancers (HR 1.23, 95% CI 1.16–1.31), esophageal cancer showed the highest risk (HR 3.96, 95% CI 2.36–6.65), followed by lung cancer (HR 2.32, 95% CI 2.00–2.69). Among autoantibody subgroups, patients with anti-Scl-70 positivity showed increased overall cancer risk (HR 1.40, 95% CI 1.03–1.92) and patients with RNA polymerase III positivity had higher rates of hematologic cancers (HR 2.20, 95% CI 1.10–4.28), whereas patients with anticentromere positivity demonstrated no increased cancer risks.

Conclusion. Patients with SSc demonstrate significantly increased risks for both hematologic and solid-organ cancers, with risk profiles varying by autoantibody status. These findings suggest the need for targeted cancer screening strategies in SSc and further research to confirm the generalizability of these findings.

INTRODUCTION

Systemic sclerosis (SSc) is a chronic autoimmune disease characterized by multisystem fibrosis affecting the skin and internal organs.[1] Elevated cancer risk in patients with SSc compared to the general population has been well described, with reported rates as high as 11.4% and cancer accounting for approximately 9% of overall mortality rates in patients with SSc.[2,3] However, risk profiles for individual cancers remain poorly characterized, with current evidence limited to case series and smaller cross-sectional studies.[1–3] The lack of large cohort with SSc cancer data and relative rarity of SSc have historically constrained the development of robust, evidence-based guidelines for cancer screening and management in this population.[1,2] The relationship between autoantibody status and cancer risk in SSc also presents a critical knowledge gap. Although some studies have

Supported by the Rheumatology Research Foundation FPS Award. Dr Sparks' work was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases, NIH (grants R01-AR-080659, R01-AR-077607, P30-AR-070253, and P30-AR-072577), and the National Heart, Lung, and Blood Institute, NIH (grant R01-HL-155522); the R. Bruce and Joan M. Mickey Research Scholar Fund; and the Llura Gund Award, funded by the Gordon and Llura Gund Foundation.

[1]Arjun Mahajan, MS, Jeffrey A. Sparks, MD, MMSc, Avery H. LaChance, MD, MPH: Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts; [2]Maria Vazquez-Machado, BS: Brigham and Women's Hospital, Boston, Massachusetts, and Ponce Health Science University School of Medicine, Ponce, Puerto Rico; [3]Nikki Zangenah, BA: Brigham and Women's Hospital and Boston University Chobanian and Avedisian School of Medicine, Boston, Massachusetts.

Authors Vazquez-Machado and Zangenah are co-second authors and contributed equally to this work. Drs Sparks and LaChance are co-senior authors and contributed equally to this work.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/art.43339).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/art.43339.

Address correspondence via email to Avery H. LaChance, MD, MPH, at alachance@bwh.harvard.edu.

Submitted for publication February 27, 2025; accepted in revised form July 7, 2025.

suggested increased cancer risk in patients with anti–RNA polymerase III, anti-RNPC-3, or anti-Scl-70 antibodies, the sparse and inconsistent evidence impedes precise oncologic risk stratification at the individual patient level.[4–6] To address these limitations, we conducted a federated analysis using the TriNetX research network to analyze cancer risk across a large-scale, international cohort of patients with SSc, further stratified by autoantibody status. Through synthesis of longitudinal clinical and serologic data, we aim to help establish a real-world evidence base for risk stratification and early detection of cancer for patients with SSc.

METHODS

Study design and data acquisition. The dataset for this retrospective cohort study was obtained from the TriNetX platform, a Health Insurance Portability and Accountability Act and General Data Protection Regulation–compliant federated network, which provides analytic access to deidentified electronic health records from more than 150 million individuals across member health care organizations in over 20 countries. Data were accessed through the Global Collaborative Network (with natural language processing) subset, encompassing 144 health care organizations. Due to the deidentified nature of the data, the study was deemed exempt by the Mass General Brigham Institutional Review Board. This study adhered to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guidelines.

Participant selection. Eligible participants with SSc were at least 18 years of age and diagnosed with SSc from January 1, 2016, to December 22, 2024, identified using validated International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10) code-based

Table 1. Baseline patient characteristics in cohort with SSc and seborrheic keratosis control cohort*

Characteristic	Before PSM, n (%)			After PSM, n (%)		
	SSc (n = 74,279), mean (±SD)	Control (n = 2,110,168), mean (±SD)	SMD	SSc (n = 66,939), mean (±SD)	Control (n = 66,939), mean (±SD)	SMD
Age at index, y	55.5 (±16.2)	61.6 (±13.8)	0.408	55.8 (±16.2)	56.5 (±15.8)	0.047
Sex						
Female	55,715 (±81.4)	1,126,913 (±55.6)	0.578	54,269 (±81.1)	54,656 (±81.7)	0.015
Male	11,165 (±16.3)	811,559 (±40.0)	0.547	11,090 (±16.6)	10,626 (±15.9)	0.019
Unknown sex	1,583 (±2.3)	88,527 (±4.4)	0.115	1,580 (±2.4)	1,657 (±2.5)	0.007
Race						
American Indian/Alaska Native	467 (±0.7)	6,262 (±0.3)	0.053	446 (±0.7)	475 (±0.7)	0.005
Asian	3,420 (±5.0)	49,911 (±2.5)	0.134	3,266 (±4.9)	3,204 (±4.8)	0.004
Black or African American	8,165 (±11.9)	52,146 (±2.6)	0.367	7,393 (±11.0)	7,464 (±11.2)	0.003
Native Hawaiian/Pacific Islander	371 (±0.5)	25,026 (±1.2)	0.074	366 (±0.5)	381 (±0.6)	0.003
Other race	2,621 (±3.8)	42,856 (±2.1)	0.101	2,550 (±3.8)	2,376 (±3.5)	0.014
White	40,280 (±58.8)	1,578,514 (±77.9)	0.418	39,935 (±59.7)	40,140 (±60.0)	0.006
Unknown race	13,139 (±19.2)	272,284 (±13.4)	0.156	12,983 (±19.4)	12,899 (±19.3)	0.003
Comorbidities						
Chronic kidney disease	3,339 (±4.9)	106,364 (±5.2)	0.017	3,085 (±4.6)	3,590 (±5.4)	0.035
Diabetes mellitus	5,123 (±7.5)	235,521 (±11.6)	0.141	4,948 (±7.4)	5,429 (±8.1)	0.027
Hypertension	13,222 (±19.3)	696,758 (±34.4)	0.345	12,695 (±19.0)	13,297 (±19.9)	0.023
Nicotine dependence	3,065 (±4.5)	156,499 (±7.7)	0.136	2,969 (±4.4)	3,289 (±4.9)	0.023
Hyperlipidemia	9,909 (±14.5)	769,677 (±38.0)	0.554	9,736 (±14.5)	10,281 (±15.4)	0.023
Sjögren disease	2,464 (±3.6)	11,679 (±0.6)	0.213	2,153 (±3.2)	2,697 (±4.0)	0.044
Rheumatoid arthritis	3,852 (±5.6)	38,011 (±1.9)	0.198	3,399 (±5.1)	4,070 (±6.1)	0.044
Systemic lupus erythematosus	4,439 (±6.5)	8,740 (±0.4)	0.336	3,341 (±5.0)	3,447 (±5.1)	0.007
Overweight and obesity	5,081 (±7.4)	289,422 (±14.3)	0.222	4,902 (±7.3)	5,128 (±7.7)	0.013
Chronic viral hepatitis B with delta-agent	10 (±0.0)	137 (±0.0)	0.008	10 (±0.0)	10 (±0.0)	<0.001
Chronic viral hepatitis B without delta-agent	76 (±0.1)	3,117 (±0.2)	0.012	70 (±0.1)	80 (±0.1)	0.004
Chronic viral hepatitis C	326 (±0.5)	10,797 (±0.5)	0.008	312 (±0.5)	357 (±0.5)	0.010
Human immunodeficiency virus disease	79 (±0.1)	6,478 (±0.3)	0.044	78 (±0.1)	74 (±0.1)	0.002
Dermatomyositis	944 (±1.4)	1,953 (±0.1)	0.150	662 (±1.0)	618 (±0.9)	0.007
Neoplasms	8,356 (±12.2)	937,449 (±46.2)	0.807	8,239 (±12.3)	8,573 (±12.8)	0.015
Medications						
Methotrexate	2,700 (±3.9)	22,602 (±1.1)	0.181	2,354 (±3.5)	2,839 (±4.2)	0.038
Mycophenolate mofetil	3,297 (±4.8)	13,222 (±0.7)	0.257	2,508 (±3.7)	2,467 (±3.7)	0.003
Cyclophosphamide	319 (±0.5)	7,253 (±0.4)	0.017	245 (±0.4)	237 (±0.4)	0.002
Tacrolimus	869 (±1.3)	28,390 (±1.4)	0.011	822 (±1.2)	1,133 (±1.7)	0.039
Azathioprine	1,330 (±1.9)	8,454 (±0.4)	0.142	1,062 (±1.6)	1,180 (±1.8)	0.014

* PSM, propensity score matching; SMD, standardized mean difference; SSc, systemic sclerosis.

approaches previously demonstrated to have a positive predictive value of ≥84% to 93%.[7,8] Subgroup analyses were performed by further stratifying the total cohort with SSc into three groups based on the presence of RNA polymerase III, anticentromere, or anti-Scl-70 antibodies, when available. A control cohort was selected, consisting of patients with seborrheic keratosis ICD-10 codes, which was chosen due to its prevalence and lack of known significant associations with cancer risk.[9]

Exclusion criteria included individuals diagnosed with any cancer before or on the index day of SSc diagnosis. These included malignancies of the lip, oral cavity and pharynx, digestive organs, respiratory and intrathoracic organs, bone and articular cartilage, melanoma, mesothelial and soft tissue, breast, female genital organs, male genital organs, urinary tract, eye, brain and other parts of central nervous system, thyroid and other endocrine glands, lymphoid, hematopoietic and related tissue, and neuroendocrine tumors. Follow-up for all cohorts extended until the onset of cancer, withdrawal from the TriNetX database, death, or the study's conclusion on December 22, 2024.

To facilitate comparative analysis, propensity score matching (1:1 using a greedy nearest neighbor algorithm) was performed on the two cohorts to balance characteristics like age, sex, race and ethnicity, and comorbidities such as hypertension, hyperlipidemia, diabetes mellitus, chronic kidney disease, and autoimmune diseases such as rheumatoid arthritis and Sjögren disease (full list in Table 1). Further details on cohort creation and definitions are provided in Supplementary Table 1.

Statistical analysis. Primary outcomes were five-year incidences of hematologic and solid-organ cancers. The follow-up period was selected based on the availability of follow-up data to ensure reliability and minimize potential bias in outcome analysis. Varicose vein and trauma (laceration of the head) were selected as negative control outcomes. A landmark was set at one year to preclude outcomes occurring before one year after the index event to lower the potential for reverse causation in which evaluation for suspected cancer may have resulted in diagnosis of SSc. A sensitivity analysis was also performed without this landmark restriction (Supplementary Tables 1 and 2). Outcomes and comorbidities were identified using ICD-10 codes that have been

Table 2. Occurrence and risk of malignancies in patients with SSc versus seborrheic keratosis controls in a propensity score–matched analysis*

Outcome	Outcome events of cohort with SSc, n (n = 66,939)	Outcome events of control cohort, n (n = 66,939)	Hazard ratio (95% CI)
All-type cancer	2,896	2,573	1.17 (1.11–1.23)
Hematologic malignancies	800	498	1.68 (1.50–1.88)
Hodgkin lymphoma	36	24	1.56 (0.93–2.62)
Non-Hodgkin lymphoma	145	80	1.89 (1.44–2.49)
Lymphoma	317	177	1.87 (1.55–2.24)
Leukemia	207	153	1.41 (1.14–1.74)
Myelodysplastic syndromes	116	60	2.03 (1.49–2.77)
Multiple myeloma	151	74	2.13 (1.61–2.81)
Solid-organ malignancies	2,142	1,810	1.23 (1.16–1.31)
Malignancy of digestive tract	477	350	1.43 (1.24–1.64)
Esophageal cancer	69	18	3.96 (2.36–6.65)
Colon cancer	159	145	1.15 (0.92–1.44)
Pancreatic cancer	82	57	1.51 (1.08–2.11)
Liver cancer	107	71	1.58 (1.17–2.13)
Malignancy of lip, oral cavity, and pharynx	123	82	1.56 (1.18–2.06)
Nasopharynx cancer	10	13	0.72 (0.31–1.69)
Breast cancer	577	739	0.81 (0.73–0.90)
Lung cancer	557	252	2.32 (2.00–2.69)
Kidney cancer	96	93	1.08 (0.81–1.43)
Bladder cancer	82	94	0.91 (0.68–1.23)
Ovarian cancer	74	74	1.04 (0.75–1.44)
Malignancy of female genital organs	233	246	0.99 (0.83–1.18)
Thyroid cancer	111	107	1.08 (0.83–1.41)
Brain cancer	54	39	1.44 (0.96–2.18)
Melanoma	133	438	0.31 (0.26–0.38)
Negative controls			
Varicose veins	342	326	1.10 (0.94–1.28)
Laceration of the head	149	191	0.82 (0.66–1.02)

* CI, confidence interval; SSc, systemic sclerosis.

validated in prior research.[10–15] Relative risks, risk difference, and hazard ratios (HRs) and their 95% confidence intervals (CIs) were calculated, the lattermost estimated using Cox proportional hazards regression. Statistical analyses were conducted within the TriNetX platform using Python version 3.7 (Python Software Foundation).

RESULTS

Demographics. Before matching, patients in the cohort with SSc were younger (mean age [±SD], 55.5 [±16.2] vs 61.6 [±13.8] years), less likely to be of White race (58.8% vs 77.9%), and more likely to be female (81.4% vs 55.6%). The prematched cohort with SSc had higher baseline rates of autoimmune comorbidities such as systemic lupus erythematosus (6.5% vs 0.4%) and Sjögren disease (3.6% vs 0.6%) but lower rates of nicotine dependence (4.5% vs 7.7%) and cardiovascular risk factors such as hypertension (19.3% vs 34.4%) and hyperlipidemia (14.5% vs 38.0%). After propensity score matching, 66,939 patients were included in each cohort, with no statistically significant differences in baseline demographics and covariates between the treatment and control groups. Median follow-up time for both matched cohorts was 1,171 days for SSc and 1,826 days for the control group.

All-type cancer risk in overall cohort with SSc. Patients with SSc had an increased overall five-year risk of cancer (Table 2) compared to controls (HR 1.17, 95% CI 1.11–1.23). For the negative control outcomes (varicose veins and trauma diagnoses), there were no clinically relevant differences between groups, in both overall and individual autoantibody cohorts.

Hematologic cancer risk in overall cohort with SSc. We identified an increased five-year risk of hematologic cancers (Table 2) in patients with SSc compared to controls (HR 1.68, 95% CI 1.50–1.88), including non-Hodgkin lymphoma (HR 1.89, 95% CI 1.44–2.49), lymphoma (HR 1.87, 95% CI 1.55–2.24), leukemia (HR 1.41, 95% CI 1.14–1.74), myelodysplastic syndromes (HR 2.03, 95% CI 1.49–2.77), and multiple myeloma (HR 2.13, 95% CI 1.61–2.81). In a sensitivity analysis without the one-year landmark period requirement, the statistical significance of all previously reported outcomes was unchanged, though Hodgkin lymphoma additionally reached significance (HR 1.52, 95% CI 1.08–2.08).

Solid-organ cancer risk in overall cohort with SSc. Patients with SSc demonstrated a higher five-year risk of solid-organ cancers (Table 2) compared to controls (HR 1.23, 95% CI

Table 3. Occurrence and risk of malignancies in patients with SSc with anti-Scl-70 antibody positivity versus seborrheic keratosis controls in a propensity score–matched analysis*

Outcome	SSc (anti-Scl-70) events, n (n = 1,246)	Control events, n (n = 1,246)	Hazard ratio (95% CI)
All-type cancer	74	49	1.40 (1.03–1.92)
Hematologic malignancies	25	≤10	2.81 (1.37–5.80)
Hodgkin lymphoma	≤10	≤10	6.21 (0.76–50.47)
Non-Hodgkin lymphoma	14	≤10	6.30 (1.18–38.12)
Lymphoma	19	≤10	3.44 (1.27–11.76)
Leukemia	≤10	≤10	3.14 (1.01–9.77)
Myelodysplastic syndromes	≤10	≤10	0.90 (0.12–6.75)
Multiple myeloma	16	≤10	7.26 (1.26–41.92)
Solid-organ malignancies	53	36	1.37 (1.02–1.94)
Malignancy of digestive tract	11	≤10	1.12 (0.46–2.75)
Esophageal cancer	≤10	≤10	1.82 (0.16–21.28)
Colon cancer	≤10	≤10	1.20 (0.26–5.51)
Pancreatic cancer	≤10	≤10	2.74 (0.11–70.88)
Liver cancer	≤10	≤10	0.54 (0.07–4.21)
Malignancy of lip, oral cavity, and pharynx	≤10	≤10	1.52 (0.19–12.08)
Nasopharynx cancer	≤10	0	–
Breast cancer	18	16	1.01 (0.51–2.00)
Lung cancer	21	≤10	2.23 (1.21–4.81)
Kidney cancer	≤10	≤10	2.06 (0.37–11.33)
Bladder cancer	≤10	≤10	0.90 (0.15–5.38)
Ovarian cancer	≤10	≤10	0.31 (0.01–8.48)
Malignancy of female genital organs	≤10	≤10	1.83 (0.39–8.58)
Thyroid cancer	≤10	≤10	2.23 (0.42–11.89)
Brain cancer	≤10	≤10	0.90 (0.09, 9.10)
Melanoma	≤10	≤10	0.35 (0.13–1.82)
Negative controls			
Varicose veins	13	11	1.42 (0.59–3.40)
Laceration of the head	≤10	≤10	1.40 (0.36–5.50)

* CI, confidence interval; SSc, systemic sclerosis.

Table 4. Occurrence and risk of malignancies in patients with SSc with RNA polymerase III antibody positivity versus seborrheic keratosis controls in a propensity score–matched analysis*

Outcome	SSc (RNA polymerase III) events, n (n = 592)	Control events, n (n = 592)	Hazard ratio (95% CI)
All-type cancer	43	30	1.25 (0.74–2.12)
Hematologic malignancies	22	≤10	2.20 (1.10–4.28)
Hodgkin lymphoma	≤10	≤10	1.10 (0.02–60.24)
Non-Hodgkin lymphoma	≤10	≤10	2.90 (0.29–29.20)
Lymphoma	≤10	≤10	0.60 (0.08–4.72)
Leukemia	≤10	≤10	1.52 (0.12–20.11)
Myelodysplastic syndromes	≤10	≤10	3.04 (0.30–30.62)
Multiple myeloma	≤10	≤10	4.86 (0.22–107.90)
Solid-organ malignancies	21	18	1.18 (0.63–2.23)
Malignancy of digestive tract	≤10	≤10	1.27 (0.33–4.85)
Esophageal cancer	0	≤10	–
Colon cancer	≤10	≤10	1.45 (0.73–2.89)
Pancreatic cancer	≤10	≤10	0.52 (0.19–1.46)
Liver cancer	≤10	≤10	2.14 (0.56–8.18)
Malignancy of lip, oral cavity, and pharynx	≤10	≤10	0.74 (0.17–3.32)
Lung cancer	≤10	≤10	1.61 (0.32–8.08)
Breast cancer	≤10	≤10	0.50 (0.16–1.59)
Kidney cancer	≤10	≤10	3.88 (0.16–91.78)
Bladder cancer	≤10	≤10	1.01 (0.06–17.17)
Ovarian cancer	≤10	0	–
Thyroid cancer	≤10	≤10	8.32 (1.04–66.53)
Malignancy of female genital organs	≤10	≤10	1.42 (0.41–7.56)
Melanoma	≤10	≤10	0.22 (0.05–1.01)
Negative controls			
Varicose veins	≤10	≤10	2.51 (0.24–26.79)
Laceration of the head	≤10	≤10	2.06 (0.06–65.93)

* CI, confidence interval; SSc, systemic sclerosis.

1.16–1.31), including cancers of the digestive tract (HR 1.43, 95% CI 1.24–1.64); esophageal cancer (HR 3.96, 95% CI 2.36–6.65); pancreatic cancer (HR 1.51, 95% CI 1.08–2.11); liver cancer (HR 1.58, 95% CI 1.17–2.13); malignancy of lip, oral cavity, and pharynx (HR 1.56, 95% CI 1.18–2.06); and lung cancer (HR 2.32, 95% CI 2.00–2.69). Patients with SSc had a decreased risk of breast cancer (HR 0.81, 95% CI 0.73–0.90) and melanoma (HR 0.31, 95% CI 0.26–0.38).

Cancer risk in cohort with anti-Scl-70 antibody positivity. After matching, 1,246 patients were included in both the group with SSc with anti-Scl-70 autoantibody positivity and control group. Patients with anti-Scl-70 antibodies showed higher rates of cancer overall (HR 1.40, 95% CI 1.03–1.92) and hematologic cancers (HR 2.81, 95% CI 1.37–5.80; Table 3). Among hematologic cancers, there were significantly increased risks of non-Hodgkin lymphoma (HR 6.30, 95% CI 1.18–38.12), lymphoma (HR 3.44, 95% CI 1.27–11.76), leukemia (HR 3.14, 95% CI 1.01–9.77), and multiple myeloma (HR 7.26, 95% CI 1.26–41.92). For solid-organ cancers, both overall (HR 1.37, 95% CI 1.02–1.94) and lung cancer (HR 2.23, 95% CI 1.21–4.81) risks were increased.

Cancer risk in cohort with RNA polymerase III antibody positivity. After matching, 592 patients were included in both the group with SSc with RNA polymerase III positivity and control group. Patients with RNA polymerase III antibodies showed significantly higher rates of hematologic cancers (HR 2.20, 95% CI 1.10–4.28) and thyroid cancer (HR 8.32, 95% CI 1.04–66.53; Table 4). In a sensitivity analysis excluding the landmark period requirement, patients with RNA polymerase III positivity demonstrated elevated rates of breast cancer compared to controls (HR 1.20, 95% CI 1.02–2.55; Supplementary Table 2). All other cancers showed no statistically significant differences between the groups.

Cancer risk in cohort with anticentromere antibody positivity. After matching, 1,032 patients were included in both the group with SSc with anticentromere autoantibody positivity and control group. Among all hematologic and solid-organ cancers, melanoma was the only cancer type that demonstrated a statistically significant difference (HR 0.26, 95% CI 0.08–0.87; Table 5).

DISCUSSION

This study offers a comprehensive and longitudinal analysis of individual hematologic and solid-organ cancer risks in patients with SSc, offering real-world evidence for risk stratification and may provide a framework for targeted screening and evaluation

Table 5. Occurrence and risk of malignancies in patients with SSc with anticentromere antibody positivity versus seborrheic keratosis controls in a propensity score–matched analysis*

Outcome	SSc (anticentromere) events, n (n = 1,032)	Control events, n (n = 1,032)	Hazard ratio (95% CI)
All-type cancer	42	48	0.89 (0.59–1.35)
Hematologic malignancies	12	11	1.21 (0.51–2.89)
Hodgkin lymphoma	0	0	–
Non-Hodgkin lymphoma	≤10	≤10	2.96 (0.31–28.46)
Lymphoma	≤10	≤10	0.86 (0.19–4.03)
Leukemia	≤10	≤10	1.30 (0.29–5.80)
Myelodysplastic syndromes	≤10	≤10	1.99 (0.18–21.89)
Multiple myeloma	≤10	≤10	2.95 (0.31–28.31)
Solid-organ malignancies	33	34	0.97 (0.60–1.57)
Malignancy of digestive tract	12	≤10	1.43 (0.58–3.53)
Esophageal cancer	≤10	≤10	2.92 (0.12–71.59)
Colon cancer	≤10	≤10	1.98 (0.36–10.79)
Pancreatic cancer	≤10	≤10	0.66 (0.05–8.32)
Liver cancer	≤10	≤10	1.73 (0.31–9.83)
Malignancy of lip, oral cavity, and pharynx	≤10	≤10	0.74 (0.08–6.17)
Nasopharynx cancer	0	0	–
Breast cancer	11	15	0.69 (0.31–1.51)
Lung cancer	≤10	≤10	1.20 (0.35–4.18)
Kidney cancer	≤10	≤10	1.61 (0.21–12.18)
Bladder cancer	≤10	≤10	0.98 (0.14–6.93)
Ovarian cancer	≤10	≤10	1.96 (0.07–58.36)
Malignancy of female genital organs	≤10	≤10	0.98 (0.17–5.71)
Thyroid cancer	≤10	≤10	1.95 (0.18–21.48)
Brain cancer	≤10	≤10	1.48 (0.19–18.63)
Melanoma	≤10	19	0.26 (0.08, 0.87)
Negative controls			
Varicose veins	19	20	0.94 (0.39–2.27)
Laceration of the head	≤10	≤10	1.39 (0.28–7.08)

* CI, confidence interval; SSc, systemic sclerosis.

of cancer. Patients with SSc had a significantly increased risk of developing any cancer compared to controls within five years of diagnosis, consistent with evidence from prior smaller cohort studies.[2,16,17] The risk of esophageal cancer was particularly elevated in the group with SSc, potentially exceeding what has been reported in several prior studies.[16,18] Cancer risk varied by SSc-related autoantibody, with the highest hematologic cancer risk for those with anti-Scl-70 positivity as well as RNA polymerase III positivity. Overall, these findings underscore the need for heightened clinical awareness of cancer risk in patients with SSc and support further research into the potential role of targeted cancer screening for select patient subgroups.

Within hematologic cancers, we observed notable increases in multiple myeloma, myelodysplastic syndromes, and non-Hodgkin lymphoma risk and more modest increases in leukemia but no increased risk of Hodgkin lymphoma in the primary analysis. In patients with SSc, chronic immune dysfunction, sustained B cell activation, and epigenetic alterations may contribute to the generation of lymphoproliferative disorders, consistent with the observed increased risk of myeloid malignancies.[4,19] The lack of observed increased risk in Hodgkin lymphoma may reflect a distinct pathogenesis and/or the relatively low event rates.[20] Prior reports have identified an increased association with hematologic neoplasms overall in patients with SSc compared to average population rates, supporting our findings; however, few have analyzed sufficient cohort sizes to identify risks of individual hematologic cancers.[16,17]

Within solid-organ cancers, we observed markedly elevated risks for lung and esophageal cancers, with more modest but significant increases in liver cancer, pancreatic cancer, and cancers of the oral cavity and pharynx. These findings are consistent with prior meta-analyses and registry studies demonstrating increased risks in these organ systems, particularly for lung and esophageal cancer.[5,17,21] In contrast, we found reduced risks of breast cancer and melanoma, with no significant risk differences for remaining solid-organ cancers. The elevated risk of cancers in organs commonly affected by inflammation and fibrosis in SSc—particularly the lungs and esophagus—suggests that chronic inflammation, tissue damage, and fibrotic remodeling may create a permissive environment for malignant transformation.[22,23] The decreased risk of melanoma may be attributable to the increased dermatologic surveillance rate in the control group (seborrheic keratosis), resulting in higher rates of melanoma diagnosis in this group. The reduced breast cancer risk in patients with SSc may reflect competing death risks, inherent disease pathogenesis and risk profile, and/or differential health care use patterns among cohorts. However, other studies have observed increased breast cancer risk in patients with SSc, whereas two meta-analyses did

not find a significant relationship.[17,24–26] Thus, additional research is needed to further understand the potential correlation between SSc and breast cancer risk. The lack of elevated risk in several solid-organ cancers should also be interpreted in the context of potential methodologic considerations, such as insufficient follow-up time for slow-developing solid-organ cancers to develop and potential survival bias given competing death risks in patients with SSc.

Though limited by sample size, analysis of autoantibody profiles in SSc demonstrated differing patterns of cancer risk. Patients with anti-Scl-70 positivity demonstrated the most extensive risk profile, with increased risks of both hematologic cancers (particularly lymphomas, leukemia, and multiple myeloma) and solid-organ cancers, particularly lung cancer. However, it is important to contextualize these findings because several cohort studies have not observed a consistent association between anti-Scl-70 positivity and overall cancer risk, whereas others have identified anti-Scl-70 positivity as a risk factor for lung cancer development.[21,27,28] The broad cancer risk observed in our study may reflect a severe inflammatory response and extensive tissue damage characteristic of the diffuse SSc phenotype often associated with anti-Scl-70 antibodies.[29] Patients with RNA polymerase III antibody positivity, though representing the smallest cohort, still demonstrated elevated risks of hematologic cancers and thyroid cancer, the latter of which has not been extensively documented in prior studies.[28,30,31] Multiple studies have demonstrated that patients with SSc with anti–RNA polymerase III antibodies have higher risks of incident cancer, with a temporal clustering between cancer and SSc onset.[30,31] In our sensitivity analysis without a landmark period exclusion, breast cancer emerged as an elevated risk, consistent with evidence from several prior studies.[25,31] In contrast, patients with anticentromere antibody positivity showed no increased cancer risks; this may reflect milder disease severity often associated with limited cutaneous SSc and/or sample size limitations, which may have masked the presence of a positive association.[32] To our knowledge, case series and smaller observational studies have reported either limited association between anticentromere antibody positivity and cancer risk or decreased cancer rates compared to those seen in patients with anti–RNA polymerase III positivity, though data on site-specific cancer risks remain limited.[28,30,32] Although autoantibody-specific subgroup sizes reflect the constraints of disease rarity, this study represents one of the largest retrospective studies to date on SSc, enabling more granular risk assessment than previously possible.

Collectively, these findings suggest that clinicians should maintain heightened awareness of cancer risk in patients with SSc. Based on our findings, pragmatic cancer surveillance considerations may include the following: ensuring adherence to age-appropriate standard cancer screenings; thorough symptom assessment at each clinical encounter, particularly for the specific cancers occurring at significantly higher rates in patients with SSc in our analysis; lower threshold for diagnostic evaluation of concerning symptoms; and consideration of additional targeted screening based on specific antibody profiles (ie, RNA polymerase III with anti-Scl-70) and risk factors. Prospective studies are needed to determine the optimal intervals, modalities, and risk-benefit tradeoffs of specific cancer screening practices for patients with SSc.

This study's strengths include its substantial cohort size spanning multiple institutions and continents internationally to enhance statistical power and generalizability, extensive covariate matching to mitigate bias from confounding variables, landmark analysis to reduce reverse confounding, and negative control analysis to support the validity of our findings. However, several limitations warrant careful consideration, including potential unmeasured confounders, inherent observational experiment-related selection biases, and cohort and outcome misclassification risk from structured clinical data and ICD-10 codes. The incomplete autoantibody data in our study population may influence study findings because documented anticentromere, anti-Scl-70, and anti–RNA polymerase III status were available in a smaller proportion of patients than may be expected based on their estimated prevalence in SSc.[1,27] Additionally, the variability in autoantibody testing methods across contributing health care organizations may introduce measurement heterogeneity and potential misclassification, compared to specialized cohort studies with standardized autoantibody assessment protocols. Our definition of RNA polymerase III positivity as requiring either RP11 or RP155 antibodies, rather than both, may impact specificity, though this approach was chosen to align with clinical practice patterns and maintain adequate cohort size for meaningful analysis.[30,31] Our analysis is also limited by a lack of data on SSc subtypes, which prior literature has identified as potentially relevant factors for cancer risk stratification, as well as other relevant disease-specific factors that may influence cancer risk in this population.[5,16] The TriNetX platform used primarily comprises electronic health record data without comprehensive linkage to claims databases, potentially leading to missed cancer outcomes when patients receive care across multiple unaffiliated health care systems. Finally, factors occurring after baseline such as immunosuppressant use, disease activity from inflammation/fibrosis, or organ damage from SSc may have mediated associations and should be investigated in future studies.

Leveraging data from a large multi-institutional cohort, our study identified increased longitudinal risks of solid-organ and hematologic cancers in patients with SSc, further stratified by autoantibody profile. In identifying the risk of individual cancers as well as their corresponding autoantibody patterns, our research may enable clinicians to be more proactive and targeted in implementing strategies for early detection, management, and prevention of significant cancer-associated morbidity and mortality rates for patients with SSc.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr LaChance confirms that all authors have provided the final approval of the version to be published and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. Bergamasco A, Hartmann N, Wallace L, et al. Epidemiology of systemic sclerosis and systemic sclerosis-associated interstitial lung disease. Clin Epidemiol 2019;11:257–273.
2. Ciołkiewicz M, Domysławska I, Ciołkiewicz A, et al. Coexistence of systemic sclerosis, scleroderma-like syndromes and neoplastic diseases. Pol Arch Med 2008;118(3):119–126.
3. Hao Y, Hudson M, Baron M, et al; Canadian Scleroderma Research Group; Australian Scleroderma Interest Group. Early mortality in a multinational systemic sclerosis inception cohort. Arthritis Rheumatol 2017;69(5):1067–1077.
4. Lazzaroni MG, Airò P. Anti-RNA polymerase III antibodies in patients with suspected and definite systemic sclerosis: why and how to screen. J Scleroderma Relat Disord 2018;3(3):214–220.
5. Maria ATJ, Partouche L, Goulabchand R, et al. Intriguing relationships between cancer and systemic sclerosis: role of the immune system and other contributors. Front Immunol 2019;9:3112.
6. Lazizi S, Hudson M, Baron M, et al. Associations between autoantibodies in systemic sclerosis and cancer in a national registry. Arthritis Rheumatol 2020;72(suppl 10). https://acrabstracts.org/abstract/associations-between-autoantibodies-in-systemic-sclerosis-and-cancer-in-a-national-registry/. Accessed December 30, 2024.
7. Jamian L, Wheless L, Crofford LJ, et al. Rule-based and machine learning algorithms identify patients with systemic sclerosis accurately in the electronic health record. Arthritis Res Ther 2019;21(1):305.
8. De Almeida Chaves S, Derumeaux H, Do Minh P, et al. Assessment of the accuracy of using *ICD-10* codes to identify systemic sclerosis. Clin Epidemiol 2020;12:1355–1359.
9. Lindelöf B, Sigurgeirsson B, Melander S. Seborrheic keratoses and cancer. J Am Acad Dermatol 1992;26(6):947–950.
10. Nedkoff L, Lopez D, Hung J, et al. Validation of ICD-10-AM coding for myocardial infarction subtype in hospitalisation data. Heart Lung Circ 2022;31(6):849–858.
11. Grams ME, Plantinga LC, Hedgeman E, et al; CDC CKD Surveillance Team. Validation of CKD and related conditions in existing data sets: a systematic review. Am J Kidney Dis 2011;57(1):44–54.
12. Park J, Kwon S, Choi EK, et al. Validation of diagnostic codes of major clinical outcomes in a National Health Insurance database. Int J Arrhythmia 2019;20(1):5.
13. Lee H, Sparks JA, Lee SB, et al. Validation of serostatus of rheumatoid arthritis using ICD-10 codes in administrative claims data. Pharmacoepidemiol Drug Saf 2023;32(5):586–591.
14. Goetz I, Choong C, Winnie J, et al. Development of a claims-based flare algorithm for systemic lupus erythematosus. Curr Med Res Opin 2022;38(9):1641–1649.
15. Dros JT, Bos I, Bennis FC, et al. Detection of primary Sjögren's syndrome in primary care: developing a classification model with the use of routine healthcare data and machine learning. BMC Prim Care 2022;23(1):199.
16. Lepri G, Catalano M, Bellando-Randone S, et al. Systemic sclerosis association with malignancy. Clin Rev Allergy Immunol 2022;63(3):398–416.
17. Bonifazi M, Tramacere I, Pomponio G, et al. Systemic sclerosis (scleroderma) and cancer risk: systematic review and meta-analysis of observational studies. Rheumatology (Oxford) 2013;52(1):143–154.
18. Derk CT, Rasheed M, Artlett CM, et al. A cohort study of cancer incidence in systemic sclerosis. J Rheumatol 2006;33(6):1113–1116.
19. Thoreau B, Chaigne B, Mouthon L. Role of B-cell in the pathogenesis of systemic sclerosis. Front Immunol 2022;13:933468.
20. Weniger MA, Küppers R. Molecular biology of Hodgkin lymphoma. Leukemia 2021;35(4):968–981.
21. Colaci M, Giuggioli D, Sebastiani M, et al. Lung cancer in scleroderma: results from an Italian rheumatologic center and review of the literature. Autoimmun Rev 2013;12(3):374–379.
22. Hassoun PM. Lung involvement in systemic sclerosis. Presse Med 2011;40(1 Pt 2):e3–e17.
23. Vijayraghavan S, Blouin T, McCollum J, et al. Widespread mutagenesis and chromosomal instability shape somatic genomes in systemic sclerosis. Nat Commun 2024;15(1):8889.
24. Zhang JQ, Wan YN, Peng WJ, et al. The risk of cancer development in systemic sclerosis: a meta-analysis. Cancer Epidemiol 2013;37(5):523–527.
25. Colaci M, Giuggioli D, Vacchi C, et al. Breast cancer in systemic sclerosis: results of a cross-linkage of an Italian rheumatologic center and a population-based cancer registry and review of the literature. Autoimmun Rev 2014;13(2):132–137.
26. Morrisroe K, Hansen D, Huq M, et al. Incidence, risk factors, and outcomes of cancer in systemic sclerosis. Arthritis Care Res (Hoboken) 2020;72(11):1625–1635.
27. Shah AA, Hummers LK, Casciola-Rosen L, et al. Examination of autoantibody status and clinical features associated with cancer risk and cancer-associated scleroderma. Arthritis Rheumatol 2015;67(4):1053–1061.
28. Watad A, McGonagle D, Bragazzi NL, et al. Autoantibody status in systemic sclerosis patients defines both cancer risk and survival with ANA negativity in cases with concomitant cancer having a worse survival. Oncoimmunology 2019;8(6):e1588084.
29. Basu D, Reveille JD. Anti-scl-70. Autoimmunity 2005;38(1):65–72.
30. Moinzadeh P, Fonseca C, Hellmich M, et al. Association of anti-RNA polymerase III autoantibodies and cancer in scleroderma. Arthritis Res Ther 2014;16(1):R53.
31. Shah AA, Rosen A, Hummers L, et al. Close temporal relationship between onset of cancer and scleroderma in patients with RNA polymerase I/III antibodies. Arthritis Rheum 2010;62(9):2787–2795.
32. Perosa F, Prete M, Di Lernia G, et al. Anti-centromere protein A antibodies in systemic sclerosis: significance and origin. Autoimmun Rev 2016;15(1):102–109.

Arthritis & Rheumatology
Vol. 78, No. 3, March 2026, pp 713–723
DOI 10.1002/art.43408

American College of Rheumatology
Empowering Rheumatology Professionals

Identification of Risk Factors for Incident Left Ventricular Systolic Dysfunction and Predictors of Cardiac Recovery in Patients With Systemic Sclerosis

Ji Soo Kim,[1,2] Rachel S. Wallwork,[1] Carrie L. Richardson,[3] Adrianne Woods,[1] Monica Mukherjee,[4] Steven Hsu,[4] Julie J. Paik,[1] Christopher A. Mecoli,[1] Laura K. Hummers,[1] Fredrick M. Wigley,[1] Scott L. Zeger,[2] and Ami A. Shah[1]

Objective. Cardiac involvement in systemic sclerosis (SSc) is a leading cause of death. We sought to investigate predictors of incident left ventricular systolic dysfunction (LVSD) and cardiac recovery in SSc.

Methods. In total, 2,303 patients in the Johns Hopkins Scleroderma Center Research Registry and 13,209 echocardiograms were analyzed. We identified predictors associated with incident LVSD defined by transitions in left ventricular ejection fraction (EF) states (EF ≥50% declining to <50% and EF >35% dropping to ≤35% [severe LVSD]) by fitting multivariate logistic regression models with time-varying and invariant variables. Variables associated with cardiac recovery were identified by fitting multivariate logistic regression models using important variables identified from random forest analysis.

Results. Male sex, Black race, diffuse skin disease, higher modified Rodnan skin score (mRSS), echocardiographic evidence of pulmonary hypertension (PH), kidney disease, and atrial fibrillation (AFib) were associated with increased odds of incident LVSD (EF < 50%), whereas anticentromere and anti–topoisomerase-1 were protective. Male sex, higher mRSS, PH, skeletal myopathy, kidney disease, AFib, and anti-Ku antibodies were associated with higher odds of incident severe LVSD (EF ≤ 35%). For previous EF <50%, tendon friction rubs were associated with lower odds of cardiac recovery and anti-RNA polymerase III (anti–RNAP III) was associated with higher odds. For previous EF ≤35%, diabetes was associated with lower odds of recovering to EF >35%.

Conclusion. Distinct demographic SSc-specific and cardiac characteristics are associated with increased risk of incident LVSD in SSc, with skeletal myopathy and anti-Ku antibodies being important risk factors for severe disease. Some patients improved, which was more likely in anti–RNAP III positive patients.

INTRODUCTION

Cardiovascular involvement is the third-most common cause of systemic sclerosis (SSc; scleroderma)–related death after pulmonary arterial hypertension (PAH) and interstitial lung disease (ILD), accounting for 12% to 30% of SSc-related death.[1–3] Left ventricular (LV) dysfunction, across various ranges of systolic dysfunction, significantly contributes to morbidity and mortality in patients with SSc and may be an underrecognized cause of adverse clinical outcomes. Patients with severely reduced LV ejection fraction (LVEF ≤ 35%) may particularly benefit from automatic implantable cardioverter-defibrillator (placement) or other goal-directed medical therapies.[4,5] However, identifying patients at the highest risk of developing cardiac involvement is a key clinical challenge. Despite the clinical importance, there are no large prospective

Supported in part by the National Institute of Arthritis and Musculoskeletal and Skin Diseases, NIH (grants P30-AR-070254, K24-AR-080217, K23-AR-073927, T32-AR-048522), National Heart, Lung and Blood Institute, NIH (grant R01-HL-162851), the Donald B. and Dorothy L. Stabler Foundation, the Jerome L. Greene Foundation, the Chresanthe Staurulakis Memorial Discovery Fund, the Martha McCrory Professorship, the Rheumatology Research Foundation, the Scleroderma Research Foundation, the National Scleroderma Foundation, the US Department of Defense (grant PR231648), and the Johns Hopkins in Health initiative.

[1]Division of Rheumatology, Johns Hopkins University School of Medicine, Baltimore, Maryland; [2]Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland; [3]Division of Rheumatology, Northwestern University Chicago, Chicago, Illinois; [4]Division of Cardiology, Johns Hopkins University, Baltimore, Maryland.

Drs Kim and Wallwork are co-first authors and contributed equally to this work.

Additional supplementary information cited in this article can be found online in the Supporting Information section (http://onlinelibrary.wiley.com/doi/10.1002/art.43408).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/art.43408.

Address correspondence via email to Ami A. Shah, MD, MHS, at Ami.Shah@jhmi.edu; or to Ji Soo Kim, PhD, at jkim478@jhu.edu.

Submitted for publication December 6, 2024; accepted in revised form September 16, 2025.

studies exploring risk of incident LV systolic dysfunction (LVSD) or factors associated with cardiac recovery in SSc.

Prior studies report varying rates of cardiac involvement and associated risk factors in SSc, ranging from clinically evident heart failure to abnormalities detected on cardiac magnetic resonance imaging (MRI), including LVEF values that are considered within the normal range per the American Society of Echocardiography (ASE)/European Association Cardiovascular Imaging (EACVI) guidelines.[4–8] These studies have variably identified age, sex, SSc disease duration, the extent of skin thickening on physical examination, the presence of ischemic digital ulcers, and myositis as risk factors for cardiac disease in SSc.[4,5,9,10] These findings suggest that clinical manifestations may accurately identify patients with SSc at increased risk of cardiac complications. Of note, none of these studies used time-varying risk factors, which may change over the course of a patient's disease or as patients age, nor did they include other important cardiac risk factors, such as atrial fibrillation (AFib) or coronary artery disease (CAD).

The primary objective of this study was to use data-driven approaches to identify demographic, SSc-specific, and cardiac risk factors for incident LVSD (LVEF < 50%) and severe LVSD (EF ≤ 35%) in SSc.[8] By using time-varying predictors, we aimed to better understand the risk factors most proximally associated with LVSD. Additionally, we explored factors associated with cardiac recovery in a large well-phenotyped, longitudinal SSc cohort.

PATIENTS AND METHODS

Study population and key variables. All patients enrolled in the Johns Hopkins Scleroderma Center (JHSC) Research Registry from October 1980 to July 2017 with at least two echocardiograms were included in this study. All included patients had SSc defined by 1980 or 2013 American College of Rheumatology (ACR)/EULAR classification criteria, at least three of five CREST (calcinosis, Raynaud phenomenon, esophageal dysmotility, sclerodactyly, telangiectasia) syndrome criteria, or definite Raynaud phenomenon, abnormal nailfold capillaries, and a SSc-specific autoantibody.[11,12] Demographics, disease onset, phenotypic characteristics, and modified Medsger organ-specific severity scores were collected at baseline and longitudinally in the context of routine clinical care for patients who consented to enroll in the JHSC Registry. Autoantibody data and serial internally and externally performed pulmonary function tests and echocardiograms were also collected. Given the risk of cardiopulmonary involvement in SSc, it is the JHSC practice to obtain annual echocardiograms on all patients.

A complete list of candidate variables is in the Supplement. Age at SSc onset was defined as the age at the patient's first non-Raynaud symptom. Patient-reported smoking status was defined as any history of smoking versus no history of smoking at baseline. Race was self-reported. Skin involvement was defined as limited cutaneous SSc if skin thickening was present only distal to the elbows or knees with or without facial involvement and was defined as diffuse cutaneous SSc if skin thickening was ever present proximal to the elbows or knees or involved the trunk.[13]

Disease duration was defined as the time from first non-Raynaud symptom until the date of each echocardiogram. LVEF was determined based on echocardiogram measurements. If EF was reported as a range, the midpoint of the range was chosen. Myopathy was defined by the presence of at least one abnormal muscle enzyme, MRI, electromyography, or muscle biopsy.[14] Organ-specific modified Medsger severity scores were dichotomized into absent–mild (score 0–1) versus moderate–severe (score ≥2).[14] Moderate–severe Raynaud included digital pits, ulceration, or gangrene. Moderate–severe gastrointestinal (GI) involvement included high-dose acid reflux medications, antibiotics for bacterial overgrowth, malabsorption syndrome, episodes of pseudo-obstruction, or total parental nutrition requirement. Moderate–severe kidney involvement included serum creatinine ≥1.7 mg/dL or at least 2+ urine protein. To investigate whether using a less inclusive definition of severity scores affects the coefficient estimate, we repeated analyses using any Raynaud phenomenon, kidney, and GI involvement, defined by absent (score 0) versus mild-to-severe (score ≥1). A restrictive ventilatory defect suggestive of ILD was defined by a forced vital capacity (FVC) ≤70% predicted, and echocardiographic evidence of pulmonary hypertension (PH) was defined as right ventricular systolic pressure (RVSP) ≥45 mm Hg.[15,16] Comorbidities were ascertained at routine clinical visits using prespecified criteria. Peripheral artery disease was defined as an abnormal ankle-brachial index (<0.9) or clinically based on claudication, lower extremity ulcers, or amputation. Suspicion of CAD was defined as angina, an abnormal exercise or pharmacologic stress test, abnormal coronary angiogram, and/or history of revascularization.[17] Hypertension was defined as a systolic blood pressure >140 mm Hg or diastolic blood pressure >90 mm Hg in two or more clinical visits, a diagnosis of hypertension by a primary care doctor, or antihypertensive use (not for Raynaud).[18]

Autoantibody status was obtained through routine clinical care in various internal and external laboratories and through research Euroimmun and enzyme-linked immunosorbent assay (ELISA) when sera were available. The autoantibody components of the Euroimmun panel included the following: centromere A and B (anticentromere antibodies [ACA]), topoisomerase 1 (Scl70), RP11 and RP155 (RNA polymerase III [RNAP III]), fibrillarin (U3RNP), Nor90, Th/To, PM75 and PM100 (PM/Scl), Ku, PDGFR, and Ro52. Positive results were defined following the manufacturer's instructions (EUROLINE Systemic Sclerosis Profile, Euroimmun), with the additional requirement that patients be positive for both subunits of multicomponent complexes (ACA, RNAP III, PM/Scl).[19] Anti-U1 RNP antibodies were assayed using an ELISA (Inova Diagnostics), and a threshold of >20 units was used to

define positivity. Patients were considered positive if ever positive by clinical or research autoantibody assays.

Statistical analysis. *Modeling EF transitions.* We chose to build statistical models that estimated the association of predictors and EF transitioning from one state to another. An alternative is to model an individual's smooth EF trajectory as a function of predictors and study the relationship between predictors and the estimated latent EF trajectory. However, unlike FVC or diffusing capacity for carbon monoxide (DLCO), EF tends to remain stable but can change abruptly from one visit to the next with a high degree of fluctuation. To investigate factors that are associated with the sudden drops in EF and cardiac recovery without the risk of them being considered random measurement errors, we built models using interval-level data to identify variables at the current visit that are most likely to predict patients' EF state at the next visit.

Our method relied on first-order Markov assumption, that is, current EF state depends only on immediate previous EF state and not on full history of the past EF states. Given the previous EF state, the EF pairs from the same or different patient were assumed to be independent intervals. This approach allowed us to build simpler and computationally efficient models that could easily generate future risk estimates given the current EF state.

For modeling, all time-varying variables were assigned to EF intervals defined by consecutive EF observations (from the time of previous EF up until the current EF). These time bins were designed to capture a patient's most recent health state before the documented change in EF. Time-varying variables, including disease duration at each echocardiogram, modified Rodnan skin score (mRSS), ILD (FVC ≤70% predicted), DLCO <60% predicted, PH (RVSP ≥ 45 mm Hg), myopathy, organ-specific severity scores, and comorbid conditions, were captured longitudinally. For all longitudinal variables, the observation closest to the current EF in time was used. Time-independent variables included sex, race, age at onset, diffuse versus limited skin involvement, and autoantibodies. We excluded EF intervals in which the time between two EF observations was greater than five years, which represented 2.2% of all echocardiogram pairs.

Missing predictors were imputed by multiple imputations using Factorial Analysis of Mixed Data (R package "missMDA").[20] We performed sensitivity analysis by comparing results from regression models using other imputation methods including random forest (R package "missForest") and multivariate imputation by chained equations (MICE; R package "mice").[21,22] A separate sensitivity analysis was performed only including patients who met 1980 or 2013 ACR criteria.

Regression models. Multivariable logistic regression models were used to identify risk factors for the development of LVSD after SSc onset. The first model estimated the probability of EF declining to <50% from a previous EF ≥50%. A cutoff of <50% was chosen as this corresponds to the definition of mildly reduced EF.[23] The second model focused on severe LVSD and estimated the probability of EF ≤35% from previous EF >35%. This value was chosen as it corresponds to when implantable cardioverter-defibrillators are recommended.[23] To evaluate the models' performance in predicting LVSD, we calculated the cross-validated area under the receiver operating characteristic curve (CV-AUC) from five-fold cross-validation. We recognize that subtle changes in LVEF may be secondary to noise in the measurement rather than reflecting a true clinical change. In a non-SSc population, the 95% confidence interval (CI) for a single EF measurement was ±2 percentage points for EF <53% and >60% and ±2.5 percentage points for EF 53% to 60%.[24] Based on these results, we performed a sensitivity analysis to investigate the robustness of the results after removing echo transitions that can potentially be created by measurement errors and requiring changes in LVEF be at least 5%.

Two additional multivariable logistic regression models were built to study the factors associated with cardiac recovery, as defined by improvement in LVEF. The probability of transitioning to EF >35% from previous EF ≤35% was modeled in the first model, and the probability of transitioning to EF ≥50 from previous EF <50% was modeled in the second model.[23] As the number of recoveries was anticipated to be small, a set of candidate predictors was selected for each outcome using a data-driven approach. We used random forest classification algorithm to identify the variables that are useful in classifying recovery status using multiple bootstrap data samples.[25] Each predictor was ranked based on mean decrease accuracy and mean decrease Gini score. The 10 variables with the highest sum of rankings based on the two metrics were chosen as candidate variables.

We also manually reviewed the medical records of patients with EF<=35% (n = 138) to better characterize the cause(s) of the severe LVSD. All statistical analyses were performed using R Statistical Software (v4.3.2; R Core Team 2021).

RESULTS

There were 2,303 patients representative of 13,209 echocardiograms included in this analysis. Most patients were female (82.6%) and White (76.2%). A substantial number had a smoking history (n = 984 [44.5%]). The mean age at SSc onset, at first non-Raynaud symptom, and at first echocardiogram were 42.8, 46.1, and 51.9 years. Patients had a mean of 5.7 echocardiograms performed. Additional clinical characteristics are provided in Table 1.

Table 2 describes all EF transitions from a previous to the current echocardiogram and the number of patients who experienced transitions. Of the 10,086 instances in which the previous echocardiogram demonstrated a normal EF (≥50%), 9,806 (97%) remained normal, whereas 280 (3%) transitioned to EF <50% on the subsequent echocardiogram. Of the 10,429 instances with previous EF >35%, 127 (1%) transitioned to EF ≤35%. Of the 236 instances of previous EF ≤35%, 100 (42%)

Table 1. Characteristics of patients (N = 2,303) with at least two echocardiograms*

Clinical feature	Mean (SD) or n (%)[a]
Age at systemic sclerosis onset, y[b]	42.8 (14.8), n = 2,299
Age at first non-Raynaud symptom, y	46.1 (14.0), n = 2,273
Follow-up duration, y[c]	9.1 (10.9), n = 2,299
Age at baseline echocardiogram, y	51.9 (13.3)
Number of echocardiograms per patient	5.7 (4.1)
Years from first to last echocardiogram	6.2 (5.2)
Female sex	1908 (82.8)
Self-identified race	
Black/African American	397 (17.2)
White	1,756 (76.2)
Other	150 (6.6)
Baseline modified Rodnan skin thickness score	9.5 (10.9), n = 2,182
Skin type	
Limited	1,298 (56.4)
Diffuse	906 (39.3)
Baseline FVC % predicted	84.5 (20.0), n = 2,210
Baseline FVC <70% predicted	511 (23.1), n = 2,210
Baseline DLco % predicted	70.6 (24.0), n = 2,210
Baseline DLco <60% predicted	713 (32.3), n = 2,210
Baseline RVSP ≥45 mm Hg	261 (11.3)[d]
Baseline moderate–severe Raynaud phenomenon[e]	801 (36.2), n = 2,212
Baseline moderate–severe gastrointestinal involvement[d]	400 (18.1), n = 2,212
Baseline moderate–severe kidney disease[f]	73 (3.3), n = 2,215
Baseline Medsger cardiac severity score[g]	200 (9.0), n = 2,215
Myopathy[h]	405 (18.3), n = 2,211
Autoantibodies	
Anticentromere	584 (26.4), n = 2,213
Anti-Scl70	446 (22.1), n = 2,022
Anti–RNAP III	339 (17.5), n = 1933
Anti-PM/Scl	38 (2.0), n = 1,861
Anti-Th/To	164 (8.8), n = 1,861
Anti-Nor90	81 (4.4), n = 1,861
Anti-Ku	80 (4.3), n = 1,861
Anti-Fibrillarin	125 (6.8), n = 1,861
Anti-Ro52	527 (28.3), n = 1,861
Anti-U1 RNP	228 (12.5), n = 1,823
Comorbidities (ever present)	
Cancer	458 (19.9)
Ever smoker	1,010 (44.2), n = 2,284
Coronary artery disease	154 (10.5), n = 1,472
Peripheral arterial disease	74 (5.1), n = 1,443
Diabetes mellitus	137 (7.9), n = 1,735
Hypertension	699 (46.8), n = 1,494
Immunosuppressive and cardiac medication exposures (ever use)	
Hydroxychloroquine	438 (21.5), n = 2,035
Methotrexate	368 (18.1), n = 2,035
Mycophenolate mofetil	568 (27.9), n = 2,035
Azathioprine	116 (5.7), n = 2,035
Cyclophosphamide	199 (9.8), n = 2,035
IVIG	116 (5.7), n = 2,035
Rituximab	19 (0.9), n = 2,035
Calcium channel blocker	1,360 (66.8), n = 2,035
Beta blocker	393 (19.3), n = 2,035
Nitrate	178 (8.7), n = 2,035
ACE-inhibitor/angiotensin receptor blockers	854 (42.0), n = 2,035
Aspirin	1,001 (49.2), n = 2,035
Diuretic	750 (36.9), n = 2,035
Statin	650 (31.9), n = 2,035
Antiplatelet	105 (5.2), n = 2,035

Table 1. (*Cont'd*)

* ACE, angiotensin-converting enzyme; DLco, diffusing capacity for carbon monoxide; EKG, electrocardiogram; FVC, forced vital capacity; GERD, gastroesophageal reflux disease; IVIG, intravenous immunoglobulin; MRI, magnetic resonance imaging; RNAP III, RNA polymerase III; RVSP, right ventricular systolic pressure.
[a] N = 2,303 unless otherwise noted.
[b] Date of systemic sclerosis onset is defined as the earlier date of Raynaud and non-Raynaud symptom onset.
[c] Follow-up duration defined as onset to first echocardiogram.
[d] Baseline moderate–severe gastrointestinal involvement defined as high dose GERD medications, antibiotics for bacterial overgrowth, malabsorption syndrome, episodes of pseudo-obstruction, or total parental nutrition requirement.
[e] Baseline moderate–severe Raynaud disease defined as Raynaud with digital pits, ulceration, or gangrene.
[f] Baseline moderate–severe kidney disease defined as serum creatinine ≥1 · 6 mg/dL, 3–4+ proteinuria, and/or history of scleroderma renal crisis.
[g] Baseline Medsger cardiac severity score is defined as clinically significant EKG conduction defect, sustained clinically important arrhythmia, or arrhythmia requiring treatment.
[h] Myopathy was defined based on the presence of one of the following: abnormal creatine kinase and/or aldolase levels, MRI findings, electromyography findings, and muscle biopsy findings.

transitioned to EF >35%, and of the 579 instances of previous EF <50%, 235 (41%) transitioned to EF ≥50%.

Among the 2,303 patients who had more than two echocardiograms separated by less than five years, 248 patients (11%) developed LVSD (EF <50%), and 114 (5%) developed severe LVSD (EF ≤35%). Most of these patients experienced cardiac recovery: 208 (66%) patients experienced recovery from LVSD, and 91 (82%) experienced improvement from severe LVSD. These percentages are higher than the probabilities of transitioning into recovery or improvement, indicating that patients take multiple visits to finally recover from LVSD or improve from severe LVSD. The median durations from SSc onset to first LVSD and first severe LVSD were 3.7 (interquartile range [IQR] 2–8.2) and 3.1 (IQR 1–7.2) years. The median durations for first recovery from LVSD and severe LVSD were 0.7 (IQR 0.2–1.5) and 0.4 (IQR 0.1–1.1) years.

Most predictors used in the regression models had missingness less than 15%, but higher missingness was observed for less frequently captured comorbid conditions, including diabetes mellitus (15.1%), CAD (24.6%), hypertension (32.1%), and AFib (31.0%) (see Supplemental Table 1). The missing values were imputed using multiple imputation of mixed data, which we used to generate all results from the analyses in this manuscript. We repeated all analyses using data imputed using random forest and MICE and verified that there were no meaningful differences in the estimated odds ratios (ORs; results not shown).

Demographic, disease-specific, and cardiac risk factors associated with increased risk of LVSD. The first multivariable logistic regression modeling the probability of EF

Table 2. Total EF transitions and numbers of patients who experienced EF transitions*

Total EF transitions					
		Current EF			
		EF ≥ 50%	35% < EF < 50%	EF ≤ 35%	Total
Previous EF	EF ≥ 50%	9806	202	78	10,086
	35% < EF < 50%	182	112	49	343
	EF ≤ 35%	53	47	136	236
	Total	10,041	361	263	10,665
Numbers of patients who experienced EF transitions					
		Current EF			
		EF ≥ 50%	35% < EF < 50%	EF ≤ 35%	Any
Previous EF	EF ≥ 50%	2189	175	73	2230
	35% < EF < 50%	158	63	41	203
	EF ≤ 35%	50	41	57	111
	Any	2226	214	121	

* Numbers of EFs transitioning from one of the three previous state categories (EF ≥ 50%, 35% < EF < 50%, or EF ≤ 35%) to current state categories (EF ≥ 50%, 35% < EF < 50%, or EF ≤ 35%) are shown in the top half of the table. Numbers of unique patients who experienced EF transitioning from one of the three previous state categories to current state categories are shown in the bottom half of the table. The last row (any for previous EF) indicates numbers of unique patients who transitioned into three current EF states regardless of their previous EF values, and the last column (any for current EF) shows numbers of unique patients whose previous EF belonged to each of the three categories regardless of current EF value. A patient can contribute to more than one cell in this table. Echocardiogram transitions observed between two echocardiograms separated by more than five years are excluded from calculations, and only patients who have more than two echocardiograms are included. EF, ejection fraction.

≥50% dropping to EF <50% demonstrated male sex (OR = 2.08 [95% CI 1.57–2.74], $P < 0.001$), Black race (OR = 1.79 [95% CI 1.32–2.43], $P < 0.001$), diffuse cutaneous type (OR = 1.49 [95% CI 1.10–2.01], $P = 0.010$), higher mRSS (OR = 1.02 [95% CI 1.00–1.03], $P = 0.039$), RVSP ≥45 mm Hg (OR = 1.64 [95% CI 1.25–2.15], $P < 0.001$), moderate–severe kidney disease (OR = 2.00 [95% CI 1.45–2.77], $P < 0.001$), and AFib (OR = 2.39 [95% CI 1.38–4.15], $P = 0.002$) were associated with higher odds of LVSD adjusting for other variables (Table 3). Anticentromere (OR = 0.49 [95% CI 0.32–0.75], $P = 0.001$) and anti-Scl70 (OR = 0.67 [95% CI 0.48–0.93], $P = 0.018$) were associated with lower odds. The model's CV-AUC was 0.70.

The second multivariate logistic regression, which modeled the probability of EF ≤35% from EF >35%, found male sex (OR = 1.73 [95% CI 1.15–2.61], $P = 0.009$), higher mRSS (OR = 1.03 [95% CI 1.01–1.05], $P = 0.005$), RVSP ≥45 mm Hg (OR = 1.74 [95% CI 1.17–2.60], $P = 0.006$), myopathy (OR = 2.26 [95% CI 1.36–3.76], $P = 0.002$), moderate–severe kidney disease (OR = 2.54 [95% CI 1.62–4.00], $P < 0.001$), AFib (OR = 2.41 [95% CI 1.13–5.13], $P = 0.023$), and anti-Ku (OR = 2.38 [95% CI 1.19–4.78], $P = 0.015$) were associated with higher odds of severe LVSD adjusting for other variables (Table 3). The model's CV-AUC was 0.70.

Given that previous research using the large EULAR Scleroderma Trial and Research group (EUSTAR) database found an association between digital ulcers and LVEF <55%, we performed a sensitivity analysis to examine whether digital ulcers and/or gangrene were associated with LVSD in our patients.[5] We did not find a statistically significant association between digital ulcers and/or gangrene and LVSD (OR = 1.29 [95% CI 0.91–1.83], $P = 0.145$) or severe LVSD (OR = 1.30 [95% CI 0.81–2.11], $P = 0.280$). The findings did not change meaningfully when a less inclusive definition of severity scores was used for kidney and GI involvement (result not shown). Our findings remained similar when restricting patient population to those who met 1980 or 2013 ACR criteria, requiring changes in LVEF be at least 5%, and adjusting for hydroxychloroquine exposure, respectively (see Supplement).

Unique factors may influence the likelihood of cardiac recovery. As shown in Table 2, a subset of patients may have had cardiac recovery or improvement. To investigate factors associated with recovery, we fit two multivariate logistic regression models. For each model, the 10 variables that were selected as the most important variables from a random forest classification model were used as predictors (Table 4). Tendon friction rubs (OR = 0.38 [95% CI 0.22–0.66], $P = 0.001$) were associated with lower odds of recovering to EF ≥50%, and anti–RNAP III (OR = 1.76 [95% CI 1.10–2.80], $P = 0.018$) was associated with higher odds. Diabetes history (OR = 0.18 [95% CI 0.04–0.83], $P = 0.028$) was associated with lower odds of improving to EF >35%. Findings were similar when restricting patients to those who meet 1980 or 2013 ACR criteria (see Supplement).

Common scleroderma-related causes of severe LVSD include arrhythmia, myocarditis, fibrosis, PH, and renal crisis. We reviewed the medical records of the 138 patients whose EF was ever ≤35% to further characterize the etiology of the severe LVSD as determined by the treating rheumatologist or cardiologist. These patients consisted of 113 patients whose EF dropped from >35% to ≤35% as well as

Table 3. Multiple logistic regression analyses examining risk factors for incident LV systolic dysfunction*

Variable	EF ≥ 50% to EF < 50%		EF > 35% to EF ≤ 35%	
	Adjusted OR (95% CI)	*P* value	Adjusted OR (95% CI)	*P* value
Disease duration, y	1.01 (0.99, 1.03)	0.512	1.00 (0.97, 1.02)	0.754
Male vs female sex	**2.08 (1.57, 2.74)**	**<0.001**	**1.73 (1.15, 2.61)**	**0.009**
Black/African American race vs all others	**1.79 (1.32, 2.43)**	**<0.001**	1.53 (0.97, 2.40)	0.065
Age at onset, y	1.01 (1.00, 1.02)	0.061	1.01 (0.99, 1.02)	0.333
Diffuse vs limited	**1.49 (1.10, 2.01)**	**0.010**	1.42 (0.90, 2.24)	0.133
mRSS	**1.02 (1.00, 1.03)**	**0.039**	**1.03 (1.01, 1.05)**	**0.005**
ILD	1.26 (0.94, 1.70)	0.122	1.22 (0.79, 1.90)	0.362
DLco < 60% predicted	1.14 (0.85, 1.52)	0.387	1.34 (0.87, 2.07)	0.187
RVSP ≥ 45 mm Hg	**1.64 (1.25, 2.15)**	**<0.001**	**1.74 (1.17, 2.60)**	**0.006**
Telangiectasia	0.93 (0.64, 1.34)	0.683	1.06 (0.61, 1.84)	0.830
Myopathy	1.45 (0.98, 2.14)	0.066	**2.26 (1.36, 3.76)**	**0.002**
Moderate–severe gastrointestinal involvement[a]	1.14 (0.87, 1.48)	0.348	0.96 (0.66, 1.42)	0.856
Moderate–severe Raynaud phenomenon[b]	1.04 (0.79, 1.36)	0.782	1.07 (0.73, 1.58)	0.731
Moderate–severe kidney disease[c]	**2.00 (1.45, 2.77)**	**<0.001**	**2.54 (1.62, 4.00)**	**<0.001**
Ever smoker	0.92 (0.55, 1.56)	0.757	1.33 (0.68, 2.61)	0.409
Diabetes mellitus	0.64 (0.31, 1.33)	0.229	0.54 (0.17, 1.75)	0.307
Coronary artery disease	0.81 (0.24, 2.74)	0.733	1.67 (0.46, 5.98)	0.433
Hypertension	0.93 (0.66, 1.32)	0.699	1.19 (0.73, 1.92)	0.490
Tendon friction rubs	0.89 (0.59, 1.35)	0.576	1.08 (0.61, 1.92)	0.792
Calcinosis	1.22 (0.88, 1.69)	0.225	1.16 (0.72, 1.88)	0.538
Atrial fibrillation	**2.39 (1.38, 4.15)**	**0.002**	**2.41 (1.13, 5.13)**	**0.023**
Anticentromere	**0.49 (0.32, 0.75)**	**0.001**	0.83 (0.46, 1.51)	0.550
Anti-Fibrillarin	0.99 (0.63, 1.55)	0.955	0.92 (0.46, 1.84)	0.822
Anti-Ku	1.33 (0.75, 2.37)	0.329	**2.38 (1.19, 4.78)**	**0.015**
Anti-Nor90	0.94 (0.50, 1.76)	0.841	1.29 (0.58, 2.87)	0.528
Anti-RNA polymerase III	0.82 (0.57, 1.17)	0.280	0.65 (0.37, 1.15)	0.143
Anti-Ro52	0.99 (0.74, 1.33)	0.947	1.03 (0.68, 1.58)	0.879
Anti-Scl70	**0.67 (0.48, 0.93)**	**0.018**	0.87 (0.54, 1.40)	0.558
Anti-Th/To	0.89 (0.54, 1.47)	0.660	1.01 (0.50, 2.07)	0.971
Anti-PM/Scl	1.05 (0.41, 2.69)	0.920	2.20 (0.73, 6.62)	0.160
Anti-U1 RNP	1.19 (0.84, 1.69)	0.330	1.59 (0.97, 2.60)	0.065

* Adjusted ORs, 95% CIs, and *P* values of each variable from two multiple logistic regressions are shown. The first regression models the probability of transitioning to EF <50% from EF ≥50%, and the second models the probability of transitioning to EF ≤35% from EF >35%. Statistically significant ORs using the threshold of *P* value <0.05 are shown in bold. CI, confidence interval; DLco, diffusing capacity for carbon monoxide; EF, ejection fraction; GERD, gastroesophageal reflux disease; ILD, interstitial lung disease; LV, left ventricular; mRSS, modified Rodnan skin score; OR, odds ratio; RVSP, right ventricular systolic pressure.
[a] Baseline moderate–severe gastrointestinal involvement defined as high dose GERD meds, antibiotics for bacterial overgrowth, malabsorption syndrome, episodes of pseudo-obstruction, or total parental nutrition requirement.
[b] Baseline moderate–severe Raynaud disease defined as Raynaud with digital pits, ulceration, or gangrene.
[c] Baseline moderate–severe kidney disease defined as serum creatinine ≥1 · 6 mg/dL, 3–4+ proteinuria, and/or history of scleroderma renal crisis.

25 whose first EF was ≤35%. The most common associations with EF ≤35% identified through medical record review were ischemic, conduction- or arrhythmia-related, fibrosis, myocarditis, right heart failure from PH, infection, and renal crisis (Table 5). Of the 25 patients with arrhythmia, 16 (68%) had an additional identified etiology for their severe LVSD such as ischemia (n = 2), infection (n = 3), myocarditis (n = 3), or biopsy-proven fibrosis (n = 4). Among those with arrhythmia as the only identifiable etiology (n = 9), five had full recovery with rate control or pacemakers, whereas four had ongoing severe LVSD. Of the 18 with myocarditis, 10 demonstrated some degree of improvement with treatment, mainly immunosuppression or immunomodulation with intravenous immunoglobulin (see Supplement). Of the 25 patients who had histologic evaluation of the myocardium, 19 (76%) had fibrosis. Six with biopsy-proven fibrosis underwent cardiac MRI. Four had delayed gadolinium enhancement indicative of fibrosis, and the remaining two had normal cardiac MRIs. Only 3 of the 19 patients with fibrosis on myocardial biopsy had full recovery; one developed severe LVSD post-partum in the setting of renal crisis; the second had concomitant myopathy and arrhythmia, which improved with amiodarone, a pacemaker, and immunosuppression; and the third was also thought to have cardiac Raynaud (wall motion abnormalities with normal cardiac catheterization). Of the eight with severe LVSD in the setting of renal crisis, six (75%) fully recovered and one improved (final EF 45%) with management of renal crisis. The patient who did not improve also had ischemic heart disease requiring a stent, PAH, and aortic stenosis. Figure 1 illustrates sample EF trajectories of

Table 4. Multiple logistic regression analyses examining factors that are associated with cardiac recovery*

Variable	EF < 50% to EF ≥ 50%		EF ≤ 35% to EF > 35%	
	Adjusted OR (95% CI)	*P* value	Adjusted OR (95% CI)	*P* value
Age at onset, y	1.01 (0.99, 1.02)	0.445	1.01 (0.99, 1.04)	0.305
Disease duration, y	0.99 (0.97, 1.02)	0.561	0.98 (0.95, 1.02)	0.299
DLco < 60	1.22 (0.84, 1.78)	0.285	1.39 (0.74, 2.61)	0.303
mRSS	1.01 (0.99, 1.03)	0.349	0.98 (0.96, 1.01)	0.248
Anti-RNA polymerase III	**1.76 (1.10, 2.80)**	**0.018**	1.71 (0.75, 3.89)	0.202
Moderate–severe kidney disease	1.48 (0.96, 2.29)	0.076	1.96 (0.96, 4.01)	0.066
Anti-Ro52	0.83 (0.55, 1.25)	0.368	1.32 (0.67, 2.59)	0.424
Diffuse	0.70 (0.45, 1.09)	0.117	–	–
Tendon friction rubs	**0.38 (0.22, 0.66)**	**0.001**	–	–
RVSP ≥ 45	0.76 (0.52, 1.10)	0.144	–	–
ILD	–	–	1.14 (0.59, 2.21)	0.691
Diabetes mellitus	–	–	**0.18 (0.04, 0.83)**	**0.028**
Raynaud phenomenon	–	–	1.32 (0.71, 2.25)	0.420

* Adjusted ORs, 95% CIs, and *P* values of two sets of 10 selected variables estimated by two multiple logistic regressions are shown. The first regression models the probability of recovering to EF ≥50% from EF <50%, and the second models the probability of improving to EF >35% from EF ≤35%. For each regression model, the 10 most important variables selected from a random forest classification model are used as predictors. Statistically significant ORs using the threshold of *P* value <0.05 are shown in bold. CI, confidence interval; DLco, diffusing capacity for carbon monoxide, EF, ejection fraction; ILD, interstitial lung disease; mRSS, modified Rodnan skin score; OR, odds ratio; RVSP, right ventricular systolic pressure.

selected patients for the five most common scleroderma-associated causes of EF ≤35%.

DISCUSSION

In this large prospective cohort study of SSc adults, we identified incident LVSD (EF < 50%) in 247 patients (11%) and severe LVSD (EF ≤ 35%) in 114 (5%). We investigated baseline and time-varying factors that were associated with incident LVSD and found that male sex, higher mRSS, PH (defined by RVSP ≥45 mm Hg on echocardiogram), moderate–severe kidney disease, and AFib were associated with increased odds of developing both LVSD and severe LVSD (EF >35% but <50% and EF ≤35%, respectively). Black race, older age at SSc onset, and diffuse cutaneous disease increased the odds, whereas anticentromere and anti-Scl70 decreased the odds of LVSD but not severe LVSD. Myopathy and anti-Ku were uniquely associated with a higher risk of severe LVSD. Anti–RNAP III was associated with cardiac recovery for previous EF <50%. To our knowledge, this study provides the most detailed analysis of scleroderma-specific and cardiac risk factors associated with varying degrees of incident LVSD in SSc using time-varying predictors and is the only study to explore factors associated with cardiac recovery.

To better understand the factors associated with severe LVSD, we reviewed the medical records of 138 patients who ever had EF ≤35%. The most common nonscleroderma-related drivers of severe LVSD included ischemic CAD and infection. The ischemic CAD may at least partially account for the association between male sex and LVSD identified in multivariate logistic regression. Several scleroderma-related complications were identified as drivers of severe LVSD including fibrosis, myocarditis, PH, and renal crisis. The PH association identified in the multivariate logistic regression was particularly important to confirm due to the use of RVSP >45 mm Hg as a surrogate marker because RVSP can be elevated secondary to left heart disease and LV dysfunction. Arrhythmias were also identified as a potential risk factor for LVSD. Through medical record review, we identified decompensated PH (excluding PH from left heart failure) as one of the top five scleroderma-related contributors to severe LVSD in our patients. In these patients, decompensated, severe PH with right-sided heart failure may have caused low EF through ventricular interdependence.[26]

Medical record review also suggests a possible reason for the association between anti–RNAP III and cardiac recovery. Patients with renal crisis, who are more likely to be anti–RNAP III positive, may develop acute LVSD in a high-afterload state, and these patients demonstrated high rates of eventual cardiac recovery or improvement (7 of 8, 88%) with successful blood pressure control. The one patient with renal crisis whose EF did not improve also had concomitant ischemic LVSD, PAH, and aortic stenosis. The high rate of recovery or improvement among patients with renal crisis has important management implications, suggesting these patients may not require an implantable defibrillator.

Although no study has explored LVSD in SSc using the same methods and definition, our findings are consistent with published studies. A study of 1,095 patients with SSc found an association between myopathy and clinically evident congestive heart failure.[4] We, too, found a significant association between myopathy and severe LVSD. Moreover, a multiple regression analysis of data from the EUSTAR database identified age, male sex, diffuse cutaneous disease, disease duration, digital ulcers, and renal and muscle involvement to be associated with LVSD, defined as EF

Table 5. Causes of severe LV systolic dysfunction in patients with SSc based on chart review*

Category[a]	Total (N = 138), n (%)	No improvement,[b] n (%)	Improvement,[b] n (%)	Full recovery,[b] n (%)
Ischemic	25 (18)	15 (60)	4 (16)	6 (24)
Arrhythmia/conduction	25 (18)	11 (44)	5 (20)	9 (36)
Fibrosis on endomyocardial biopsy[c]	19 (14)	13 (68)	3 (16)	3 (16)
Myocarditis[d]	18 (13)	8 (44)	3 (17)	7 (39)
PHTN[e]	14 (10)	6 (43)	3 (21)	5 (36)
Infection	13 (9)	6 (46)	1 (8)	6 (46)
Renal crisis	8 (6)	1 (13)	1 (13)	6 (75)
Valvular disease	6 (4)	3 (50)	0 (0)	3 (50)
Hypertension	5 (4)	3 (60)	0 (0)	2 (40)
Bleed/anemia[f]	5 (4)	3 (60)	0 (0)	2 (40)
Medication induced	5 (4)	4 (80)	0 (0)	1 (20)
Postpartum	4 (3)	2 (50)	0 (0)	2 (50)
Pulmonary embolism	3 (2)	2 (67)	0 (0)	1 (33)
Renal disease (non-SSc)	3 (2)	1 (33)	0 (0)	2 (67)
Postoperative[g]	2 (1)	0 (0)	0 (0)	2 (100)
Microvascular disease[h]	2 (1)	1 (50)	0 (0)	1 (50)
Familial	2 (1)	0 (0)	1 (50)	1 (50)
Stress cardiomyopathy	2 (1)	0 (0)	0 (0)	2 (100)
Alcohol	1 (<1)	1 (100)	0 (0)	0 (0)
Myasthenia gravis flare	1 (<1)	1 (100)	0 (0)	0 (0)
Unknown	24 (17)	10 (42)	3 (13)	11 (46)

* CK, creatine kinase; CKMB, CK myocardial band; EF, ejection fraction; GI, gastrointestinal; LVSD, left ventricular systolic dysfunction; MRI, magnetic resonance imaging; PHTN, pulmonary hypertension; SSc, systemic sclerosis; TNF, tumor necrosis factor.
[a] Patients with multiple etiologies for severe LVSD were included in each relevant category.
[b] No improvement is defined as a final EF of 35% or below; partial improvement is defined as a final EF of 36%–49%; full recovery is defined as EF ≥50%.
[c] One patient with biopsy-proven fibrosis was on a TNF-inhibitor at the time of developing LVSD.
[d] Myocarditis defined by elevated CK with high CKMB, EF fluctuation with peripheral myositis, and patchy delayed enhancement on MRI (2 patients).
[e] Excluding PHTN secondary to left heart failure.
[f] In one patient, the bleed was both GI and massive hemoptysis; three patients had GI bleed, and one patient had significant anemia requiring transfusion of unclear etiology (upper endoscopy and colonoscopy were normal).
[g] Aortic valve replacement (1 patient), lung cancer resection (1 patient).
[h] Microvascular disease was suspected due to an abnormal stress test or significant regional wall motion abnormalities on echocardiogram with a subsequent normal left heart catheterization.

<55%, in SSc.[5] Despite using more stringent EF thresholds that excluded patients with low–normal cardiac function, we also identified onset age, male sex, diffuse cutaneous disease, renal disease, and muscle involvement as SSc-specific risk factors associated with increased odds of LVSD. Severe Raynaud phenomenon, defined by digital ulceration or gangrene, was not statistically significantly associated with LVSD in our study. We also identified an increased risk of LVSD in patients with SSc with self-identified Black race. Although Black patients with SSc have more severe disease than self-identified White patients, it is possible that non-SSc factors, such as a higher prevalence of cardiovascular disease or social determinants of health, may contribute to this increased risk.[27] We identified an association between anti-Ku and increased odds of severe LVSD. The literature regarding the association between anti-Ku and cardiac involvement in SSc is mixed. One study of 60 patients with SSc, 6 of whom were anti-Ku positive, found that patients with anti-Ku had significantly higher odds of cardiac disease, defined as EF <45%, pericarditis, arrhythmia requiring treatment, or conduction defect (OR 12.50 [95% CI 1.88–83.3]).[28] In particular, those with scleroderma and myositis who are anti-Ku positive may have a higher risk of severe arrhythmias.[29] Two other studies, with similar sample sizes and definition of cardiac involvement, found no association between cardiac involvement and anti-Ku in patients with SSc.[30,31] All three of these studies had small sample sizes and used broad definitions of cardiac involvement, encompassing evidence of diastolic dysfunction, arrhythmia, and pericarditis. Future research should take a more nuanced approach, exploring the association between anti-Ku and specific types of cardiac involvement in SSc.

Our study has several strengths, including the use of a large, well-phenotyped, prospective SSc cohort with long-term follow-up. Our modeling approach, using time-varying variables, was designed to identify dynamic predictors associated with incident LVSD and cardiac recovery. Instead of modeling progression or recovery regardless of patients' past cardiac state, we modeled the transition to LVSD from a past normal state or transition to recovery from past LVSD. This approach, while facilitating our understanding of population-level risk factors for LVSD, lays the foundation for future work to develop an individualized, dynamic

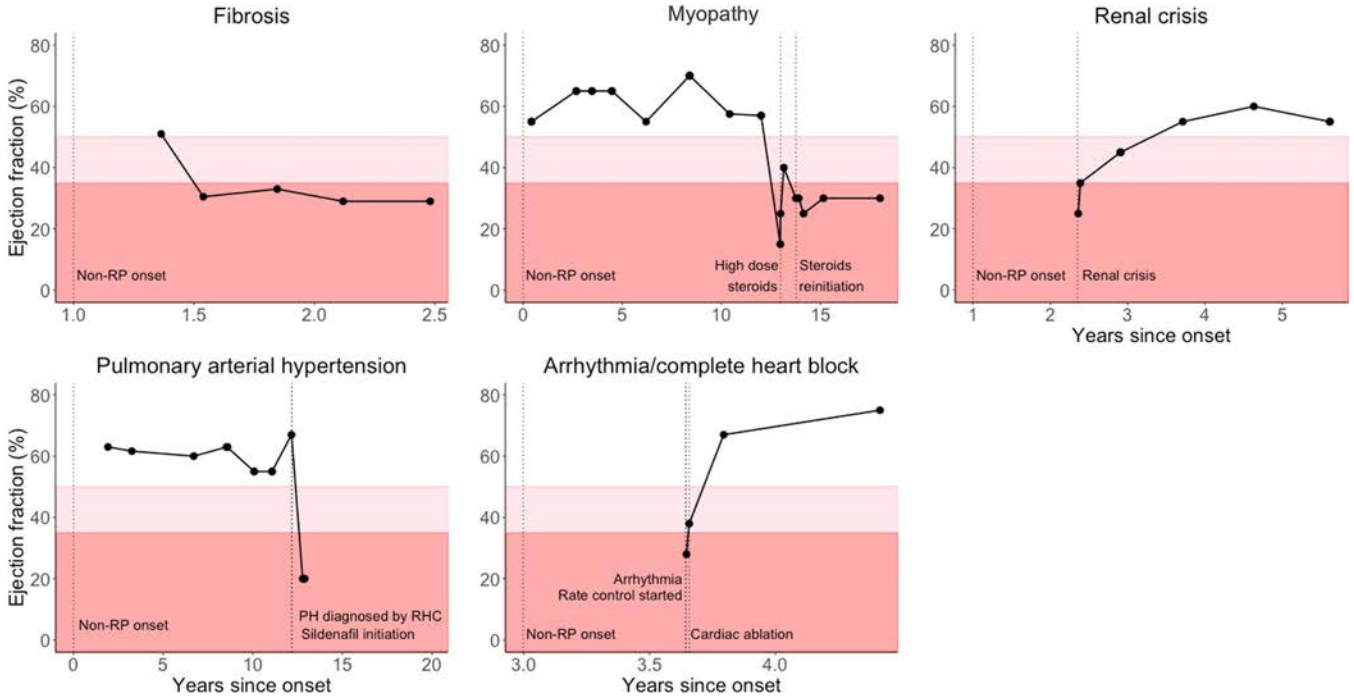


Figure 1. Sample ejection fraction trajectories of patients with ejection fraction ≤35% due to cardiac fibrosis (top left), myopathy (top center), renal crisis (top right), pulmonary arterial hypertension (bottom left), and arrhythmia/complete heart block (bottom center). RP, Raynaud phenomenon.

cardiac risk stratification tool for patients with SSc. Rather than relying on baseline variables, which might not accurately reflect a patient's current health status, such dynamic risk calculations could provide risk estimates at any time during a patient's disease course. Additional work is required to improve model performance, including adding relevant laboratory and echocardiographic biomarkers as predictors.

Fewer samples were available for our two recovery outcomes, and random forest models were used to select the most relevant variables to be included as predictors. By performing data-driven variable selections, we avoided relying heavily on a priori judgments about the importance of individual variables. This approach can be reproduced in other cohorts.

Our study has some limitations. Data on cardiac comorbidities were not systematically collected for the entire study period. We attempted to limit missingness by comprehensive medical record review for data before 2010. Further, there is inherent variability and subjectivity in assessment of LVEF by echocardiogram. We attempted to manage this "noise" by examining optimal cut points for reliable EFs that are clinically meaningful and relevant for rheumatologists receiving echocardiographic reports. Although we could have used a technique that smooths over the random measurement errors to obtain a smooth trajectory of EF, such a technique would limit our ability to directly investigate factors that are associated with sudden changes in EF as these changes could be interpreted as noise. To test the robustness of our findings, we performed a sensitivity analysis requiring changes in LVEF be at least 5%; our findings were qualitatively similar although some parameters were no longer statistically significant. We also did not have data on whether patients had clinical signs and symptoms of heart failure or whether LVSD was treated, such as with goal-directed medical therapy. Additionally, we likely underestimated the number of patients with cardiac fibrosis as most patients do not undergo cardiac biopsy. We also did not have systematic creatine phosphokinase, troponin, pro–brain natriuretic peptide, electrocardiogram, and cardiac MRI data captured in this study population. This limited our ability to fully assess for primary cardiac involvement attributable to SSc that may have affected the myocardium, vasculature, valves, pericardium, and conduction system, possibly due to microvascular dysfunction, inflammation, and fibrosis among other mechanisms. Given the long period under study and changes in practice patterns over time, right heart catheterization was not performed in all patients with echocardiographic evidence of PH to confirm the diagnosis with gold standard methods. Lastly, we were unable to investigate the role of an overlapping diagnosis of an idiopathic inflammatory myopathy, autoimmune thyroid disease, HIV status, and other factors that may contribute to LVSD.

In conclusion, timely intervention for SSc myocardial disease has been hindered by the lack of tools for predicting the development and course of LVSD. Our study demonstrates that LVSD is a common complication in SSc, with an incidence of 11%. We identified distinct risk factors, including unique immune response (anti-Ku), associated with the development of LVSD. With early

identification, patients at risk for LVSD may be candidates for close monitoring and treatment of arrhythmias or inflammatory myocarditis. Additionally, our study found the important prognostic value of anti–RNAP III regarding the likelihood of cardiac recovery once LVSD has developed.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding authors, Drs Shah and Kim confirm that all authors have provided the final approval of the version to be published and take responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. Hao Y, Hudson M, Baron M, et al; Canadian Scleroderma Research Group; Australian Scleroderma Interest Group. Early mortality in a multinational systemic sclerosis inception cohort. Arthritis Rheumatol 2017;69(5):1067–1077.
2. Steen VD, Medsger TA Jr. Severe organ involvement in systemic sclerosis with diffuse scleroderma. Arthritis Rheum 2000;43(11):2437–2444.
3. Bruni C, Ross L. Cardiac involvement in systemic sclerosis: getting to the heart of the matter. Best Pract Res Clin Rheumatol 2021;35(3):101668.
4. Follansbee WP, Zerbe TR, Medsger TA Jr. Cardiac and skeletal muscle disease in systemic sclerosis (scleroderma): a high risk association. Am Heart J 1993;125(1):194–203.
5. Allanore Y, Meune C, Vonk MC, et al; EUSTAR co-authors. Prevalence and factors associated with left ventricular dysfunction in the EULAR Scleroderma Trial and Research group (EUSTAR) database of patients with systemic sclerosis. Ann Rheum Dis 2010;69(1):218–221.
6. Tipparot T, Foocharoen C, Mahakkanukrauh A, et al. Clinical and laboratory predictions of myocardial inflammation as detected by cardiac magnetic resonance imaging in patients with systemic sclerosis: a pilot study. Int J Rheum Dis 2019;22(12):2125–2133.
7. Lin CY, Chen HA, Chang TW, et al. Association of systemic sclerosis with incident clinically evident heart failure. Arthritis Care Res (Hoboken) 2023;75(7):1452–1461.
8. Galderisi M, Cosyns B, Edvardsen T, et al; 2016–2018 EACVI Scientific Documents Committee; 2016–2018 EACVI Scientific Documents Committee. Standardization of adult transthoracic echocardiography reporting in agreement with recent chamber quantification, diastolic function, and heart valve disease recommendations: an expert consensus document of the European Association of Cardiovascular Imaging. Eur Heart J Cardiovasc Imaging 2017;18(12):1301–1310.
9. Fernández-Codina A, Simeón-Aznar CP, Pinal-Fernandez I, et al. Cardiac involvement in systemic sclerosis: differences between clinical subsets and influence on survival. Rheumatol Int 2017;37(1):75–84.
10. West SG, Killian PJ, Lawless OJ. Association of myositis and myocarditis in progressive systemic sclerosis. Arthritis Rheum 1981;24(5):662–668.
11. Masi AT. Preliminary criteria for the classification of systemic sclerosis (scleroderma). Subcommittee for scleroderma criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. Arthritis Rheum 1980;23(5):581–590.
12. van den Hoogen F, Khanna D, Fransen J, et al. 2013 classification criteria for systemic sclerosis: an American College of Rheumatology/European League against Rheumatism collaborative initiative. Arthritis Rheum 2013;65(11):2737–2747.
13. LeRoy EC, Black C, Fleischmajer R, et al. Scleroderma (systemic sclerosis): classification, subsets and pathogenesis. J Rheumatol 1988;15(2):202–205.
14. Medsger TA Jr, Bombardieri S, Czirjak L, et al. Assessment of disease severity and prognosis. Clin Exp Rheumatol 2003;21(3 suppl 29):S42–S46.
15. Shah AA, Wigley FM, Hummers LK. Telangiectases in scleroderma: a potential clinical marker of pulmonary arterial hypertension. J Rheumatol 2010;37(1):98–104.
16. Rudski LG, Lai WW, Afilalo J, et al. Guidelines for the echocardiographic assessment of the right heart in adults: a report from the American Society of Echocardiography endorsed by the European Association of Echocardiography, a registered branch of the European Society of Cardiology, and the Canadian Society of Echocardiography. J Am Soc Echocardiogr 2010;23(7):685–713.
17. Gulati M, Levy PD, Mukherjee D, et al. 2021 AHA/ACC/ASE/CHEST/SAEM/SCCT/SCMR guideline for the evaluation and diagnosis of chest pain: a report of the American College of Cardiology/American Heart Association Joint Committee on clinical practice guidelines. Circulation 2021;144(22):e368–e454.
18. Colantonio LD, Booth JN III, Bress AP, et al. 2017 ACC/AHA blood pressure treatment guideline recommendations and cardiovascular risk. J Am Coll Cardiol 2018;72(11):1187–1197.
19. Mecoli CA, Gutierrez-Alamillo L, Yang Q, et al. PM-Scl and Th/To in systemic sclerosis: a comparison of different autoantibody assays. Clin Rheumatol 2021;40(7):2763–2769.
20. Josse J, Husson F. missMDA: a package for handling missing values in multivariate data analysis. J Stat Softw 2016;70(1):1–31.
21. Stekhoven DJ, Bühlmann P. MissForest–non-parametric missing value imputation for mixed-type data. Bioinformatics 2012;28(1):112–118.
22. van Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. J Stat Softw 2011;45(3):1–67.
23. Heidenreich PA, Bozkurt B, Aguilar D, et al; ACC/AHA Joint Committee Members. 2022 AHA/ACC/HFSA guideline for the management of heart failure: a report of the American College of Cardiology/American Heart Association Joint Committee on clinical practice guidelines. Circulation 2022;145(18):e895–e1032.
24. Suwatanaviroj T, He W, Pituskin E, et al. What is the minimum change in left ventricular ejection fraction, which can be measured with contrast echocardiography? Echo Res Pract 2018;5(2):71–77.
25. Breiman L. Random Forests. Mach Learn 2001;45(1):5–32.
26. Naeije R, Badagliacca R. The overloaded right heart and ventricular interdependence. Cardiovasc Res 2017;113(12):1474–1485.
27. Kuchinad KE, Kim JS, Woods A, et al. Racial variability in immune responses only partially explains differential systemic sclerosis disease severity. Ann Rheum Dis 2024;83(11):1513–1521.
28. Rodriguez-Reyna TS, Hinojosa-Azaola A, Martinez-Reyes C, et al. Distinctive autoantibody profile in Mexican Mestizo systemic sclerosis patients. Autoimmunity 2011;44(7):576–584.
29. Bhalodia A, Bermea K, Schmidt J, et al. Increased risk of myocarditis and arrythmias in anti-Ku-positive scleroderma-myositis overlap patients: a case series. Rheumatology (Oxford) 2024;63(9):e268–e269.

30. Rozman B, Cucnik S, Sodin-Semrl S, et al. Prevalence and clinical associations of anti-Ku antibodies in patients with systemic sclerosis: a European EUSTAR-initiated multi-centre case-control study. Ann Rheum Dis 2008;67(9):1282–1286.

31. Kaji K, Fertig N, Medsger TA Jr, et al. Autoantibodies to RuvBL1 and RuvBL2: a novel systemic sclerosis-related antibody associated with diffuse cutaneous and skeletal muscle involvement. Arthritis Care Res (Hoboken) 2014;66(4):575–584.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis & Rheumatology
Vol. 78, No. 3, March 2026, pp 724–733
DOI 10.1002/art.43406

Longitudinal Changes in Serum Urate Levels From Premenopause Through Postmenopause: Interrupted Time-Series Analyses

Shreya Billa,[1] Sho Fukui,[1,2,3] Misti L. Paudel,[1] Takahiro Suzuki,[4] Ryosuke Imai,[5] Yuntae Kim,[6] Takehiro Nakai,[2] Hiromichi Tamaki,[2] Mitsumasa Kishimoto,[2,7] Hilde S. Ørbo,[1] Sara K. Tedeschi,[1] Hyon K. Choi,[8] Masato Okada,[2] and Daniel H. Solomon[1]

Objective. Hyperuricemia (HU) and gout are common in postmenopausal women. We aimed to identify the longitudinal changes in serum urate (SU) levels during and after the menopausal transition and its interaction with coexisting SU-modifying conditions.

Methods. This longitudinal study included Japanese women who underwent annual medical examinations from April 2004 to September 2024 and had at least one visit before and after self-reported menopause. Menopausal transition stages were categorized into premenopause, perimenopause (5 years before and up to the menopause), and postmenopause. Longitudinal changes in SU and HU (SU ≥6.8 mg/dL or taking medications for gout or HU) were examined by interrupted time-series analyses and evaluated across stratified subgroups.

Results. We analyzed 8,169 eligible participants with 93,511 visits over a median follow-up of 13.8 years. SU levels gradually increased during premenopause, rose sharply over perimenopause, and stabilized in postmenopause. Compared to premenopause, the mean SU level was 0.41 mg/dL (95% confidence interval: 0.38–0.43) higher in postmenopause. HU prevalence increased from <1.0% during premenopause to 4% to 5% during postmenopause. Compared to premenopause, the associations of a low estimated glomerular filtration rate (<60 mL/min per 1.73 m^2) and a high body mass index (≥25) with HU were greater in postmenopause; HU was observed in approximately 18% of overweight or obese women at menopause.

Conclusion. SU levels rapidly increase during perimenopause and are already elevated by the time of menopause. Maintaining a normal body weight and preserving kidney function before menopause may decrease postmenopausal HU and potentially prevent subsequent gout in women.

INTRODUCTION

Gout is the most common inflammatory arthritis caused by a deposition of monosodium urate crystals.[1,2] Gout has been increasing worldwide[1–5] and poses substantial health and financial burden,[5,6] necessitating strategies for prevention. Given that hyperuricemia (HU) is a strong determinant of developing gout,[7–9] a clear understanding of the longitudinal serum urate (SU) levels throughout the life course is essential for gout prevention.

Menopause has been considered an inflection point in gout risk for women. Gout is uncommon in young women,[3] but its incidence markedly increases in older postmenopausal women.[3,8,10–12] Although SU levels may increase with aging,[13] cross-sectional studies have shown an independent association between HU and menopause.[14,15] Estrogen is known to promote

Dr Solomon's work was supported by the NIH (grant AR-P30072577).

[1]Division of Rheumatology, Inflammation, and Immunity, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts; [2]Immuno-Rheumatology Center, St Luke's International Hospital, Tokyo, Japan; [3]Department of Emergency and General Medicine, Kyorin University School of Medicine, Tokyo, Japan; [4]Department of Cardiovascular Medicine, St Luke's International Hospital, Tokyo, Japan; [5]Department of Pulmonary Medicine, St Luke's International Hospital, Tokyo, Japan; [6]Department of Gastroenterology, St Luke's International Hospital, Tokyo, Japan; [7]Department of Nephrology and Rheumatology, Kyorin University School of Medicine, Tokyo, Japan; [8]Rheumatology Unit, Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston.

Ms Billa and Dr Fukui contributed equally to this work.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/art.43406).

Author disclosures and graphical abstract are available at https://onlinelibrary.wiley.com/doi/10.1002/art.43406.

Address correspondence via email to Sho Fukui, MD, MPH, at shofukui@ks.kyorin-u.ac.jp.

Submitted for publication April 24, 2025; accepted in revised form August 26, 2025.

renal excretion of SU, and the decreased SU excretion postmenopause may partially explain HU in this group of women.[16] Other conditions that are common during or after the menopausal transition, such as obesity and kidney dysfunction, may also contribute to postmenopausal HU and gout.[17–20]

However, the precise timing of increases in SU across the menopausal transition are not fully understood, limiting the ability to develop evidence-based prevention strategies. The trend in HU prevalence and specific time when HU becomes more common during a woman's midlife is not well-known. Previous studies have primarily used cross-sectional analyses, with one study of 7,662 women in the United States finding that menopause is independently associated with higher SU levels. Another study of 58,870 Korean women found that increased prevalence of HU occurred during the late menopausal transition stage.[14,15,21] However, as these studies did not follow the same patients, it is difficult to understand how aging, menopause, and other relevant clinical risk factors impact SU over time, creating the need for more longitudinal studies.

We studied a longitudinal dataset from a large sample of women in Japan to examine the temporal changes in SU levels and the prevalence of HU throughout the menopausal transition and the postmenopause stages. Furthermore, we sought to elucidate how changes in SU levels are related to other contributing factors, such as body mass index (BMI) and kidney function, over the midlife of women.

METHODS

Study design, setting, and participants. This longitudinal study is conducted as part of the St. Luke's Initiative on Menopause and Women's Health Study, which investigates the impact of menopause on various health outcomes using longitudinal data from annual health checkups in Japan. This study used St. Luke's Health Checkup Database (SLHCD) from April 2004 to September 2024. SLHCD is a database of a large annual medical examination program at St. Luke's International Hospital, a tertiary care center in Tokyo, Japan. The details have been described previously.[22–27] The annual health checkup is a health-promotion program that serves all Japanese citizens. All full-time employees are mandated to undergo annual medical examinations, and other people can voluntarily participate in the examination. Because the program takes place in central Tokyo, the study cohort primarily consists of local residents and people working in the area, including hospital staff from the institution. SLHCD systematically collected data about participant demographics, medical histories, patient questionnaires, and examination results. This set of analyses included female participants who reported the occurrence of menopause during the follow-up; participants were required to have at least one visit both before and after the menopause. We excluded patients whose menopause timing was not determined and those who refused to use their data for research, received hormonal therapy, and underwent surgical procedures that may relate to surgical menopause (ie, total or extensive hysterectomy and ovarian resection).

This study was conducted following the Declaration of Helsinki and relevant ethical guidelines for medical research in Japan. Informed consent was obtained by offering opt-out options to the participants. This study was approved by the St. Luke's International University Institutional Review Board (24-R124). The study followed the Strengthening the Reporting of Observational Studies in Epidemiology reporting guideline.[28]

Exposure of interest and outcomes. The exposure of interest was menopause. At every visit of the annual medical examination, female participants were asked whether they underwent menopause or not in the questionnaire. We assumed that the patient had undergone menopause when the answer changed from "no" to "yes" for the first time. If a patient answered "yes" at one visit but later answered "no" at a subsequent visit, the timing of menopause was considered undetermined. For statistical analyses, the median date between the two consecutive visits before and after menopause occurrence was considered the date of menopause.

The primary outcome of this study was SU level (mg/dL) and the secondary outcome was HU (SU ≥6.8 mg/dL or taking medications for HU/gout). In the annual medical examination, SU was routinely measured at every visit by the uricase oxidization method (Clinical Chemistry Analyzers, JCA-BM2250/6070, JEOL Ltd). In SLHCD, information on gout and HU was collected as a single category, and thus gout could not be distinguished from HU. In Japan, asymptomatic HU is routinely monitored in many annual health checkup programs and occasionally treated in certain high-risk individuals, such as those at risk for gout or comorbid conditions, in accordance with national guidelines.[29]

Covariates. We collected both time-fixed and time-varying covariates for this study. The time-fixed covariates were age at menopause. The time-varying covariates included BMI, estimated glomerular filtration rate (eGFR) based on a Japanese-specific formula,[30] smoking status (current, previous, or never), and alcohol intake (drink per day). One drink of alcohol was defined as a drink that included 10 g of ethanol, and the ethanol amount was calculated as previously described.[22,23,27] Given medications that can affect SU levels,[31,32] we extracted time-varying medication use status for hypertension, diabetes, dyslipidemia, cerebrovascular disease, ischemic heart disease, kidney dysfunction, and mycobacterium infections (yes or no). The database only contains information on treatment status for each condition, such as whether the patient is "receiving pharmacological treatment"; it does not include data on the details of medications. The results of the dietary frequency questionnaire became available in October 2012 and included information on the intake of carbohydrates, meat and eggs, seafood, soy, milk and dairy products,

vegetables, fruits, and sweets. In the sensitivity analyses, the lifestyle questionnaire results, including dietary information, were used as categorical time-varying covariates.

Statistical analyses. Participant characteristics at the first visit of the study period and the first visit after menopause were described using numbers of women (%) and median values (interquartile range; IQR). To visualize a longitudinal trend, SU level was summarized as the mean with 95% confidence intervals (CI) for each integer year relative to the menopause. Similarly, HU was summarized as the prevalence with 95% CI computed using the Wald method for binomial data. The trend of eGFR was also depicted because eGFR declines with aging and is influential on SU levels. In addition, we supplementally described the SU trend in male participants matched on the age at the first visit and the total number of visits, to understand the difference in SU trends between men and women.

Based on the observed trends, we defined three stages based on menopause: stage 1 (premenopause) was defined as more than five years before menopause; stage 2 (perimenopause) was defined as the window from five years before up until menopause (year 0); and Stage 3 (postmenopause) was defined as the window after menopause. With these three stages, we performed interrupted time-series analyses to determine their relationship with SU levels. Because the study dataset contained repeated annual measurements for a given woman, the analyses were nested in a mixed-effect linear model with random intercepts and random slopes to account for intrapersonal correlations.

Three sets of models were considered with increasing adjustment. Model 1 included continuous time (years from menopause), categorical menopause stages, and their product term (equation is described in Supplemental Methods). Model 2 included additional covariates, such as age at menopause, BMI, eGFR, smoking status, alcohol intake, and medication use for hypertension, diabetes, and dyslipidemia. Model 3 additionally considered product (interaction) terms of menopausal stages with BMI and with eGFR; these variables are major contributors to SU levels. This allowed us to evaluate whether the association of eGFR or BMI with SU differed according to menopause stage. We conducted three sensitivity analyses based on model 2. First, we included only women who had at least two visits in each of three menopause stages. Second, we included dietary frequency questionnaire results in the model. We used data at visits after October 2012, and the participants were limited to those who underwent menopause after October 2012. Third, we excluded patients who were receiving medications for the aforementioned conditions to minimize their potential effect on SU levels. Moreover, we performed an additional analysis in which time from menopause was replaced with age in model 2.

Next, we examined the association between categorical menopause stages and HU by modified Poisson regression with participant-clustered robust variance.[33] Crude models included only menopause stages. Multivariable analyses included the same covariates as model 2, except for continuous time and its product term with menopause stages.

Stratified subgroup analyses were also pursued: the trends in SU and HU were visualized according to the subgroups stratified by median menopausal age, kidney function (eGFR <60 or ≥60 mL/min per 1.73 m^2), BMI (<25 or ≥25), and alcohol intake (drinkers or nondrinkers) at the first visit after menopause. For the SU trend, subgroup results based on model 1 were presented. For HU, we examined both additive and multiplicative effect modification by these subgroups on the association between menopausal stage and HU, adjusting for the same covariates as in the primary HU analysis. Specifically, for additive effect modification, we estimated relative excess risk due to

Table 1. Characteristics of participants at the first visits during follow-up and after menopause*

	First visit during follow-up (n = 8,169)	First visit after menopause (n = 8,169)
Age, y	45.0 (41.0–49.0)	52.0 (51.0–54.0)
Body mass index	20.5 (19.0–22.4)	20.8 (19.2–23.0)
Body mass index category, n (%)		
<18.5 (underweight)	1,371 (16.8)	1,338 (16.4)
18.5–25 (normal weight)	6,040 (73.9)	5,810 (71.1)
25–30 (overweight)	631 (7.7)	829 (10.1)
≥30 (obese)	127 (1.6)	192 (2.4)
Time from menopause, y	−5.9 (−10.4 to −2.5)	0.5 (0.5 to 0.8)
Menopausal category, n (%)		
Premenopause	4,632 (56.7)	0 (0)
Perimenopause	3,537 (43.3)	0 (0)
Postmenopause	0 (0)	8,169 (100)
Serum urate, mg/dL	4.2 (3.7–4.8)	4.8 (4.2–5.4)
Hyperuricemia (≥6.8 mg/dL), n (%)	81 (1.0)	328 (4.0)
eGFR, mL/min per 1.73 m^2	83.1 (74.0–93.3)	73.1 (65.7–81.7)
Smoking status, n (%)		
Never	6,663 (81.6)	6,770 (82.9)
Previous	874 (10.7)	969 (11.9)
Current	632 (7.7)	430 (5.3)
Regular alcohol drinker	4,263 (52.2)	4,113 (50.3)
Alcohol intake among drinkers, drink per day	1.0 (0.4–2.0)	1.0 (0.4–2.1)
Diagnosis of gout and/or hyperuricemia, n (%)	4 (0.0)	9 (0.1)
Medication use for conditions, n (%)		
Hyperuricemia/gout	3 (0.0)	9 (0.1)
Hypertension	165 (2.0)	572 (7.0)
Diabetes	22 (0.3)	89 (1.1)
Dyslipidemia	68 (0.8)	404 (4.9)
Cerebrovascular disease	10 (0.1)	24 (0.3)
Ischemic heart disease	5 (0.1)	17 (0.2)
Kidney disease	4 (0.0)	7 (0.1)
Tuberculosis or NTM infection	0 (0.0)	6 (0.1)

* eGFR, estimated glomerular filtration rate; NTM, nontuberculous mycobacteria.

interaction (RERI) by creating the joint categories of menopausal stage and each subgroup, and then estimating prevalence ratios for HU using modified Poisson regression. For the multiplicative effect modification, the model additionally included a product term of menopausal stages and each subgroup variable.

We performed complete case analyses because only 16 participants (<1.0%) had missing values. Point estimates and 95% CI were reported without *P* values, considering the observational nature of this study. We conducted all statistical analyses with Stata (Version 18.0, Stata Corp LLC).

Availability of data and material. Data are not publicly available due to our institution's regulations. The scripts for Stata software (version 18.0, StataCorp, College Station) in this study are available upon reasonable request to the corresponding author (SF).

RESULTS

Participant characteristics at first visit and menopause visit. We identified 8,169 eligible participants with 93,511 visits for the analyses (Supplemental Figure 1). The participants had a median of 12 visits (IQR 6–16), with a median follow-up of 13.8 years (IQR 8.3–17.9). The median number of visits per participant was one (IQR 0–4) in premenopause, four (IQR 2–5) in perimenopause, and four (IQR 2–8) in postmenopause. The first

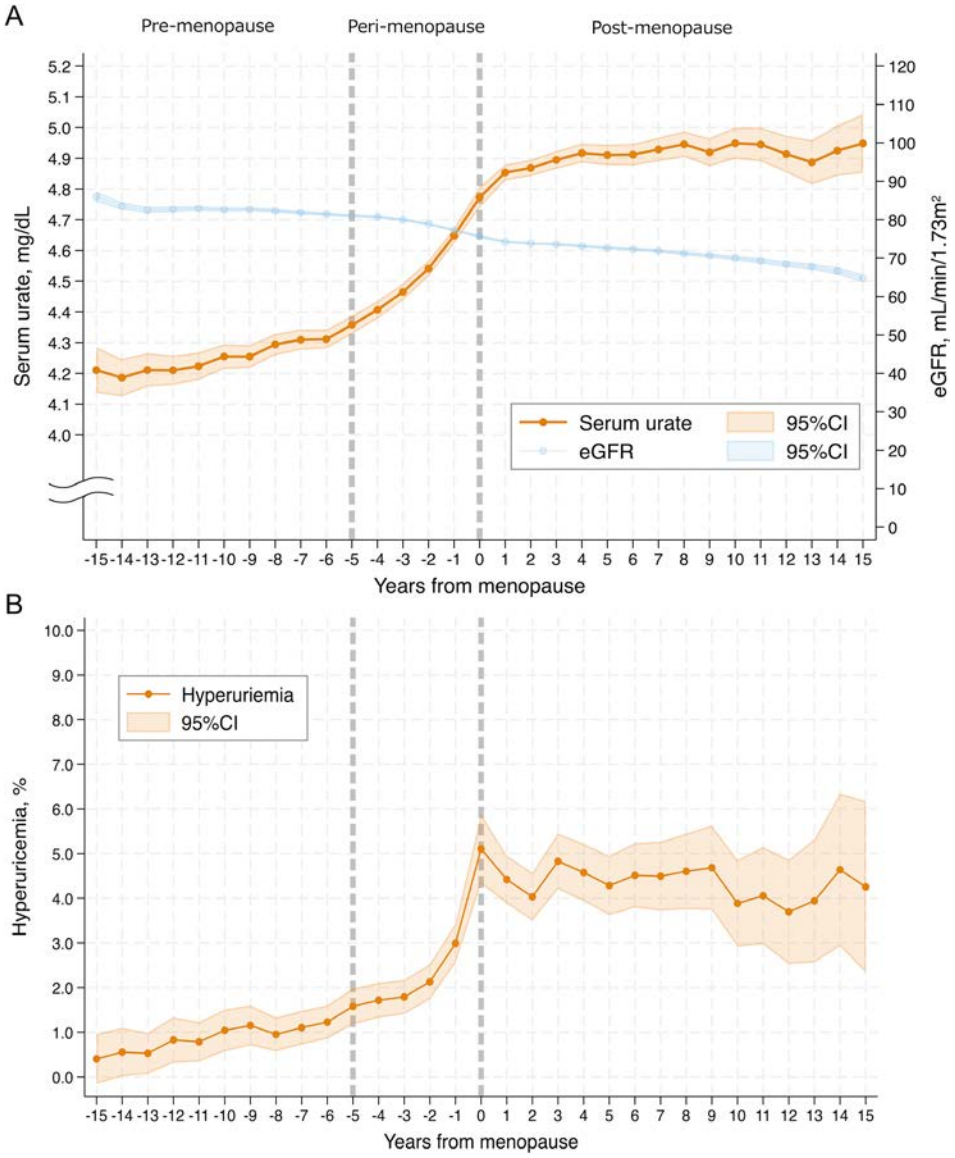


Figure 1. Longitudinal trends in serum urate and hyperuricemia. Mean serum urate levels (A) and the proportion of hyperuricemia (B) with a 95% CI were described. The date of menopause was defined as the median date between the two consecutive visits immediately before and after the self-reported menopause occurrence. The time from menopause was calculated for each visit. Participant data were then summarized by years from menopause. CI, confidence interval; eGFR, estimated glomerular filtration rate. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.43406/abstract.

visits occurred 5.9 years (IQR 2.5–10.4) before menopause. The participants' median age at the first visit was 45.0 years, and median BMI was 20.5. At the first visit, 4,632 (56.7%) were premenopausal and 3,537 (43.3%) were perimenopausal. The median SU level was 4.2 mg/dL (IQR 3.7–4.8), and 81 participants (1.0%) had HU (Table 1). At the first visit after menopause (+0.5 years [IQR 0.5–0.8] from menopause), the median age was 52 years, with a median BMI of 20.8. The median SU level increased to 4.8 mg/dL (IQR 4.2–5.4), and the proportion of those with HU increased to 328 participants (4.0%) (Table 1). Participants characteristics at the first visit within each menopausal stage were described in Supplemental Table 1. Characteristics of the subpopulation with dietary questionnaire results were similar to the primary study population (Supplemental Result and Supplemental Table 2).

Trends in SU levels and HU across menopause. Mean SU levels gradually increased during the premenopausal stage, followed by a rapid rise during perimenopause. Thereafter, the mean SU level remained stable throughout the postmenopausal stage, although eGFR steadily decreased over time (Figure 1A; Supplemental Table 3). The proportion of HU was between 0.5% to 1.0% during premenopause and began to increase approximately two to three years before menopause. In the postmenopausal stage, 4% to 5% of total female participants had HU (Figure 1B; Supplemental Table 3). This trend was specific to women compared to matched male controls. (Supplemental Figure 2).

Results of interrupted time-series analyses for SU level and HU. Compared to premenopause, model 1 (crude model) showed the mean SU level was higher by 0.286 mg/dL (95% CI: 0.261–0.310) in perimenopause and 0.410 mg/dL (95% CI: 0.387–0.433) in postmenopause. In terms of slopes (annual SU change), SU increased by 0.033 mg/dL per year (95% CI: 0.030–0.035) during premenopause. The slope then changed to 0.095 mg/dL (change from premenopause: +0.062 mg/dL; 95% CI: 0.057–0.067) during perimenopause, and 0.016 mg/dL (change from premenopause: –0.017 mg/dL; 95% CI: –0.021 to –0.014) in postmenopause. Model 2 (adjusted model) showed similar results. Models 2 and 3 showed that the slope of SU in the postmenopausal stage was close to 0, based on β coefficients for time and its product term with menopause stages (Table 2). These model-based estimates are visualized in Figure 2.

As shown in model 3 results, the association between SU and eGFR or SU and BMI became larger during perimenopause and postmenopause compared with premenopause (Table 2). For instance, a one-unit increase in BMI was associated with an increase of 0.076 mg/dL (95% CI: 0.072–0.080) in SU levels during the premenopausal stage; the association became 0.089 mg/dL (95% CI: 0.085–0.092) in the postmenopausal stage (Table 2).

The first sensitivity analysis included 2,878 participants with 45,430 visits who had at least two visits for each of the menopause stages. The results in this population did not differ from the primary analysis. Another sensitivity analysis using the lifestyle questionnaire results was broadly similar to the results of the main analyses. The results were similar after we excluded patients who were receiving medications for the aforementioned conditions. (Supplemental

Table 2. Result of interrupted time-series analyses for serum urate levels*

	Model 1, β coefficient (95%CI)	Model 2, β coefficient (95%CI)	Model 3, β coefficient (95%CI)
Mean SU level difference by the menopausal stages			
Premenopause	0 (Reference)	0 (Reference)	0 (Reference)
Perimenopause	0.286 (0.261–0.310)	0.261 (0.238–0.285)	0.334 (0.242–0.426)
Postmenopause	0.410 (0.387–0.433)	0.387 (0.365–0.409)	0.369 (0.267–0.471)
SU slope in premenopausal stage (annual SU change)	0.033 (0.030–0.035)	0.020 (0.017–0.023)	0.022 (0.019–0.025)
Additional slope change			
Premenopause	0 (Reference)	0 (Reference)	0 (Reference)
Perimenopause	0.062 (0.057–0.067)	0.055 (0.051–0.060)	0.053 (0.049–0.058)
Postmenopause	−0.017 (−0.021 to −0.014)	−0.019 (−0.022 to −0.015)	−0.022 (−0.025 to −0.019)
eGFR, per mL/min per 1.73 m^2		−0.013 (−0.014 to −0.013)	−0.011 (−0.012 to −0.011)
Product term of eGFR and menopause stages			
Premenopause			0 (Reference)
Perimenopause			−0.002 (−0.003 to −0.002)
Postmenopause			−0.003 (−0.004 to −0.003)
BMI		0.084 (0.081–0.087)	0.076 (0.072–0.080)
Product term of BMI and menopause stages			
Premenopause			0 (Reference)
Perimenopause			0.005 (0.001–0.008)
Postmenopause			0.013 (0.009–0.017)

* Model 1 included time from menopause + menopause stage + time from menopause*menopause stage. Model 2 included model 1 variables + age at menopause + BMI + eGFR + smoking status + alcohol intake + medication use for hypertension, diabetes, and dyslipidemia. Model 3 included model 2 variables + eGFR*menopause stage + BMI*menopause stage. CI, confidence interval; BMI, body mass index; eGFR, estimated glomerular filtration rate; SU, serum urate.

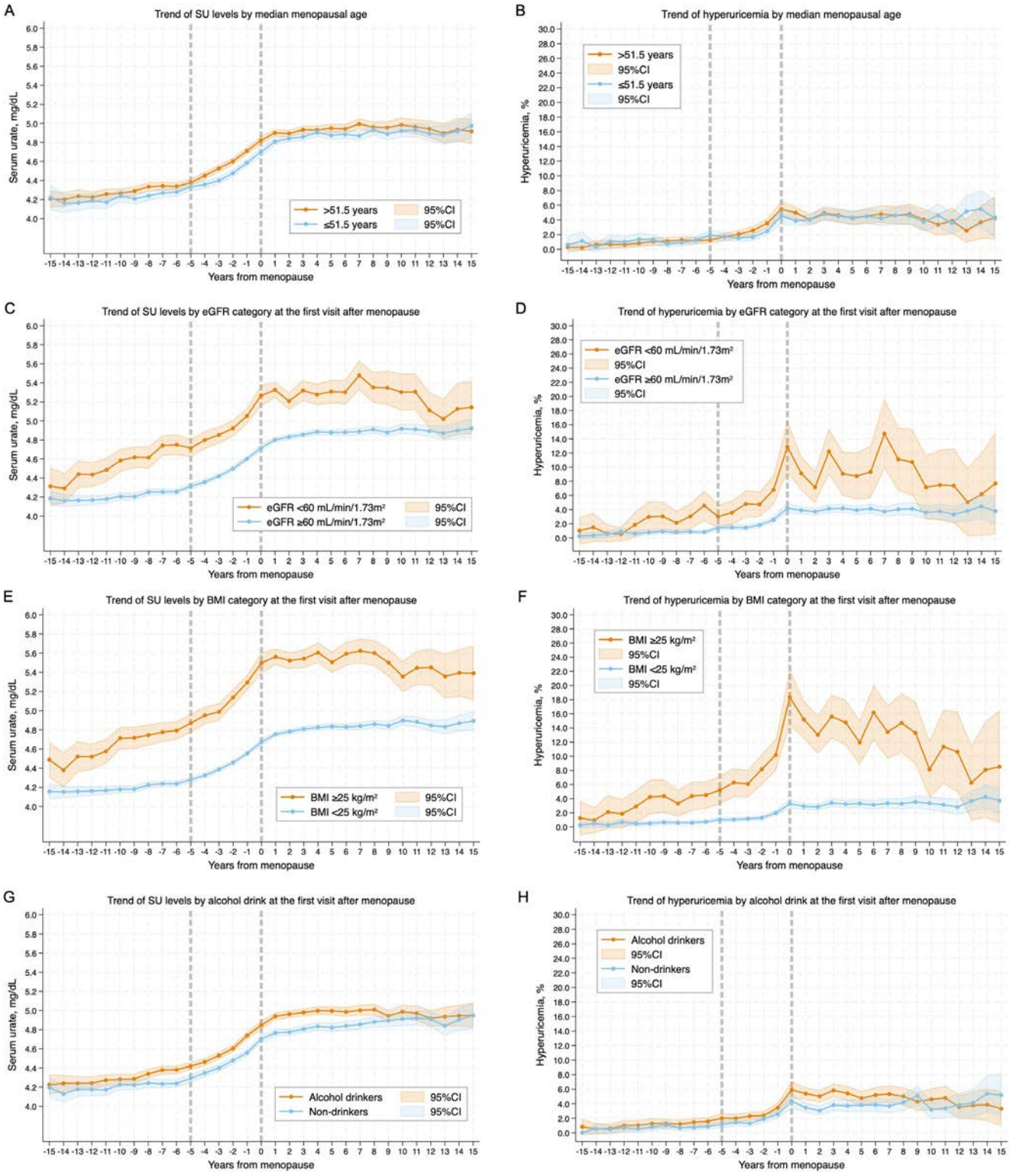


Figure 2. Longitudinal trends in SU by subgroups. Mean SU levels (A, C, E, G) and the proportions of hyperuricemia (B, D, F, H) with a 95% CI were described by stratified subgroups of menopausal age (A, B), eGFR (C, D), BMI (E, F), and alcohol intake (G, H). The date of menopause was defined as the median date between the two consecutive visits immediately before and after the self-reported menopause occurrence. The time from menopause was calculated for each visit. Participant data were then summarized by years from menopause. The information at the first visit after menopause was used for stratification. CI, confidence interval; BMI, body mass index; eGFR, estimated glomerular filtration rate; SU, serum urate.

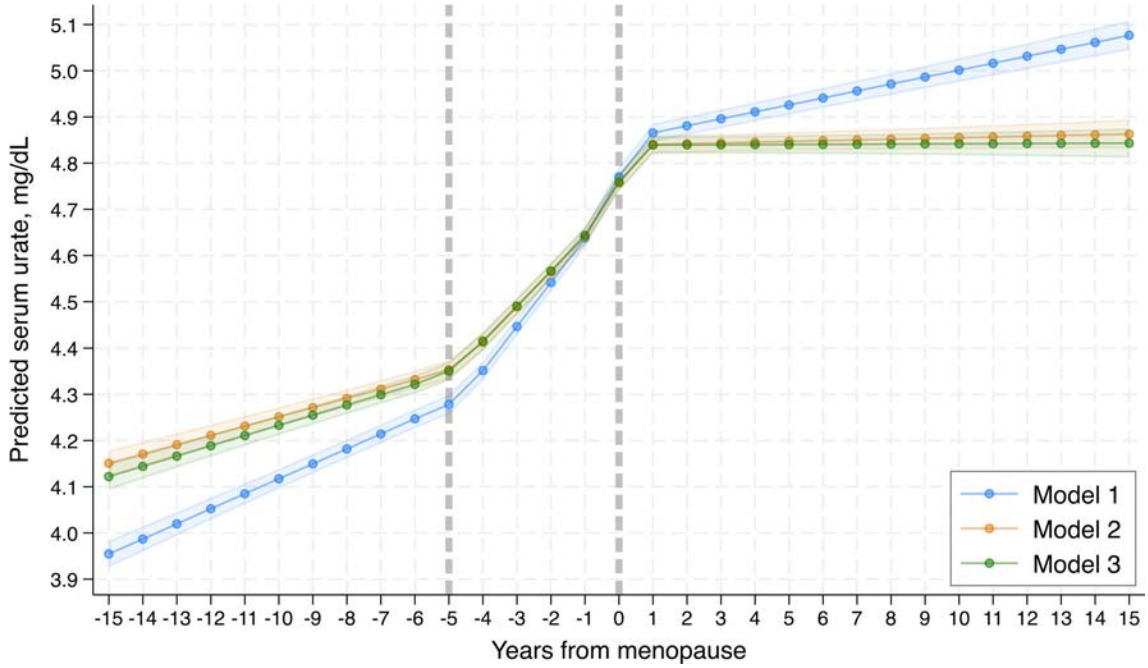


Figure 3. Predicted serum urate level by the different models. Mean serum urate levels and 95% confidence intervals were estimated using mixed-effect models (model 1–3). The date of menopause was defined as the median date between the two consecutive visits immediately before and after the self-reported menopause occurrence. The time from menopause was calculated for each visit. Model 1 included continuous time (years from menopause), categorical menopause stages, and their product term. Model 2 included additional covariates, such as age at menopause, BMI, eGFR, smoking status, alcohol intake, and medication use for hypertension, diabetes, and dyslipidemia. Model 3 additionally adjusted for product terms of menopause stages with BMI and with eGFR. In this analysis, we assumed a hypothetical patient who was a nonsmoker and received no medications. For continuous covariates (menopausal age, body mass index, alcohol intake, and estimated glomerular filtration rate), we set each value to the population mean. BMI, body mass index; eGFR, estimated glomerular filtration rate. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.43406/abstract.

Table 4). The results of the analysis that replaced time from menopause with age are presented in Supplemental Figure 3.

Compared to premenopause, both perimenopause and postmenopause stages were positively associated with HU. With premenopause as a reference, the relative risk of postmenopause for HU was 4.27 (95% CI: 3.33–5.47) in the univariable model and 2.32 (95% CI: 1.79–3.00) in the multivariable model (Supplemental Table 5).

Trends in SU levels and HU by subgroups. Longitudinal trends in eGFR, BMI, and alcohol intake across subgroups are presented in Supplemental Table 6. The trends in SU and HU did not substantially differ between the subgroups stratified by median menopausal age or alcohol intake at the first visit after menopause. However, participants with a lower eGFR (<60 mL/min per 1.73 m^2) and those who were overweight or had obesity (BMI ≥25) at menopause had more frequent HU during the postmenopausal stage (Figure 3; Supplemental Table 7). HU was observed in 3% of participants with a BMI <25 at the first visit after menopause; in contrast, 18% of female participants with a BMI ≥25 had HU (Figure 3). With an increase in BMI quartile, SU levels and HU prevalence also increased (Supplemental Figure 4). Positive RERI values indicated that the association of postmenopause (vs premenopause) with HU was stronger in women with kidney dysfunction and in those who were overweight or obese (Table 3).

DISCUSSION

This longitudinal study examined temporal trends in SU levels and HU prevalence across the midlife of women. The same women of varying ages were observed from the pre- or perimenopausal stage to the postmenopausal stage. We identified unique patterns in SU and HU specific to women in relation to the menopausal transition. SU levels gradually increased during premenopause, had a sharp rise during perimenopause, and became stable in the postmenopausal stage. Compared to premenopause, the mean SU level was 0.41 mg/dL higher in postmenopause. HU prevalence increased from <1% during premenopause to 4% to 5% during postmenopause. Overweight or obese women and those with kidney dysfunction had much higher SU levels over the menopausal transition. HU was much more commonly observed in these participants in the postmenopausal stage.

The present study found SU levels increase over the menopausal transition with a rapid rise in perimenopause. Other factors associated with SU levels, such as BMI, kidney function, and lifestyle factors, may also change during a women's midlife. However, the association between menopausal stages and SU levels

Table 3. Hyperuricemia by menopausal stage across subgroups: prevalence ratios and effect modification*

	Prevalence ratio for hyperuricemia (95% CI)		Additive effect modification (RERI), β coefficient (95% CI)	Multiplicative effect modification, β coefficient (95% CI)
Subgroup 1: age at menopause, y	≤51.5	>51.5		
Premenopause	1 (Reference)	0.67 (0.41, 1.08)	0 (Reference)	0 (Reference)
Perimenopause	1.19 (0.88–1.62)	1.35 (0.92–1.98)	0.49 (0.18–0.79)	0.52 (0.11–0.94)
Postmenopause	1.94 (1.35–2.79)	1.74 (1.19–2.56)	0.13 (−0.33 to 0.59)	0.29 (−0.20 to 0.78)
Subgroup 2: eGFR at menopause, mL/min per 1.73m^2	≥60	<60		
Premenopause	1 (Reference)	2.60 (1.57–4.29)	0 (Reference)	0 (Reference)
Perimenopause	2.06 (1.59–2.67)	4.76 (3.20–7.09)	1.10 (−0.13 to 2.34)	−0.12 (−0.56 to 0.33)
Postmenopause	4.06 (3.00–5.48)	8.83 (6.10–12.79)	3.18 (0.92–5.45)	−0.18 (−0.71 to 0.36)
Subgroup 3: BMI at menopause	<25	≥25		
Premenopause	1 (Reference)	5.73 (3.53–9.32)	0 (Reference)	0 (Reference)
Perimenopause	1.92 (1.47–2.51)	9.46 (6.50–13.77)	2.80 (0.57–5.03)	−0.15 (−0.57 to 0.27)
Postmenopause	3.02 (2.20–4.16)	11.54 (7.96–16.73)	3.78 (0.74–6.82)	−0.41 (−0.90 to 0.09)
Subgroup 4: Alcohol intake at menopause	Nonregular drinkers	Regular drinkers		
Premenopause	1 (Reference)	1.75 (1.08–2.85)	0 (Reference)	0 (Reference)
Perimenopause	1.77 (1.27–2.46)	2.76 (1.84–4.14)	0.24 (−0.35 to 0.83)	−0.11 (−0.54 to 0.31)
Postmenopause	2.58 (1.75–3.81)	3.77 (2.53–5.63)	0.44 (−0.32 to 1.20)	−0.18 (−0.67 to 0.31)

* Modified Poisson regression with patient-clustered robust variance were used for the analysis. Adjusted for age at menopause, BMI, eGFR, smoking status, alcohol intake, and medication use for hypertension, diabetes, and dyslipidemia; the covariate used for making subgroup was removed from each analysis. CI, confidence interval; BMI, body mass index; eGFR, estimated glomerular filtration rate; RERI, relative excess risk due to interaction.

was maintained after adjusting for these factors, suggesting a direct influence of menopause on SU levels. Based on the observed temporal trends, SU is expected to be influenced by changes occurring years before menopause. The effects of estrogen and progesterone on renal urate elimination may partially explain the SU rise before menopause. In a murine model, estrogen suppresses the protein expression of Urat1, Glut9, and Abcg2; as well, progesterone suppresses levels of Smct1; these are all renal urate transporters.[34] Approximately 70% of uric acid is excreted through the kidneys, with 30% excreted through the gastrointestinal tract.[35,36] Estrogen levels fluctuate during the first half of perimenopause and begin to decline one to two years before menopause, with the decline being substantial in postmenopause.[37,38] A study in 259 premenopausal women observed endogenous estrogen and progesterone having an inverse association with SU, whereas follicle-stimulating hormone (FSH) was positively associated with SU levels.[39] In the Study of Women's Health Across the Nation, FSH started to increase 6.1 years before the menopause,[40] and the 2011 Stages of Reproductive Aging Workshop (STRAW + 10) noted that FSH gradually increased two to three years before the menopause.[41] In addition to estrogen and progesterone, biologic studies on FSH effects on SU during premenopause or early perimenopause may potentially help provide the mechanism behind the present study's observation that SU begins to rise approximately five years before the menopause.

In contrast to previous cross-sectional studies,[14,15,21] the current study followed the same participants longitudinally and therefore is able to detect the timing of HU onset more precisely. In this study, the prevalence of HU started to increase in the two to three years before menopause and was elevated by the time of menopause. This result is consistent with a Korean study that showed the highest prevalence of HU during the late menopausal transition.[15] The current study findings suggest that the risk of gout may already be elevated by the time of menopause. Indeed, previous research reported that the incidence rate of gout in women in their 50s is already more than twice as high as in those under 45 years of age.[10] Gout risk may further increase in women in their 70s.[3,10] However, this increase could be influenced by other factors such as medications that affect SU levels as well as increased comorbidities, including kidney dysfunction.

The current study examined the coexisting conditions associated with SU levels over the menopausal transition. Although age at menopause and alcohol intake did not yield clinically relevant differences, we found that BMI and eGFR play an important role in the relationship between menopause and SU. Of note, the associations of SU or HU with BMI and eGFR become greater in postmenopause compared to premenopause. It is biologically plausible that menopause interacts with other factors impacting SU. Menopause, obesity, and kidney dysfunction are all associated with decreased renal urate elimination.[42–44] Subgroup analyses revealed that SU levels and HU prevalence in postmenopause were markedly higher in overweight or obese women and those with a lower eGFR. These findings suggest that maintaining a healthy body weight and preserving kidney function before menopause may help reduce postmenopausal HU, potentially lowering the subsequent risk of gout.

The present study has limitations. First, the quasi-experimental nature of this study may limit the ability to establish a causal relationship between menopause and SU increases.

However, results were consistent after adjusting for relevant confounders and across multiple sensitivity analyses. Biologic mechanisms also support the observed association between menopause and SU levels. Second, hormone levels were not measured in the annual medical examinations. Menopausal stages were determined based on self-reported menopause and observed SU trend in this study, although staging based on the 2011 STRAW +10 is recommended for reproductive aging.[41] Nevertheless, the STRAW+10 criteria also define menopause primarily based on the absence of menstruation, with hormonal measurements serving only as supporting information. Prior studies have demonstrated high reproducibility of self-reported menopausal status.[45] Moreover, the median age at menopause in this study was similar to that from the Study of Women's Health Across the Nation, in which the median age at menopause was 52.5.[46] Third, the SLHCD did not contain detailed information on medications. However, our results remained consistent after excluding participants who received medications for relevant comorbidities. We were unable to conduct certain subgroup analyses, such as among users of hormone replacement therapy or urate-lowering therapy, because individuals using hormone therapy were excluded from the study, and the number of participants receiving treatment for gout or HU was limited. Finally, the inclusion of only Japanese women who participated in annual medical examinations presents a cohort that may not translate globally, including differences in socioeconomic status and education levels. In this study, only a limited proportion of participants (approximately 10%) had comorbidities treated with medications at the time of menopause. In comparison, more than 30% of women in the United States use cardiovascular medications, such as antihypertensives and lipid-lowering therapies, during menopause.[47] Based on the results of subgroup analyses, the impact of menopause on SU might be more pronounced in the United States and in other countries where obesity is more frequent.

In conclusion, this longitudinal study found that the rise in SU levels and the prevalence of HU begins several years before menopause, with the largest increase occurring during perimenopause. These changes are more pronounced in overweight or obese women and those with a lower eGFR. Maintaining a healthy body weight and preserving kidney function in the early stages before menopause may help prevent postmenopausal HU and potentially reduce the risk of developing gout. Further longitudinal and biologic studies will help elucidate how changing biologic factors during the midlife of women impact SU.

ACKNOWLEDGMENTS

We express our gratitude to the health care workers who contributed to St. Luke's annual medical examination program. We greatly appreciate the advice on SLHCD from Dr. Kimura. We sincerely thank Mr. Nakajima and the Information System Unit in St. Luke's International Hospital for their data collection. During the preparation of this work the authors used ChatGPT in order to check grammar and styles. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr Fukui confirms that all authors have provided the final approval of the version to be published and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. Dehlin M, Jacobsson L, Roddy E. Global epidemiology of gout: prevalence, incidence, treatment patterns and risk factors. Nat Rev Rheumatol 2020;16(7):380–390.
2. Roddy E, Doherty M. Epidemiology of gout. Arthritis Res Ther 2010;12(6):223.
3. Rai SK, Aviña-Zubieta JA, McCormick N, et al. The rising prevalence and incidence of gout in British Columbia, Canada: population-based trends from 2000 to 2012. Semin Arthritis Rheum 2017;46(4):451–456.
4. Yokose C, McCormick N, Lu N, et al. Trends in prevalence of gout among US Asian adults, 2011-2018. JAMA Netw Open 2023;6(4):e239501.
5. Cross M, Ong KL, Culbreth GT, et al; GBD 2021 Gout Collaborators. Global, regional, and national burden of gout, 1990-2020, and projections to 2050: a systematic analysis of the Global Burden of Disease Study 2021. Lancet Rheumatol 2024;6(8):e507–e517.
6. Kyu HH, Abate D, Abate KH, et al; GBD 2017 DALYs and HALE Collaborators. Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018;392(10159):1859–1922.
7. Dalbeth N, Phipps-Green A, Frampton C, et al. Relationship between serum urate concentration and clinically evident incident gout: an individual participant data analysis. Ann Rheum Dis 2018;77(7):1048–1052.
8. Bhole V, de Vera M, Rahman MM,et al. Epidemiology of gout in women: fifty-two-year followup of a prospective cohort. Arthritis Rheum 2010;62(4):1069–1076.
9. McCormick N, Yokose C, Challener GJ, et al. Serum urate and recurrent gout. JAMA 2024;331(5):417–424.
10. Hak AE, Curhan GC, Grodstein F, et al. Menopause, postmenopausal hormone use and risk of incident gout. Ann Rheum Dis 2010;69(7):1305–1309.
11. Kumar S, Gupta R, Suppiah R. Gout in women: differences in risk factors in young and older women. N Z Med J 2012;125(1363):39–45.
12. Patel AV, Gaffo AL. Managing gout in women: current perspectives. J Inflamm Res 2022;15:1591–1598.
13. Wallace KL, Riedel AA, Joseph-Ridge N, et al. Increasing prevalence of gout and hyperuricemia over 10 years among older adults in a managed care population. J Rheumatol 2004;31(8):1582–1587.
14. Hak AE, Choi HK. Menopause, postmenopausal hormone use and serum uric acid levels in US women--the Third National Health

and Nutrition Examination Survey. Arthritis Res Ther 2008;10(5): R116.
15. Cho SK, Winkler CA, Lee SJ, et al. The prevalence of hyperuricemia sharply increases from the late menopausal transition stage in middle-aged women. J Clin Med 2019;8(3):296.
16. Nicholls A, Snaith ML, Scott JT. Effect of oestrogen therapy on plasma and urinary levels of uric acid. BMJ 1973;1(5851):449–451.
17. He H, Pan L, Liu F, et al. The mediation effect of body composition on the association between menopause and hyperuricemia: evidence from China National Health Survey. Front Endocrinol (Lausanne) 2022;13:879384.
18. Luo Y, Wu Q, Meng R, et al. Serum uric acid levels and their association with renal function decline and kidney disease progression in middle-aged and elderly populations: a retrospective cohort study. J Multidiscip Healthc 2023;16:3357–3366.
19. Marrugo J, Santacroce LM, Paudel ML, et al. Longitudinal analysis of serum urate in prediabetic phase. Rheumatology (Oxford) 2024; 63(11):2976–2980.
20. Kaneshita S, Fukui S, Niku S, et al. Association of chronic kidney disease with increased glucose uptake-associated metabolic activity of visceral adipose and subcutaneous adipose tissue in patients with gout. Arthritis Care Res (Hoboken) 2025;77(8):1022–1031.
21. Krishnan E, Bennett M, Chen L. Aging, not menopause, is associated with higher prevalence of hyperuricemia among older women. Menopause 2014;21(11):1211–1216.
22. Fukui S, Okada M, Rahman M, et al. Differences in the association between alcoholic beverage type and serum urate levels using standardized ethanol content. JAMA Netw Open 2023;6(3):e233398.
23. Fukui S, Okada M, Shinozaki T, et al. Changes in alcohol intake and serum urate changes: longitudinal analyses of annual medical examination database. Ann Rheum Dis 2024;83(8):1072–1081.
24. Fukui S, Okada M, Shinozaki T, et al. Weight reduction and target serum urate level: a longitudinal study of annual medical examination. Arthritis Rheumatol 2025;77(3):346–355.
25. Imai R, Tomishima Y, Nakamura T, et al. Prognosis of equivocal interstitial lung abnormalities in a health check-up population. Ann Am Thorac Soc 2025;22(3):387–394.
26. Fukui S, Ikeda Y, Kataoka Y, et al. Clinical significance of monitoring hypothyroidism in patients with autoimmune rheumatic disease: a retrospective cohort study. Sci Rep 2021;11(1):13851.
27. Suzuki T, Fukui S, Shinozaki T, et al. Lipid profiles after changes in alcohol consumption among adults undergoing annual checkups. JAMA Netw Open 2025;8(3):e250583.
28. von Elm E, Altman DG, Egger M, et al; STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Ann Intern Med 2007;147(8):573–577.
29. Hisatome I, Ichida K, Mineo I, et al. Japanese Society of Gout and Uric & Nucleic Acids 2019 Guidelines for Management of Hyperuricemia and Gout 3rd edition. Gout and Uric & Nucleic Acids. 2020;44-(Supplement):1–40.
30. Matsuo S, Imai E, Horio M, et al; Collaborators developing the Japanese equation for estimated GFR. Revised equations for estimated GFR from serum creatinine in Japan. Am J Kidney Dis 2009; 53(6):982–992.
31. Leung N, Yip K, Pillinger MH, et al. Lowering and raising serum urate levels: off-label effects of commonly used medications. Mayo Clin Proc 2022;97(7):1345–1362.
32. Marrugo J, Santacroce LM, Paudel ML, et al. Gout risk in adults with pre-diabetes initiating metformin. Ann Rheum Dis 2024;83(10): 1368–1374.
33. Zou G. A modified poisson regression approach to prospective studies with binary data. Am J Epidemiol 2004;159(7):702–706.
34. Takiue Y, Hosoyamada M, Kimura M, et al. The effect of female hormones upon urate transport systems in the mouse kidney. Nucleosides Nucleotides Nucleic Acids 2011;30(2):113–119.
35. Bobulescu IA, Moe OW. Renal transport of uric acid: evolving concepts and uncertainties. Adv Chronic Kidney Dis 2012;19(6): 358–371.
36. Maesaka JK, Fishbane S. Regulation of renal urate excretion: a critical review. Am J Kidney Dis 1998;32(6):917–933.
37. Su HI, Freeman EW. Hormone changes associated with the menopausal transition. Minerva Ginecol 2009;61(6):483489.
38. Burger HG, Dudley EC, Robertson DM, et al. Hormonal changes in the menopause transition. Recent Prog Horm Res 2002;57(1): 257–275.
39. Mumford SL, Dasharathy SS, Pollack AZ, et al. Serum uric acid in relation to endogenous reproductive hormones during the menstrual cycle: findings from the BioCycle study. Hum Reprod 2013;28(7): 1853–1862.
40. Randolph JF Jr, Zheng H, Sowers MR, et al. Change in follicle-stimulating hormone and estradiol across the menopausal transition: effect of age at the final menstrual period. J Clin Endocrinol Metab 2011;96(3):746–754.
41. Harlow SD, Gass M, Hall JE, et al; STRAW 10 Collaborative Group. Executive summary of the Stages of Reproductive Aging Workshop + 10: addressing the unfinished agenda of staging reproductive aging. Menopause 2012;19(4):387–395.
42. Facchini F, Chen YD, Hollenbeck CB, et al. Relationship between resistance to insulin-mediated glucose uptake, urinary uric acid clearance, and plasma uric acid concentration. JAMA 1991;266(21):3008–3011.
43. Freeman EW, Sammel MD, Lin H, et al. Obesity and reproductive hormone levels in the transition to menopause. Menopause 2010;17(4): 718–726.
44. Farahmand M, Ramezani Tehrani F, et al. Endogenous estrogen exposure and chronic kidney disease; a 15-year prospective cohort study. BMC Endocr Disord 2021;21(1):155.
45. Colditz GA, Stampfer MJ, Willett WC, et al. Reproducibility and validity of self-reported menopausal status in a prospective cohort study. Am J Epidemiol 1987;126(2):319–325.
46. El Khoudary SR, Greendale G, Crawford SL, et al. The menopause transition and women's health at midlife: a progress report from the Study of Women's Health Across the Nation (SWAN). Menopause 2019;26(10):1213–1227.
47. Solomon DH, Ruppert K, Greendale GA, et al. Medication use by race and ethnicity in women transitioning through the menopause: a Study of Women's Health Across the Nation drug epidemiology study. J Womens Health (Larchmt) 2016;25(6):599–605.

Arthritis & Rheumatology
Vol. 78, No. 3, March 2026, pp 734–742
DOI 10.1002/art.43412

American College of Rheumatology
Empowering Rheumatology Professionals

Diagnostic Implications and Correlates of Plasma Adenosine Deaminase 2 Activity and *ADA2* Variants

Jian Yue,[1,2] Courtney E. LeSon,[1] Casey A. Rimland,[1,3] Kailey Brodeur,[1,4] Rachel Weng,[1] Evan E. Hsu,[1] Michael T. Lam,[1] Manpreet Meyer,[1] Seigo Okada,[1,5] Andrew L. D. Hsu,[1] Musaab A. Alhezam,[1] Megan M. Perron,[1] Olha Halyabar,[1] Fatma Dedeoglu,[1] Peter A. Nigrovic,[1,3] Akiko Shimamura,[6] R. Grant Rowe,[6] Megan A. Cooper,[7] Tiphanie P. Vogel,[8] Miriah Gillispie-Taylor,[9] Rosemary Peterson,[10] Clara Lin,[11] Katherine D. Nowicki,[11] Kaylee Dollerschell,[12] Taizo A. Nakano,[13] Roxane Labrosse,[14] Alexis-Virgil Cochino,[15,16] Andreea Ioan,[15,16] Leonardo Oliveira Mendonca,[17,18] Samar Freschi Barros,[17,18] Jorge Kalil,[17,18] Javier Perez-Rojas,[19] Alicia Diaz-Kuan,[20] Eric S. Sobel,[21] Taha Moussa,[22] Samuel Shang Ming Lee,[23] Li Wearn Koh,[23] Archana Khan,[24] Sodality Sutnga,[24] Raju Khubchandani,[24] Walaa Shoman,[25] Yasmine El Chazli,[26] Benzeeta Pinto,[27] Chengappa Kavadichanda,[28] Roshini S. Abraham,[29] Xiangwei Sun,[30] Qing Zhou,[30] Gsrsnk Naidu,[31] Aman Sharma,[31] Carl Esperanzate,[32] Daniel L. Kastner,[32] Ivona Aksentijevich,[32] Susan J. Kelly,[33] Michael Hershfield,[33,34] Teresa K. Tarrant,[33,35,36] and Pui Y. Lee[1]

Objective. Deficiency of adenosine deaminase 2 (DADA2) is a monogenic autoinflammatory disease manifested as polyarteritis nodosa, stroke, and bone marrow failure. Leveraging an international cohort of 200 DADA2 cases, we aimed to characterize the diagnostic utility of a plasma ADA2 enzyme activity assay and understand the implications of residual ADA2 activity.

Methods. Data were collected from individuals who underwent ADA2 testing from 2018 to 2025. Plasma ADA2 activity was determined using an established spectrophotometric assay. ADA2 variants were analyzed in transfected cells by enzyme assay and western blotting.

Results. We determined that plasma ADA2 activity is 99.0% and 96.0% sensitive and 99.7% and 98.8% specific in distinguishing genetically confirmed DADA2 cases from controls and carriers, respectively. Eighteen individuals with DADA2 (9%) possessed detectable ADA2 activity, including several cases with levels seen in carriers. Residual ADA2 activity was associated with the vasculitis/inflammatory phenotype but not with disease severity. Genotype analysis revealed that 14 of 18 cases with residual plasma activity possessed at least one hypomorphic missense variant with greater than 20% residual ADA2 function when overexpressed in 293T cells, often occurring in trans with a more deleterious variant. In vitro analysis revealed that missense ADA2 variants exert variable dominant-negative effects by forming large intracellular protein aggregates via disulfide bond formation at a cysteine residue (Cys408).

Conclusion. We confirmed the utility of plasma ADA2 activity as a diagnostic assay and showed that the inflammatory phenotype of DADA2 occurred in cases with residual activity. In vitro findings illustrate potential interactions of ADA2 variants to synergistically disrupt protein function.

Dr Rimland's work was supported by the Brigham and Women's Department of Medicine and the Rheumatology Fellowship Training Program. Dr Lam's work was supported by the Institutional National Research Service Award (T32; grant 232A1007512-38), NIH, and a Joint Biology Consortium Microgrant off parent award from National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS), NIH, grant 2P30-AR-070253. Mr Meyer's work was supported by Max Delbrück Center International PhD Program/Berlin Institute of Health Fellowship. Dr Nigrovic's work was supported by NIAMS, NIH (grant 2P30-AR-070253). Dr Sharma's work was supported by National Institute of Diabetes and Digestive and Kidney Diseases, NIH (grant 5RC2-DK-122533). Dr Lee's work was supported by the Boston Children's Hospital Faculty Career Development Fellowship, the Rheumatology Research Foundation R-Bridge Award and Innovative Research Award, and National Institute of Allergy and Infectious Diseases, NIH (grant 1R01-AI-187854-01A1).

[1]Division of Immunology, Boston Children's Hospital, Harvard Medical School, Boston, Massachusetts; [2]Department of Rheumatology and Immunology, Guangdong Second Provincial General Hospital, Guangzhou, China; [3]Division of Rheumatology, Inflammation, and Immunity, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts; [4]Medical Scientist Training Program, UMass Chan Medical School, Worcester, Massachusetts; [5]Department of Pediatrics, Yamaguchi University Graduate School of Medicine, Yamaguchi, Japan; [6]Dana-Farber/Boston Children's Cancer and Blood Disorders Center, Harvard Medical School, Boston, Massachusetts; [7]Division of Rheumatology and Immunology, Department of Pediatrics, Washington University in St. Louis, St. Louis, Missouri; [8]Division of

INTRODUCTION

Deficiency of adenosine deaminase 2 (DADA2), due to biallelic pathogenic variants in *ADA2*, is a monogenic disease with pleiotropic phenotypes including systemic vasculitis and bone marrow failure.[1–3] Since the discovery of DADA2 in 2014, over 600 cases have been described in the literature, and the carrier frequency of pathogenic *ADA2* variants is estimated to be around 1 of 236 individuals, which ranks DADA2 among the more common monogenic autoinflammatory syndromes.[4–6] However, an understanding of the physiologic function of ADA2 remains incomplete, and how impaired production of ADA2 incites heterogenous pathology has not been fully elucidated.

ADA2 catalyzes the conversion of free adenosine and 2'deoxyadenosine to inosine and 2'deoxyinosine, respectively.[7] Accurate measurement of ADA2 enzymatic activity requires both the use of saturating concentration of substrate and concomitant selective inhibition of ADA1, which has log-fold higher substrate affinity.[8,9] ADA2 may also act as a lysosomal protein that deaminates deoxyadenosine in DNA.[10]

Although a deaminase function of ADA2 remains to be confirmed in vivo, plasma ADA2 activity is accepted as a diagnostic assay for DADA2.[5,11–13] Different ADA2 enzymatic assays have been described, and near-absent activity is generally considered confirmatory for DADA2.[14] The utility of plasma ADA2 activity in distinguishing DADA2 cases from carriers and controls has been demonstrated in a number of publications, but this selectivity has not been systematically established because of the relatively small sample size of cohort studies and differences in methodology among laboratories.

In this study, we establish the utility of plasma ADA2 activity for diagnosis with an international cohort of 200 DADA2 cases. We describe a subgroup of DADA2 patients with measurable plasma ADA2 activity, including cases with activity in the range for carriers. We explore the relationships among residual plasma ADA2 activity, disease phenotype, disease severity, and causal *ADA2* variants. We further examine synergistic effects of *ADA2* variants in disrupting ADA2 protein expression.

MATERIALS AND METHODS

Study design and participants. We conducted a retrospective analysis of ADA2 plasma activity measured from individuals with DADA2, carriers, and healthy controls between January 2018 and May 2025. This study was approved by the Boston Children's Hospital Institutional Review Board (IRB-A00045581). The geographic distribution of DADA2 cases is shown in Figure 1A.

ADA2 enzymatic activity assay. ADA2 activity was determined in human plasma and culture media using an established spectrophotometric assay (see Supplemental Methods for details).[5,6,15,16] Plasma samples were assayed in duplicate. For evaluation of *ADA2* variants, enzyme activity in the supernatant of transfected 293T cells was measured 48 hours after transfection. Each variant was assessed by at least three independent experiments.

Analysis of ADA2 variants. Construction of expression plasmids for wildtype and mutant *ADA2* using pcDNA3.1-Myc/His(–) vector was described previously.[6,17] Plasmids were prepared using Monarch Plasmid Miniprep Kit (NEB) and transfected into HEK293T cells using Fugene 6 reagent (Promega). Western blotting of ADA2 was performed as described.[18] Additional

Rheumatology, Department of Pediatrics, Baylor College of Medicine and Center for Human Immunobiology, Texas Children's Hospital, Houston; [9]Division of Rheumatology, Department of Pediatrics, Huffington Department of Education, Innovation, and Technology, Baylor College of Medicine, Texas Children's Hospital, Houston; [10]Pediatric Rheumatology, Dell Children's Medical Group, Dell Medical School, University of Texas at Austin; [11]Section of Pediatric Rheumatology, Department of Pediatrics, Children's Hospital Colorado, University of Colorado School of Medicine, Aurora; [12]Department of Pediatrics, Children's Hospital Colorado, University of Colorado School of Medicine, Aurora; [13]Center for Cancer and Blood Disorders, Children's Hospital Colorado, University of Colorado School of Medicine, Aurora; [14]Section of Allergy and Clinical Immunology, Department of Pediatrics, Centre Hospitalier Universitaire Sainte-Justine, Montreal, Quebec, Canada; [15]Carol Davila University of Medicine and Pharmacy, Bucharest, Romania; [16]Alessandrescu Rusescu National Institute of Mother and Child Health, Bucharest, Romania; [17]Discipline of Clinical Immunology and Allergy, Faculdade de Medicina da Universidade de São Paulo, Sao Paulo, Brazil; [18]Laboratory for Medical Investigation (LIM-19), Instituto do Coração, Faculdade de Medicina da Universidade de São Paulo, Sao Paulo, Brazil; [19]Division of Clinical Immunology and Allergy, Department of Internal Medicine, Guillermo Almenara Irigoyen National Hospital, EsSalud, Lima, Peru; [20]Genetica Molecular, Universidad Peruana Cayetano Heredia, Lima, Peru; [21]Department of Medicine, University of Florida College of Medicine, Gainesville; [22]Section of Pediatric Rheumatology, University of Florida College of Medicine – Jacksonville; [23]Department of Rheumatology, Allergy and Immunology, Tan Tock Seng Hospital, Novena, Singapore; [24]Pediatric Rheumatology, SRCC Children's Hospital, Mumbai, India; [25]Pediatric Immunology and Rheumatology Unit, Department of Pediatrics, Faculty of Medicine, Alexandria University, Alexandria, Egypt; [26]Pediatric Hematology Unit, Department of Pediatrics, Faculty of Medicine, Alexandria University, Alexandria, Egypt; [27]Department of Clinical Immunology and Rheumatology, St. John's Medical College, St. John's National Academy of Health Sciences, Bengaluru, India; [28]Clinical Immunology and Rheumatology, RheumaCARE, Mysuru, India; [29]Diagnostic Immunology Laboratory, Department of Pathology and Laboratory Medicine, Nationwide Children's Hospital, Columbus, Ohio; [30]Department of Rheumatology, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine and Liangzhu Laboratory, Zhejiang University, Hangzhou, China; [31]Clinical Immunology and Rheumatology Division, Department of Internal Medicine, Postgraduate Institute of Medical Education and Research, Chandigarh, India; [32]Inflammatory Disease Section, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland; [33]Department of Medicine, Duke University School of Medicine, Durham, North Carolina; [34]Department of Biochemistry, Duke University School of Medicine, Durham, North Carolina; [35]Department of Pathology, Duke University School of Medicine, Durham, North Carolina; [36]Durham Veteran's Health Administration Health Care System, Durham, North Carolina.

Dr Yue, Ms LeSon, and Dr Rimland contributed equally to this work.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/art.43412).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/art.43412.

Address correspondence via email to Pui Y. Lee, MD, PhD, at pui.lee@childrens.harvard.edu.

Submitted for publication June 17, 2025; accepted in revised form August 27, 2025.

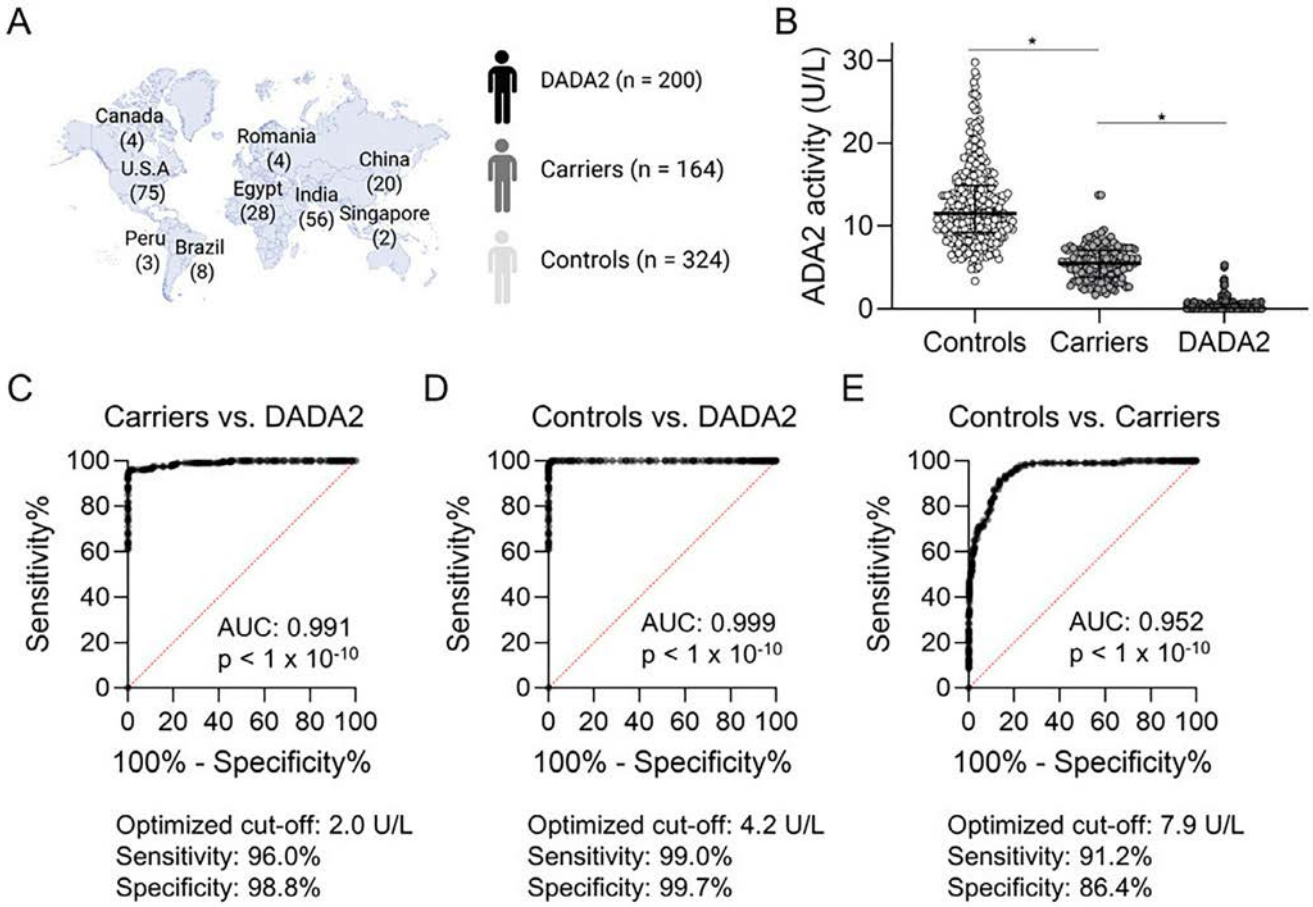


Figure 1. Plasma ADA2 activity as a diagnostic test for DADA2. (A) Geographic distribution of DADA2 cases included in the current study. Image was generated using BioRender. (B) Plasma ADA2 activity from DADA2 cases (n = 200), healthy controls (n = 324), and carriers with a heterozygous pathogenic ADA2 variant (n = 164). *$P < 1.0 \times 10^{-10}$ by Student's *t*-test. ROC curve depicting the utility of plasma ADA2 activity in discriminating carriers vs DADA2 cases (C), controls vs DADA2 cases (D), and controls vs carriers (E). AUC, *P* value, and optimized cut-off as determined by Youden's method are displayed. AUC, area-under-the-curve; DADA2, deficiency of adenosine deaminase 2; ROC, receiver operating characteristics. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.43412/abstract.

details of these experiments are provided in the Supplemental Methods. Predicted splice-site variants and large insertions/deletions were not analyzed because of the unknown impact at the complementary DNA level.

Statistics. Differences between groups were analyzed by the unpaired Student's *t*-test. All tests were two-sided, and *P* value <0.05 was considered significant. Receiver operating characteristics (ROC) curve was used to analyze the utility of ADA2 testing. Area-under-curve (AUC) was calculated, and optimized cut-off was determined using the Youden's index. Prism 10.0 software (GraphPad Software) was used for statistical analyses and graphical illustrations.

RESULTS

Characterization of ADA2 enzymatic activity as a diagnostic test. From 2018 to 2025, we measured plasma ADA2 activity in 200 individuals with DADA2 from different parts of the world using a validated coupled enzymatic assay that measures adenosine-derived ammonia spectrophotometrically (Figure 1A). This assay was cross-validated by re-analysis of a subset of samples (n = 73) in a Clinical Laboratory Improvement Amendment (CLIA)-certified laboratory using the same spectrophotometric assay protocol as well as using an independent high performance liquid chromatography-based ADA2 assay that measures inosine production (Supplemental Figure 1A and B).

The average age of disease onset was 6.7 years. At least 110 of these cases, including the first cases in Peru and Singapore, have not been described in previous cohort studies. We also analyzed plasma ADA2 activity from 324 healthy controls and 164 carriers that included parents of DADA2 cases as well as siblings of DADA2 cases and unrelated individuals confirmed to have a heterozygous deleterious *ADA2* variant (Figure 1B).

Based on the recent International Consensus Statement, a diagnosis of DADA2 can be established by near-absent plasma ADA2 activity and/or genetic studies demonstrating biallelic deleterious *ADA2* variants.[14] Genetic confirmation was available from

169 cases (84.5%); undetectable ADA2 activity alone was considered diagnostic for the remaining cases, for which genetic testing is pending or not available. Although undetectable ADA2 activity (below the lower limit of quantification of 1 U/L for our spectrophotometric assay) was observed in greater than 90% of individuals with DADA2, 18 cases among those with genetic confirmation (10.7%) exhibited detectable activity, including several with levels found in the range for carriers (Figure 1B). The highest level of activity detected in this subgroup was 5.3 U/L, near the median of the carrier group (5.5 U/L). As described previously,[5,19] higher levels of plasma ADA2 were observed in pediatric controls compared with adult controls (Supplemental Figure 2A). However, we did not find age-associated differences in ADA2 levels among carriers or in individuals with DADA2 (Supplemental Figure 2B and C).

The ROC curve showed that plasma ADA2 activity is effective in discriminating DADA2 cases from healthy controls (AUC 0.999) and carriers (AUC 0.991) (Figure 1C and D). Using the detection limit (1.0 U/L) as cut-off, this test delineated DADA2 cases from healthy controls and carriers with 88.5% sensitivity and 100% specificity. Using an optimized cut-off of 2 U/L, plasma ADA2 activity was able to distinguish DADA2 cases from carriers with 96.0% sensitivity and 98.8% specificity (Figure 1C). Six DADA2 cases possessed plasma ADA2 activity above 2 U/L, whereas three carriers displayed activity below this level. These data indicate that DADA2 cases can possess residual ADA2 activity, and genetic studies are necessary to distinguish individuals with detectable ADA2 activity from carriers. ADA2 activity was also less effective in separating carriers from controls (AUC 0.952; 91.2% sensitivity and 86.4% specificity based on an optimized cut-off of 7.9 U/L) (Figure 1E).

DADA2 cases with residual ADA2 activity. The predominant disease phenotypes of DADA2 include vasculitis/inflammation and severe hematologic involvement (pure red cell aplasia and bone marrow failure) and immunodeficiency.[14] Information on disease phenotype was available from 192 participants: 145 cases (75.5%) exhibited vasculitis/inflammation as the predominant phenotype, 16 cases (8.4%) had predominantly hematologic impairment, and 19 cases (10.0%) displayed overlapping disease phenotype. Presymptomatic siblings (n = 13; 6.8%) were found by family screening of confirmed cases. Residual plasma ADA2 activity was associated with the vasculitis/inflammation phenotype of DADA2 (Figure 2A). Patients with other disease phenotypes did not exhibit ADA2 activity above the detection threshold. All presymptomatic cases in this cohort also exhibited undetectable activity (Figure 2A), suggesting that disease activity is not explained by the level of plasma ADA2 activity.

The subset of DADA2 cases with residual ADA2 plasma activity included both pediatric and adult cases (Table 1). Clinical features of this subgroup were classic manifestations of DADA2, including peripheral vasculitis, ischemic stroke, abdominal aneurysms, and intestinal hemorrhage (Figure 2B–D and Table 1). The presence of residual activity was not an indicator of milder disease, as highlighted by a three-year-old patient with carrier level activity (3.4 U/L) who exhibited diffuse abdominal aneurysms and died of bowel ischemia and hemorrhage (Figure 2D).

Correlations between genetic variants and residual ADA2 activity. Analysis of DADA2 cases with genetic confirmation (n = 173) in the current study revealed 87 unique *ADA2* variants (Figure 2E and Supplemental Table 1). Two missense variants, p.G47R and p.R169Q, were the most common disease-associated variants in the cohort (Figure 2F). These represent the identified founder mutations in South Asian and Northern European populations.[12,20,21] Most of these pathogenic variants were functionally confirmed by our previous studies.[6,17] Impaired production of ADA2 by 12 novel missense variants and loss-of-function variants in the current study was confirmed by expression in 293T cells (Figure 2G). Integration of functional data from all missense and null variants revealed an enrichment of variants with greater residual activity in DADA2 cases with detectable plasma ADA2 activity (Figure 2H and I and Supplemental Table 1). Moreover, 14 of 18 cases with detectable ADA2 activity possessed a hypomorphic variant with greater than 20% residual activity, which was typically in trans with a more detrimental variant (Figure 2J). These data suggest that residual plasma ADA2 activity is associated with the vasculitis/inflammatory phenotype and the presence of hypomorphic ADA2 variants.

Interactions of ADA2 variants in vitro. Because ADA2 is produced as a dimer,[7] we investigated whether interactions between ADA2 mutants may explain the pathogenicity of hypomorphic variants. For cases with residual ADA2 activity and compound heterozygous missense variants, we expressed the variants individually and in tandem, then measured ADA2 activity in the supernatant. We found that co-transfection of pathogenic variants with at least 80% loss of function (ie, p.P251L and p.V458D; highlighted in red) further reduced the residual function of the hypomorphic variants (ie, p.P344S and p.G25C; highlighted in blue), suggesting a potential synergistic effect of select *ADA2* variants (Figure 3A and B and Supplemental Figure 3A).

Next, we co-transfected each variant with wildtype ADA2 to assess a potential dominant-negative effect of individual variants. To increase generalizability, we also analyzed the p.R169Q and p.Y453C variants (most common missense variants after p.G47R).[1,17] We found that co-expression of these and other ADA2 mutants with significant loss of function (≥80%; highlighted in red) partially disrupted the production of wildtype ADA2, as measured by a reduction in secreted ADA2 activity in a dose-dependent manner (Figure 3C and Supplemental Figure 3B).

Most missense ADA2 variants cause protein misfolding and retention in the endoplasmic reticulum.[18] Western blotting under nonreducing conditions revealed that most ADA2 mutants

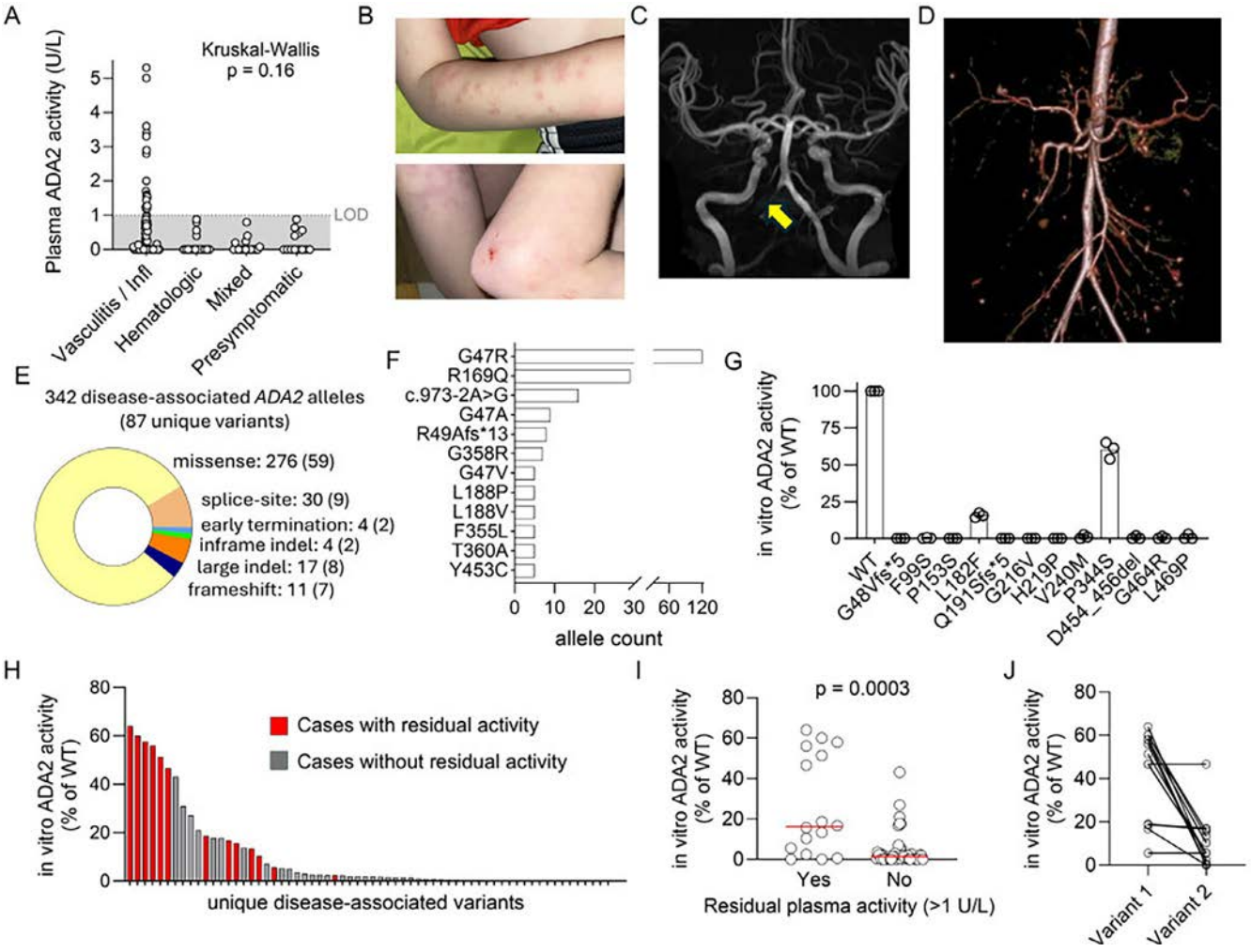


Figure 2. Residual plasma ADA2 activity in patients with DADA2. (A) Plasma ADA2 levels in individuals with DADA2 arranged according to disease phenotype. Vasculitis/ inflammation, n = 144; hematologic phenotype, n = 16; mixed phenotype, n = 19; presymptomatic cases, n = 3. (B) Images of cutaneous vasculitis and livedo racemosa in a three-year-old boy with DADA2. (C) Vertebral artery vasculitis (arrow) in a 49-year-old woman with DADA2 (MRI). (D) Multiple arterial aneurysms in a three-year-old boy with DADA2 who later died because of intestinal hemorrhage and shock (CT). (E) Classification of disease-associated ADA2 variants. The number of unique variants among each group is provided in parentheses. (F) Allelic count of the most common pathogenic ADA2 variants in the current cohort. (G) ADA2 activity of novel *ADA2* variants was determined in the supernatant of transfected 293T cells. (H) ADA2 activity of all missense and loss-of-function *ADA2* variants in the current cohort and (I) comparison of variants from cases with or without residual plasma ADA2 activity. (J) ADA2 activity of paired disease-associated variants from DADA2 cases with residual plasma ADA2 activity. For each variant, the enzymatic activity in the supernatant relative to wildtype ADA2 is provided. Data displayed in (G)–(J) represent the average from three independent experiments. CT, computed tomography; DADA2, deficiency of adenosine deaminase 2; MRI, magnetic resonance imaging; WT, wild type.

formed large protein aggregates (Figure 3D). Mutant protein aggregation required disulfide bonding and was eliminated with the addition of 2-mercaptoethanol (Figure 3D). The propensity to form aggregates correlated inversely with the residual enzymatic activity of these ADA2 mutants (Figure 3E). Co-expression studies further showed that mutant ADA2 (Myc-tagged) aggregated with wildtype ADA2 (HA-tagged), providing a mechanistic explanation for the dominant-negative impact observed earlier (Figure 4A).

To further investigate how disulfide bonding results in aggregate formation, we tested the requirement of the four highly conserved cysteine residues present on ADA2 (Supplemental Figure 4A). Three cysteine residues (at amino acid positions 134, 137, and 159) are located within the putative receptor binding domain, whereas Cys408 is found adjacent to the zinc coordinate site (Supplemental Figure 4B and C). Structural analysis of ADA2 suggests the existence of a disulfide bond between Cys137 and Cys159.[22] Experimentally, all four residues were required for the enzymatic function of ADA2 (Supplemental Figure 4D). Replacement of Cys134, Cys137, or Cys159 with serine led to significant protein aggregation, whereas substitution of Cys408 resulted in minimal aggregation (Supplemental Figure 4E). Unlike most pathogenic missense variants associated

Table 1. DADA2 cases with residual plasma ADA2 activity*

Case ID	ADA2 variants		ADA2 activity, U/L	Age at onset, y	Clinical features
R1	p.F355L	c.1442+2T>C	5.3	8	Fever and recurrent oral ulcers
R2	p.L188V	p.L188V	5	7	PAN, subarachnoid hemorrhage, mononeuritis multiplex, GI hemorrhage
R3	p.P344S	p.P251L	3.4	4	Livedo racemosa, seizures, encephalomyelitis, cutaneous nodules, renal infraction, abdominal aneurysms, intestinal perforation (deceased)
R4	p.F355L	Exon 6 del	3.3	0.5	Fever, abdominal pain, diarrhea, elevated markers of inflammation
R5	p.G47R	p.L188V	2.9	7	Cutaneous PAN, recurrent leg ulcerations, livedo racemosa
R6	p.G47R	p.L188V	2.8	23	Cutaneous PAN, arthralgia, brain aneurysm, hearing loss
R7[a]	p.F355L	p.C408S	2	20	Livedo racemosa, amaurosis fugax, leg ulcers
R8	p.F355L	p.E328K	1.7	2	Livedo racemosa, recurrent fever, polyarthralgia, erythema nodosum, abdominal arterial aneurysms, renal infarction
R9[a]	p.F355L	p.C408S	1.6	12	Livedo racemosa, leg ulcers, digital ulcers, ischemic strokes, vertebral artery vasculitis, labyrinthitis
R10[a]	p.D336G	p.F80Sfs*23	1.6	12	Cutaneous PAN, Raynaud phenomenon, oral ulcers, liver nodules
R11	p.G47R	p.G47R	1.6	12	Livedo reticularis, skin ulcers and necrosis, renal infract, peripheral neuropathy
R12[a]	p.D336G	p.F80Sfs*23	1.5	11	Cutaneous PAN, Raynaud phenomenon, oral ulcers
R13	p.G47A	Exon 7 del	1.4	3	Cutaneous PAN confirmed on biopsy
R14	p.G47R	p.G47R	1.3	19	Recurrent stroke, hearing loss
R15	p.P344S	p.L182P	1.2	13	Livedo reticularis, ischemic stroke, hypertension, elevated markers of inflammation
R16	p.V458D	p.G25C	1.2	15	PAN, recurrent ischemic stroke
R17	p.G47A	p.M436T	1.2	3	Cutaneous PAN confirmed on biopsy, recurrent fever, arthritis
R18	p.G47R	p.P435A	1.2	14	Livedo racemosa, recurrent ischemic strokes

* DADA2, deficiency of adenosine deaminase 2; ID, identification; GI, gastrointestinal; PAN, polyarteritis nodosa.
[a] Sibling cases.

with DADA2, Cys408S did not exhibit significant aggregation or dominant-negative effects (Figure 3C–E).

Cys408 is not implicated to form any intramolecular disulfide bond, making it theoretically accessible for intermolecular interactions. Therefore, we hypothesized that Cys408 is required for the formation of mutant ADA2 aggregates. Indeed, the replacement of Cys408 with serine in cis with the common ADA2 mutants lessened the degree of aggregate formation and partially reversed the dominant-negative effect of these variants on wildtype ADA2 (Figure 4B and C). Taken together, these studies illustrate that mutant ADA2 protein can impose dominant-negative effects and may explain how mutants with significant residual function can act in synergy with more detrimental variants.

DISCUSSION

Quantification of ADA2 enzymatic activity is a rapid and cost-effective method used to evaluate DADA2.[6,11,12,14,23,24] The utility of ADA2 activity as a diagnostic test has not been thoroughly evaluated because of differences among laboratories in ADA2 assay conditions and limitations in patient sample size. In this study, we employed a validated ADA2 assay that has been used extensively for diagnostic purposes in a CLIA certified laboratory. Leveraging the recruitment of 200 DADA2 cases through international collaborations, we assessed the ability of plasma ADA2 activity determined with this assay to discriminate DADA2 cases from carriers and carriers from controls. Although ADA2 activity is highly sensitive and specific in distinguishing DADA2 from other groups, a subset of DADA2 cases exhibited residual activity that extends into the carrier range. These findings indicate that genetic evaluation is necessary for suspected cases with low but detectable plasma ADA2 activity, especially for those with activity below the mean value of the carrier group. Similarly, given the overlap in ADA2 activity in carriers compared with both DADA2 cases and healthy controls, carrier status should be determined genetically.

Our previous analysis of *ADA2* variants in vitro projected that the vasculitis/inflammatory phenotype of DADA2 is associated with missense variants that possess greater residual function, whereas the hematologic phenotype is linked to null variants and missense variants with minimal residual activity.[17] The association of vasculitis/inflammatory phenotype with residual plasma ADA2 activity in the current study provides further support of the proposed genotype-phenotype association, although generalizability is limited by fewer patients with predominantly hematologic

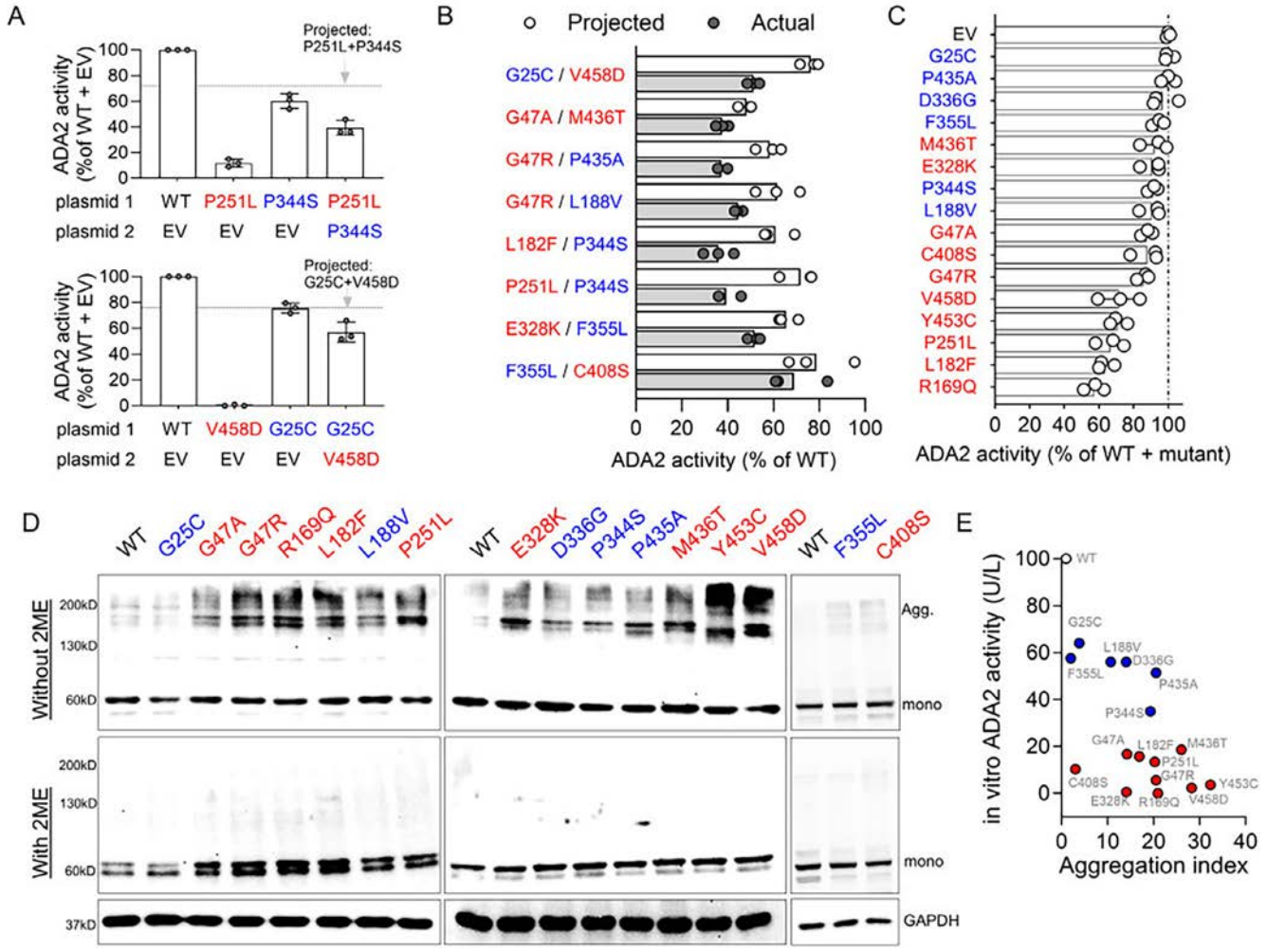


Figure 3. Interactions between pathogenic *ADA2* variants. (A) ADA2 activity in the supernatant of 293T cells transfected with wildtype and mutant ADA2 plasmids. Dotted line indicates the projected level (calculated by adding the mean levels from each mutant ADA2 pair). (B) Projected and actual ADA2 activity from co-transfection of paired ADA2 variants from patients with residual plasma ADA2 activity. (C) ADA2 activity in the supernatant of 293T cells co-transfected with WT ADA2 and the indicated ADA2 variant. Dotted line indicates ADA2 activity from co-transfection of WT ADA2 plasmid with empty vector. A reduction of ADA2 activity levels suggests a dominant-negative effect of the pathogenic variant on WT ADA2. Data in (A)–(C) are combined from three independent experiments. (D) Western blot of whole cell lysates from 293T cells transfected with missense ADA2 variants. Blots with or without 2-mercaptoethanol are displayed. Results are representative of three independent experiments. (E) Correlation between in vitro ADA2 activity and formation of large mutant protein aggregates in transfected cells. Aggregation index represents the ratio of ADA2 aggregate to monomers as quantified by densitometry analysis using ImageJ software. In all panels, red and blue text indicate variants with <20% or ≥20% residual activity, respectively. ADA2, adenosine deaminase 2; EV, empty vector; WT, wildtype.

phenotype. A similar trend was observed in a recent study from the Dutch cohort.[25]

Residual plasma ADA2 activity aligns with the presence of hypomorphic *ADA2* variants. In most cases, the hypomorphic allele is present in trans with a variant that profoundly disrupted ADA2 function. We demonstrate here that some ADA2 variants exert a dominant-negative impact when co-expressed in 293T cells, which is an effect explained mechanistically by the generation of large mutant protein aggregates due to intermolecular disulfide bond formation at a specific cysteine residue. Although replacement of Cys408 can partially abrogate the dominant-negative effect and aggregate formation, this strategy cannot correct the function of mutant ADA2 proteins because the Cys408 mutant is a pathogenic mutation associated with DADA2, and all four cysteine residues are intrinsically required for ADA2 function.

A dominant-negative effect of ADA2 mutants on the wildtype protein raises the question of whether carriers can also develop the pathology of DADA2. This concept of symptomatic carrier secondary to dominant-negative *ADA2* variants was suggested by a recent study.[26] In our collective clinical experience, however, carriers (including siblings, parents, and other relatives of DADA2 cases) are healthy and do not manifest clinical features associated with DADA2. Several carriers with very low plasma ADA2 activity

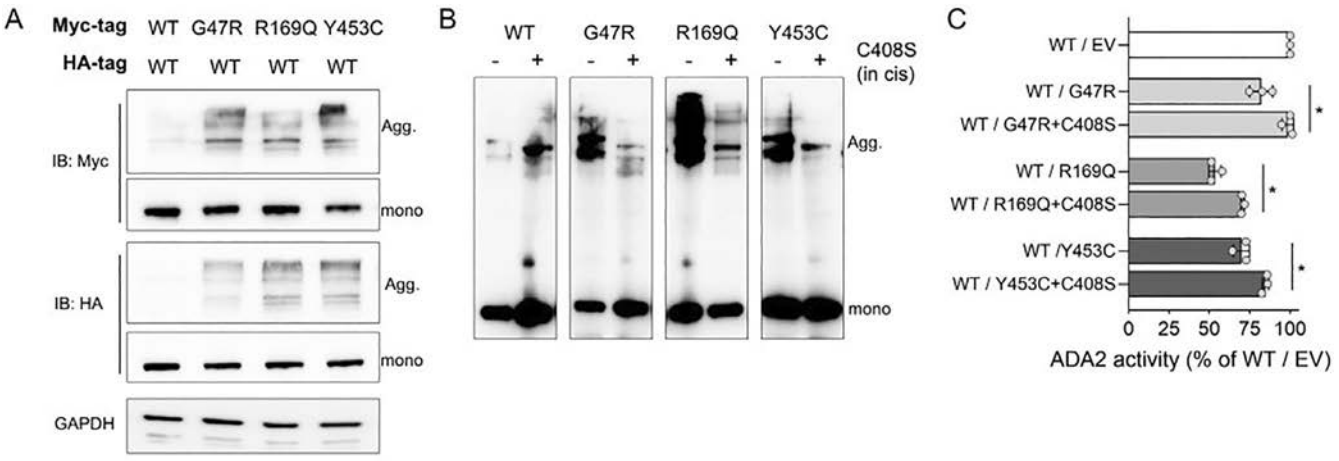


Figure 4. Dominant-negative impact of pathogenic *ADA2* variants mediated by disulfide bonding at Cys408. (A) Western blot of whole cell lysates from 293T cells transfected with Myc-tagged WT or mutant ADA2 and HA-tagged WT ADA2. (B) Western blot of cell lysates and (C) ADA2 activity in the supernatant of 293T cells transfected with mutant ADA2 and the same mutants with C408S substitution in cis. Results in all panels are representative of at least three independent experiments. *p<0.05 by Student's *t*-test. ADA2, adenosine deaminase 2; WT, wildtype.

(<3 U/L) in the current study were healthy individuals. A more systematic comparison of health status, medical history, laboratory findings, and immune parameters in carriers and matched controls is needed to better understand the impact of heterozygous *ADA2* variants. The possibility of unbalanced allelic expression and regulation by noncoding regions should also be considered in future studies.[27]

Our work is limited by an incomplete understanding of the physiologic function of ADA2. Although plasma ADA2 activity serves as an excellent diagnostic biomarker, this activity level does not correlate well with disease activity. Severe vasculitis was noted in patients with residual activity, whereas undetectable activity was found in presymptomatic individuals. These observations, along with the low affinity of ADA2 for adenosine and deoxyadenosine and other considerations, argue against the deaminase function of ADA2 in the peripheral blood as the main determinant of disease.[8] Recent studies suggest a role of ADA2 as a lysosomal enzyme that binds to DNA and deaminates deoxyadenosine nucleotides, but the connection to the pathogenesis of DADA2 has not been elucidated.[10,28]

Another limitation of the study cohort is predominance of DADA2 cases with vasculitis/ inflammation compared with other phenotypes, which is in part attributable to referral bias as well as variable awareness of DADA2 among subspecialities. Data from DADA2 cases were collected from one timepoint; the relationship between residual ADA2 levels and long-term outcome is not clear. Moreover, our analysis of *ADA2* variants and inference of their interactions are based on an overexpression system in cells that do not endogenously express ADA2. Further studies using primary cells are needed to validate these observations.

In conclusion, our study provides a comprehensive evaluation of plasma ADA2 activity as a diagnostic assay. We demonstrate that DADA2 remains possible in cases with some residual activity and illustrate a potential mechanism on how *ADA2* variants interact to synergistically disrupt protein function.

ACKNOWLEDGMENT

We thank the patients and family members for their participation and contribution to our studies.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr Lee confirms that all authors have provided the final approval of the version to be published, and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. Meyts I, Aksentijevich I. Deficiency of adenosine deaminase 2 (DADA2): updates on the phenotype, genetics, pathogenesis, and treatment. J Clin Immunol 2018;38(5):569–578.
2. Kasap Cuceoglu M, Sener S, Batu ED, et al. Systematic review of childhood-onset polyarteritis nodosa and DADA2. Semin Arthritis Rheum 2021;51(3):559–564.
3. Lee PY, Aksentijevich I, Zhou Q. Mechanisms of vascular inflammation in deficiency of adenosine deaminase 2 (DADA2). Semin Immunopathol 2022;44(3):269–280.
4. Navon Elkan P, Pierce SB, Segel R, et al. Mutant adenosine deaminase 2 in a polyarteritis nodosa vasculopathy. N Engl J Med 2014; 370(10):921–931.

5. Zhou Q, Yang D, Ombrello AK, et al. Early-onset stroke and vasculopathy associated with mutations in ADA2. N Engl J Med 2014; 370(10):911–920.
6. Jee H, Huang Z, Baxter S, et al. Comprehensive analysis of ADA2 genetic variants and estimation of carrier frequency driven by a function-based approach. J Allergy Clin Immunol 2022;149(1): 379–387.
7. Zavialov AV, Engström A. Human ADA2 belongs to a new family of growth factors with adenosine deaminase activity. Biochem J 2005; 391(Pt 1):51–57.
8. Tarrant TK, Kelly SJ, Hershfield MS. Elucidating the pathogenesis of adenosine deaminase 2 deficiency: current status and unmet needs. Expert Opin Orphan Drugs 2022;10(3):257–264.
9. Muraoka T, Katsuramaki T, Shiraishi H, et al. Automated enzymatic measurement of adenosine deaminase isoenzyme activities in serum. Anal Biochem 1990;187(2):268–272.
10. Greiner-Tollersrud OK, Krausz M, Boehler V, et al. ADA2 is a lysosomal deoxyadenosine deaminase acting on DNA involved in regulating TLR9-mediated immune sensing of DNA. Cell Rep 2024;43(11): 114899.
11. Caorsi R, Penco F, Grossi A, et al. ADA2 deficiency (DADA2) as an unrecognised cause of early onset polyarteritis nodosa and stroke: a multicentre national study. Ann Rheum Dis 2017;76(10):1648–1656.
12. Özen S, Batu ED, Taşkıran EZ, et al. A monogenic disease with a variety of phenotypes: deficiency of adenosine deaminase 2. J Rheumatol 2020;47(1):117–125.
13. Nanthapisal S, Murphy C, Omoyinmi E, et al. Deficiency of adenosine deaminase type 2: a description of phenotype and genotype in fifteen cases. Arthritis Rheumatol 2016;68(9):2314–2322.
14. Lee PY, Davidson BA, Abraham RS, et al; DADA2 Foundation. Evaluation and management of deficiency of adenosine deaminase 2: an International Consensus Statement. JAMA Netw Open 2023;6(5): e2315894.
15. Lee PY, Huang Z, Hershfield MS, et al. Response to: 'total adenosine deaminase highly correlated with adenosine deaminase 2 activity in serum' by Gao et al. Ann Rheum Dis 2022;81(2):e31.
16. Schnappauf O, Sampaio Moura N, Aksentijevich I, et al; NISC Comparative Sequencing Program; Vasculitis Clinical Research Consortium. Sequence-based screening of patients with idiopathic polyarteritis nodosa, granulomatosis with polyangiitis, and microscopic polyangiitis for deleterious genetic variants in ADA2. Arthritis Rheumatol 2021;73(3):512–519.
17. Lee PY, Kellner ES, Huang Y, et al. Genotype and functional correlates of disease phenotype in deficiency of adenosine deaminase 2 (DADA2). J Allergy Clin Immunol 2020;145(6):1664–1672.e10.
18. Chen L, Mamutova A, Kozlova A, et al. Comparison of disease phenotypes and mechanistic insight on causal variants in patients with DADA2. J Allergy Clin Immunol 2023;152(3):771–782.
19. Lee PY, Schulert GS, Canna SW, et al. Adenosine deaminase 2 as a biomarker of macrophage activation syndrome in systemic juvenile idiopathic arthritis. Ann Rheum Dis 2020;79(2):225–231.
20. Sharma A, Naidu G, Sharma V, et al. Deficiency of adenosine deaminase 2 in adults and children: experience from India. Arthritis Rheumatol 2021;73(2):276–285.
21. Van Montfrans JM, Hartman EA, Braun KP, et al. Phenotypic variability in patients with ADA2 deficiency due to identical homozygous R169Q mutations. Rheumatology (Oxford) 2016;55(5):902–910.
22. Zavialov AV, Yu X, Spillmann D, et al. Structural basis for the growth factor activity of human adenosine deaminase ADA2. J Biol Chem 2010;285(16):12367–12377.
23. Ito M, Nihira H, Izawa K, Yasumi T, et al. Enzyme activity in dried blood spot as a diagnostic tool for adenosine deaminase 2 deficiency. Anal Biochem 2021;628:114292.
24. Cafaro A, Pigliasco F, Barco S, et al. A novel LC-MS/MS-based method for the diagnosis of ADA2 deficiency from dried plasma spot. Molecules 2021;26(18):5707.
25. Andriessen MVE, Legger GE, Bredius RGM, et al. Clinical symptoms, laboratory parameters and long-term follow-up in a national DADA2 cohort. J Clin Immunol 2023;43(7):1581–1596.
26. Wouters M, Ehlers L, Van Eynde W, et al. Dominant negative ADA2 mutations cause ADA2 deficiency in heterozygous carriers. medRxiv. Preprint posted December 9, 2024. doi:10.1101/2024.12.09.24317629
27. Stewart O, Gruber C, Randolph HE, et al. Monoallelic expression can govern penetrance of inborn errors of immunity. Nature 2025; 637(8048):1186–1197.
28. Dong L, Luo W, Maksym S, et al. Adenosine deaminase 2 regulates the activation of the toll-like receptor 9 in response to nucleic acids. Front Med 2024;18(5):814–830.

Arthritis & Rheumatology
Vol. 78, No. 3, March 2026, pp 743–751
DOI 10.1002/art.43401

American College of Rheumatology
Empowering Rheumatology Professionals

Effect of Time to Start of Biologic Therapy on Treatment Response in Childhood Arthritis: Results From the UCAN CAN-DU Cohort

Jelleke B. de Jonge,[1,2,3] Sytze de Roock,[1,2] Dieneke Schonenberg-Meinema,[4] J. Merlijn van den Berg,[4] Deborah A. Marshall,[5,6] Sebastiaan J. Vastert,[1,2,3] Rae S. M. Yeung,[7] Joost F. Swart,[1,2] and Susanne M. Benseler,[6,8,9] on behalf of the UCAN CAN-DU and UCAN CURE consortia

Objective. To estimate the effect of time from symptom onset to start of biologic treatment on achieving inactive arthritis within six months in a cohort of patients with juvenile idiopathic arthritis (JIA).

Methods. The international UCAN CAN-DU study prospectively enrolled patients with JIA across Canada and the Netherlands. A nested cohort study was performed and biologic-naive patients with nonsystemic JIA were included at the start of biologic therapy. The primary outcome was inactive arthritis at six months. Demographics, disease-related parameters, and treatment response were compared using (non)parametric tests among early (time symptom onset to biologic start: 0–6 months), intermediate (7–12 months), and late (13–24 months) treatment groups. A logistic regression model analyzed the effect of time to biologic start on the response at six months, adjusting for active joint count and physician global assessment. A graphical representation of the model was created.

Results. One hundred and thirty children with JIA were included (early: n = 35; intermediate: n = 46; late: n = 49), 66% were female, and the median age at symptom onset was 11.0 years. The proportion of patients that reach inactive arthritis in the early starters (83%) was significantly higher than in late starters (57%). For each month of delay to the start of biologic treatment, the adjusted odds of having active arthritis after six months of therapy was 1.09 (interquartile range: 1.02–1.17, $P = 0.009$).

Conclusion. Early start of biologic therapies in patients with JIA was associated with a higher proportion of patients reaching inactive arthritis within six months, suggesting a window of opportunity to control disease activity.

INTRODUCTION

Juvenile idiopathic arthritis (JIA) is the most common chronic rheumatic disease in children and a significant cause of short- and long-term disability.[1] For the majority of patients, JIA will persist into adulthood, leading to a diminished quality of life.[2] By definition, arthritis needs to persist for more than six weeks before the diagnosis can be made.[3] The diagnostic delay is usually much longer with a median interval of 20 weeks from symptom onset to first assessment by a pediatric rheumatology team.[4] Another

The preliminary data of this work were previously presented during the 31st European Paediatric Rheumatology Congress, Gothenburg, Sweden, September 11–14, 2024.

This study was undertaken as part of the UCAN CAN-DU and UCAN CURE consortia, which were supported by the Canadian Institutes of Health Research (grant 381280), Genome Canada (grant OGI-150), Genome Alberta (Canada), Ontario Genomics (Canada), ZonMw (the Netherlands; grant 848.006.001), the Dutch Arthritis Foundation (ReumaNederland), and Long-standing Support Program LLP10 of the Dutch Arthritis Foundation.

[1]Division of Pediatrics, Department of Pediatric Rheumatology, Wilhelmina Children's Hospital, University Medical Center Utrecht, the Netherlands; [2]Faculty of Medicine, Utrecht University, the Netherlands; [3]Center for Translational Immunology, University Medical Center Utrecht, Utrecht University, the Netherlands; [4]Department of Pediatric Immunology, Rheumatology and Infectious Diseases, Emma Children's Hospital, Amsterdam University Medical Centers, University of Amsterdam, the Netherlands; [5]Departments of Community Health Services and Medicine, McCaig Institute for Bone and Joint Health, O'Brien Institute for Public Health, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada; [6]Alberta Children's Hospital Research Institute, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada; [7]Division of Rheumatology, Department of Paediatrics, Immunology and Institute of Medical Science, The Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada; [8]Division of Rheumatology, Department of Pediatrics, Alberta Children's Hospital, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada; [9]Rheumatology, Paediatrics, Children's Health Ireland, Dublin, Ireland.

De Jonge and Dr de Roock are co-first authors and contributed equally to this work. Drs Swart and Benseler are co-senior authors and contributed equally to this work.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/art.43401).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/art.43401.

Address correspondence via email to Jelleke B. de Jonge, MSc, at j.b.dejonge-10@umcutrecht.nl; or Joost F. Swart, MD, PhD, at j.f.swart@umcutrecht.nl.

Submitted for publication March 7, 2025; accepted in revised form August 29, 2025.

barrier to fast effective therapy are the current JIA treatment guidelines that are anchored in a "step-up approach."[5]

JIA therapy commonly follows the sequence of nonsteroidal anti-inflammatory drugs, intra-articular steroid injections, conventional synthetic disease-modifying antirheumatic drugs (csDMARDs), and, as a final step, biologic therapies. Since their introduction more than 25 years ago, the dogma prevailed that biologic treatment is to be started only after the patient failed to respond to csDMARD, an approach mainly driven by a difference in drug costs. Methotrexate (MTX) is the most frequently used csDMARD and controls arthritis and induces inactive disease in only around 30% of patients with JIA at two years.[6] More importantly, MTX is poorly tolerated. Up to 50% of patients receiving MTX mention nausea and discomfort,[7] amplifying the suffering of children living with arthritis.

This mandatory step-up approach causes a long trial-and-error period between the diagnosis of JIA and the start of biologics, which is the most effective therapy for the majority of patients with JIA. Exposure to biologic agents is relatively uncommon in the first year of follow-up; however, a consistent increase is found at both the three- and five-year time points in both cumulative exposure and point prevalence use, such that at five years, 42% of the cohort had been exposed to a biologic.[8] Previous studies have focused on the start time of biologic therapy by comparing a cohort of patients following different treatment strategies, such as the step-up approach and immediate biologic start.[9–11] These studies showed that an early start of biologic therapy after diagnosis can alter the disease trajectory compared to patients with JIA starting on csDMARDs.[9–11] From a biologic perspective, the time between inflammation with subsequent symptom onset and the initiation of biologic therapy might be more relevant than the time between diagnosis and biologic therapy.

There are convincing data that there is a window of opportunity in the systemic form of JIA, in which an early start of a biologic leads to successful discontinuation of therapy in the majority of patients.[12] The hypothesis that there is a limited period since onset in which biologic therapy can be most successful has been postulated for nonsystemic JIA as well.[13] This hypothesis is that the inflammation increases progressively in a step-wise manner, reflecting the accumulation of joints that have acquired durable changes, placing them at a higher risk of subsequent flare or persistent active disease. Effective treatment should arrest this step-wise escalation of risk, providing a rolling window of opportunity to prevent further joint accumulation. Proof for this hypothesis has not yet been delivered but is only anecdotal.

Optimal timing of effective treatment is beneficial for clinical outcomes, patients' quality of life, and patient and societal burden in the short- and long-term. Therefore, the hypothesis that a longer time between symptom onset and initiation of biologic therapy negatively affects the probability of reaching inactive arthritis is formally tested here in an international, prospectively observed cohort of patients with JIA. The aims of the study were (1) to describe a cohort of patients with nonsystemic JIA starting with biologic treatment, (2) to compare the proportion of patients that reach inactive arthritis in early, intermediate, and late biologic treatment groups, and (3) to analyze the impact of time from symptom onset to start of biologic treatment on early outcome.

METHODS

The international multicenter UCAN CAN-DU study prospectively enrolled children with arthritis across Canada and the Netherlands (see Appendix A for a list of UCAN CAN-DU and UCAN CURE consortia members). All patients participating in the UCAN CAN-DU study provided informed consent to participate in the study. Between April 2019 and November 2024, a nested cohort study was performed within the UCAN CAN-DU cohort, which included patients from eight centers in Canada and six centers in the Netherlands. Patients with JIA were included in this nested study if they (1) started biologic treatment for arthritis (active joint count [AJC] > 0) within 24 months after symptom onset, (2) were biologic-naive at the start of biologic therapy, and (3) had a six-month follow-up visit after the start of biologic treatment (six-month follow-up visits are defined as a follow-up visit within four to eight months by the UCAN CAN-DU study protocol). Patients with systemic JIA (sJIA) were excluded from this study. Patients with a six-month follow-up visit were compared to those not included in the study because their follow-up visit was outside of the defined timeframe. Patients were comparable regarding the disease status at the start of biologic therapy (Supplementary Table 1). Physician- and patient/family-derived data were prospectively captured in the designated UCAN eHealth platform. The study was approved by the Institutional Research Ethical Board of the University Medical Center Utrecht (no. 18-474) for the Netherlands and by the Conjoint Health Research Ethics Board, University of Calgary for Canada (REB17-1563).

Demographic, clinical data, and global measures. Demographic characteristics included age at symptom onset, biologic sex, and country of residence. International League of Associations for Rheumatology JIA subtypes were reported by the treating physicians as diagnostic criteria. In the analysis, extended oligoarticular JIA and polyarticular JIA were collectively referred to as polyarticular course JIA, whereas (persistent) oligoarticular JIA were referred to as oligoarticular course JIA. Clinical variables included JIA-related articular and extra-articular manifestations, including the number and distribution of affected joints and enthesitis and overall AJC. Active joints were assessed in 83 joints (carpometacarpal I joints [2×], distal interphalangeal joints in the foot [8×], and sacroiliac joints [2×]) in addition to the

clinical Juvenile Arthritis Disease Activity Score (cJADAS) 71 joints,[14] and a joint was considered active if it was swollen or had both signs of tenderness and limited range of motion. In case of doubt (limitation and/or pain), imaging was performed on the joints under supervision of the treating physician. All global measures were captured on digital visual analog scales ranging from 0 to 10 with 0.5-point increments, including Physician Global Assessment (PhGA) of disease activity and Patient/Parent Global Assessment (PPGA). cJADAS10 scores were calculated based on the AJC, PhGA, and PPGA. Variables were collected, and assessments were performed at all visits.

Biologic therapy and stratification. Treatment data included the medication type, dose and route, and start and stop date of all therapies, including biologic and conventional JIA treatments. Time to biologic treatment was defined as time from symptom onset to the start date of the first biologic treatment. Based on this, patients were stratified into three different treatment groups: (1) early treatment group, in which the time from symptom onset to biologic therapy was 0 to 6 months, (2) intermediate treatment group (7–12 months), and (3) late treatment group (13–24 months) (Supplementary Figure 1).

Outcome. The outcome of biologic therapy was inactive arthritis defined as no evidence of active joints (AJC = 0) at six months after the start of biologic treatment. An AJC of zero was chosen as an outcome measurement, as this frequently guides decisions around effectiveness in real life. To account for extra-articular manifestations, an AJC + PhGA score of <1 was used as a secondary outcome measurement.

Analysis. Descriptive statistics summarize the patient characteristics. Comparisons were made using parametric and non-parametric tests as appropriate, including Pearson's chi-square test, Fisher's exact test, and Kruskal–Wallis rank sum test. To test whether patients were comparable when treatment was initiated during their first hospital visit, comparisons were made using clinical data from available diagnosis visits in which patients were still treatment-naive. Moreover, to show that patients had a comparable chance of starting a biologic early or later, the baseline characteristics age at symptom onset, AJC, and PhGA were modeled to predict the time to biologic treatment start, and the predicted and observed time to biologic treatment start were compared. Next, a logistic regression model was constructed to assess the effect of time between symptom onset and biologic start on the primary outcome, adjusted for other variables. Variables, measured at the start of biologic therapy, with a statistically relevant ($P < 0.05$) effect on the primary outcome were included in the model, as well as variables with clinical relevance (prior use of csDMARD). A potential interaction between AJC and PhGA was examined to account for their possible relationship given that PhGA is partially dependent on AJC but also on the presence of enthesitis, uveitis, and other patient characteristics. cJADAS10 was not included in the model because this composite score simply adds the information from the variables PhGA, PPGA, and AJC that are part of the univariate analysis. In addition, differences across centers were not accounted for in the models, as these partially underlie differences in treatment. A continuous variable for time between symptom onset and biologic start was used. To identify the optimal model, a backward step-wise selection of variables was performed based on Akaike's information criterion and variance inflation factors were calculated to check for collinearity in the model. Models were compared using the likelihood ratio test. All analyses were performed in R version 4.4.1 software[15] using the package tidyverse (version 2.0.0)[16] and lme4 (version 1.1.35.3).[17]

RESULTS

A total of 130 children was included, 86 were female patients (66%), and the median age at symptom onset was 11.0 years (interquartile range [IQR]: 3.8–13.4). Of these patients, 35 (26.9%) were in the early treatment group, 46 (35.4%) in the intermediate treatment group, and 49 (37.7%) in the late treatment group. Overall, 98 (75%) of the patients were treated in Dutch centers, with a significantly higher proportion of the patients treated in Dutch centers in the early treatment group (94%) compared to the intermediate (72%) and late (65%) treatment groups. A total of 83 children (64%) used csDMARDs before the start of biologic therapy, with a significantly lower proportion of patients using csDMARDs in the early treatment group (37%) compared to the intermediate (78%) and late (69%) treatment groups (Pearson's chi-square test, $P < 0.001$). The use of csDMARDs during biologic treatment was comparable between treatment groups. The predominantly targeted biologic pathway was tumor necrosis factor α (97%) across all treatment groups; other targeted biologic pathways were interleukin-6 (2%, tocilizumab) and JAK (1%, tofacitinib). Moreover, age at symptom onset, percentage of female patients, AJC, global assessment of both physicians and patients/parents, cJADAS10, and JIA subtype were consistent across the groups (Table 1). Complete data were available in all 130 patients for all variables, except for PPGA and cJADAS10, for which 68% was complete.

Sensitivity analysis at diagnosis visit. A total of 91 of the 130 patients had an available diagnosis visit on which a subanalysis was performed to check whether the early, intermediate, and late treatment groups were comparable during the visits in which treatment was initiated. This analysis showed that at diagnosis visits, in which all patients were still treatment-naive for csDMARDs and biologics, the variables AJC, PhGA, PPGA, and JIA subtype were all comparable across the three treatment groups (Table 2). In addition, only weak correlation was found

Table 1. Characteristics of children with JIA at the start of biologic treatment stratified by the timing of treatment since symptom onset in three treatment groups*

Characteristic	Overall n = 130	Start of biologic treatment			*P* value
		Early (0–6 months) n = 35	Intermediate (7–12 months) n = 46	Late (13–24 months) n = 49	
Age at onset	11.0 (3.8, 13.4)	10.2 (3.5, 14.2)	10.1 (3.5, 12.9)	11.4 (5.3, 13.5)	0.5[c]
Sex (female)	86 (66)	24 (69)	33 (72)	29 (59)	0.4[d]
Country of residence (the Netherlands)	98 (75)	33 (94)	33 (72)	32 (65)	0.008[a,d]
AJC	4.0 (2.0, 7.0)	5.0 (2.0, 10.5)	4.5 (3.0, 8.0)	3.0 (2.0, 6.0)	0.12[c]
PhGA	3.4 (2.0, 5.0)	4.5 (2.0, 6.0)	3.0 (2.0, 4.9)	3.5 (2.5, 4.5)	0.3[c]
PPGA[b]	4.8 (2.4, 6.8)	5.5 (2.1, 7.0)	5.5 (3.2, 7.0)	4.0 (2.2, 6.0)	0.3[c]
cJADAS10[b]	12.5 (8.9, 17.0)	15.1 (9.5, 18.0)	14.0 (10.8, 18.0)	11.0 (7.3, 14.3)	0.058[c]
Months to biologic treatment	9.0 (5.9, 15.4)	4.0 (2.8, 5.1)	8.6 (7.3, 9.5)	16.8 (14.7, 19.0)	<0.001[a,c]
csDMARDs use before biologics	83 (64)	13 (37)	36 (78)	34 (69)	<0.001[a,d]
csDMARDs continued or started	95 (73)	24 (69)	36 (78)	35 (71)	0.6[d]
Biologic therapy					>0.9[e]
TNFα inhibition	126 (97)	34 (97)	44 (96)	48 (98)	
IL-6 inhibition	3 (2.3)	1 (2.9)	1 (2.2)	1 (2.0)	
JAK inhibition	1 (0.8)	0 (0)	1 (2.2)	0 (0)	
JIA subtype					0.6[e]
Enthesitis arthritis	31 (24)	6 (17)	11 (24)	14 (29)	
Oligoarticular course JIA	43 (33)	14 (40)	11 (24)	18 (37)	
Polyarticular course JIA	50 (38)	13 (37)	21 (46)	16 (33)	
Psoriatic arthritis	3 (2)	1 (3)	2 (4)	0 (0)	
Undifferentiated JIA	3 (2)	1 (3)	1 (2)	1 (2)	

* Values presented as median (interquartile range) or n (%). AJC, Active Joint Count; cJADAS10, clinical Juvenile Arthritis Disease Activity Score 10 joints; csDMARD, conventional synthetic Disease-Modifying Antirheumatic Drug; IL-6, Interleukin-6; JIA, Juvenile Idiopathic Arthritis; PhGA, Physician Global Assessment; PPGA, Patient/Parent Global Assessment; TNFα, Tumor Necrosis Factor α.
[a] Statistically significant at $P < 0.05$.
[b] Data complete for 68% of the patients.
[c] Kruskal-Wallis rank sum test.
[d] Pearon's chi-square test.
[e] Fisher's exact test.

between the predicted "months to biologic start" (based on a linear model including the variables age at symptom onset, AJC, and PhGA at first hospital visit) and the observed "months to biologic start," indicating that the variables of treatment-naive patients could not predict when a biologic was initiated (Supplementary Figure 2).

Effect of time to biologic treatment on reaching inactive arthritis at six months. The early treatment group received biologic treatment at a median of 4 months (IQR: 2.8–5.1), the intermediate group at 8.6 months (IQR: 7.3–9.5), and the late treatment group at 16.8 months (IQR: 14.7–19.0) (Table 1). The proportion of patients reaching inactive arthritis at six months after start of biologic treatment was the highest in the early treatment group (83%) compared to the intermediate (67%, Pearson's chi-square test, $P = 0.12$) and late treatment group (57%, Pearson's chi-square test, $P = 0.013$, Figure 1). Of the 88 patients with inactive arthritis, 4 still had active uveitis (physician reported) after six months of biologic treatment. To account for extra-articular manifestations, the proportion of patients reaching AJC + PhGA <1 was measured, resulting in comparable results (83% of the patients reaching AJC + PhGA <1 in the early treatment group, 57% in the intermediate treatment group, and 53% in the late treatment group; Supplementary Figure 3).

Multivariate logistic regression model for active arthritis at six months. The time between symptom onset and the start of biologic treatment had a significant effect on the probability of reaching inactive arthritis at six months after treatment start (Figure 1 and Table 3). To assess the corrected effect of time between symptom onset and biologic start on the probability of reaching inactive arthritis six months after starting biologic treatment, a multivariate logistic regression model was created. Univariate analysis identified lower AJC and PhGA at the start of biologic therapy as significantly associated with reaching inactive arthritis (Wilcoxon rank sum test, $P < 0.05$, Table 3). Both were included in the initial model with an interaction term between these variables to account for nonindependence. Prior treatment with csDMARDs, although not statistically significant in the univariate analysis (Table 3), was also included in the initial multivariate model because of its possible effect on subsequent treatment response. After backward step-wise selection, the model included AJC interacting with PhGA next to the time between symptom onset and biologic start. Variance inflation factor

Table 2. Comparison of JIA patient- and disease-related factors at the treatment-naive diagnosis visit, stratified by timing of treatment since symptom onset in three treatment groups*

Characteristic at diagnosis visit	Overall n = 91	Start of biologic treatment: Early (0–6 months) n = 27	Intermediate (7–12 months) n = 28	Late (13–24 months) n = 36	*P* value
Age at onset	11.4 (4.9, 13.4)	10.2 (3.6, 13.9)	10.9 (4.3, 12.9)	11.6 (9.0, 13.5)	0.7[d]
Sex (female)	56 (62)	18 (67)	17 (61)	21 (58)	0.8[e]
Country of residence (the Netherlands)	72 (79)	26 (96)	24 (86)	22 (61)	0.002[a,e]
AJC	4.0 (2.0, 7.5)	4.0 (2.0, 11.0)	4.5 (2.0, 7.0)	2.5 (1.8, 4.3)	0.4[d]
PhGA	4.0 (3.0, 6.0)	3.5 (2.6, 6.8)	4.0 (2.8, 5.0)	3.8 (3.0, 5.3)	>0.9[d]
PPGA[b]	5.0 (3.6, 7.0)	4.5 (1.8, 6.8)	5.5 (4.8, 6.7)	4.6 (3.9, 6.7)	0.6[d]
cJADAS10[b]	13.5 (9.4, 17.3)	15.0 (9.6, 17.5)	13.6 (11.0, 17.2)	11.5 (8.5, 16.8)	0.7[d]
JIA subtype[c]					0.4[f]
Enthesitis arthritis	20 (25)	3 (14)	7 (28)	10 (29)	
Oligoarticular course JIA	29 (36)	6 (27)	8 (32)	15 (44)	
Polyarticular course JIA	27 (33)	11 (50)	8 (32)	8 (24)	
Psoriatic arthritis	2 (2)	1 (5)	1 (4)	0 (0)	
Undifferentiated JIA	3 (4)	1 (5)	1 (4)	1 (3)	

* Patient characteristics were measured at the diagnosis visit (first available hospital visit, treatment-naive for csDMARDs and biologics). Values presented as median (interquartile range) or n (%). AJC, Active Joint Count; cJADAS10, clinical Juvenile Arthritis Disease Activity Score 10 joints; csDMARD, conventional synthetic Disease-Modifying Antirheumatic Drug; JIA, Juvenile Idiopathic Arthritis; PhGA, Physician Global Assessment; PPGA, Patient/Parent Global Assessment.
[a] Statistically significant at $P < 0.05$.
[b] Data complete for 58%.
[c] Data complete for 89% (formal JIA subtype yet unknown at first hospital visit).
[d] Kruskal-Wallis rank sum test.
[e] Pearson's chi-square test.
[f] Fisher's exact test.

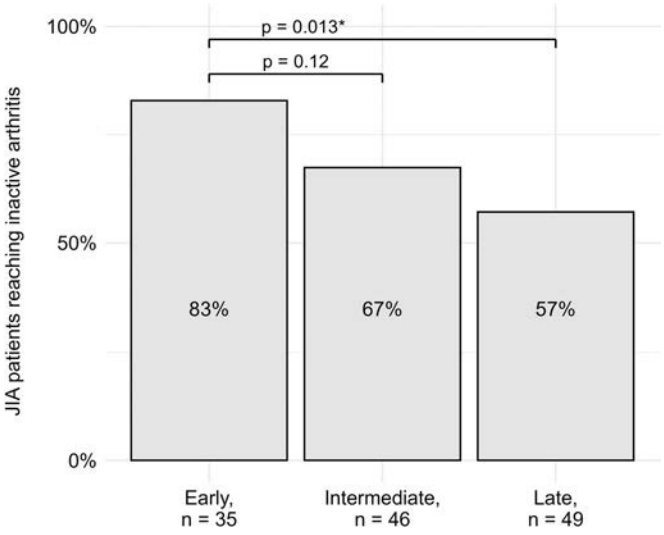


Figure 1. Proportion of patients with JIA reaching inactive arthritis (AJC = 0) after six months of early biologic treatment compared to intermediate and late biologic treatment. Proportion of patients with JIA reaching inactive arthritis after six months of biologic treatment across three treatment groups: early treatment group (time from symptom onset to biologic start: 0–6 months), intermediate treatment group (7–12 months), and late treatment group (13–24 months). Inactive arthritis was defined as having an AJC of zero at six months. The statistical significance of the observed differences was tested with the chi-square test. *Statistically significant at $P < 0.05$. AJC, Active Joint Count; JIA, Juvenile Idiopathic Arthritis.

calculations showed that collinearity was acceptable in the model without the interaction term. Therefore, the final model included AJC and PhGA next to time between symptom onset and biologic start.

Although PPGA at the start of biologic treatment was not significant in the univariate analysis, PPGA at the start of biologic treatment tended to be lower for children reaching inactive arthritis. For this reason, a separate model was tested with PPGA as an extra variable next to AJC, PhGA, and time to biologic treatment. In this analysis, only those patients for whom PPGA was available were included (89 of 130 patients, 29 with active and 60 with inactive arthritis, corresponding to 69% and 68% of all patients with inactive and active arthritis, respectively). The inclusion of PPGA next to AJC, PhGA, and time to biologic treatment did show similar results compared to the model without PPGA.

Visualization of risk model. The model shows that active arthritis at six months was associated with time between symptom onset and biologic start after adjusting for AJC and PhGA (adjusted odds ratio per month: 1.09 [confidence interval 1.02–1.17]; Table 4). Odds ratios for AJC and PhGA at the start of treatment are shown in Table 4. Similarly, in a separate model using AJC and PhGA <1 as the outcome, comparable associations with time between symptom onset and biologic start were observed after adjusting for AJC and PhGA (Supplementary Table 2). To

Table 3. Comparison of patient-, disease-, and treatment-related factors at the start of biologic treatment between patients with inactive and active arthritis six months after biologic treatment in JIA*

	Active arthritis (AJC ≥ 1) n = 42	Inactive arthritis (AJC = 0) n = 88	*P* value
Age at onset	11.4 (4.1, 13.7)	10.3 (3.7, 13.4)	0.3[c]
Sex (female)	28 (67)	58 (66)	>0.9[d]
Country of residence (the Netherlands)	30 (71)	68 (77)	0.5[d]
AJC	6.0 (3.0, 9.0)	3.5 (2.0, 6.0)	0.013[a,c]
PhGA	4.00 (3.00, 5.58)	3.00 (2.00, 4.63)	0.018[a,c]
PPGA[b]	6.00 (4.00, 7.00)	4.00 (1.69, 6.43)	0.077[c]
Months to biologic start	11.7 (7.3, 16.6)	8.8 (5.2, 13.5)	0.034[a,c]
csDMARDs use before biologics	29 (69)	54 (61)	0.4[d]
csDMARDs continued or started	30 (71)	65 (74)	0.8[d]
Biologic therapy			>0.9[e]
TNFa inhibition	41 (98)	85 (97)	
IL-6 inhibition	1 (2.4)	2 (2.3)	
JAK inhibition	0 (0)	1 (1.1)	
JIA subtype			0.8[e]
Enthesitis arthritis	10 (24)	21 (24)	
Oligoarticular course JIA	13 (31)	30 (34)	
Polyarticular course JIA	16 (38)	34 (39)	
Psoriatic arthritis	2 (4.8)	1 (1.1)	
Undifferentiated JIA	1 (2.4)	2 (2.3)	

* Patient characteristics were measured at the start of biologic treatment. Values presented as median (interquartile range) and n (%). AJC, Active Joint Count; csDMARD, conventional synthetic Disease-Modifying Antirheumatic Drug; IL-6, Interleukin-6; JIA, Juvenile Idiopathic Arthritis; PhGA, Physician Global Assessment; PPGA, Patient/Parent Global Assessment; TNFα, Tumor Necrosis Factor α.
[a] Statistically significant at $P < 0.05$.
[b] Data complete for 68% of the patients.
[c] Wilcoxon rank sum test.
[d] Pearson's chi-square test.
[e] Fisher's exact test.

visualize the effect of time between symptom onset and biologic start adjusted for AJC and PhGA, a graphical representation is shown in Figure 2. The graphical representation of the model shows how the chance of reaching inactive arthritis at six months decreases with time for patients with more extreme values for AJC and PhGA at the start of biologic treatment (Figure 2).

DISCUSSION

This study provides real-life evidence that the delay in biologic treatment negatively impacts the treatment response in children with nonsystemic JIA. The percentage of children with inactive arthritis at six months after the start of biologics decreased dramatically from 83% to 57% comparing early versus late biologic treatment starters. Moreover, the regression model revealed that for each month of delay to the start of biologic treatment, the odds for still having active arthritis after six months of treatment increases by factor 1.09. These results suggest a window of opportunity to control nonsystemic JIA in which the biologic therapy is initiated early after symptom onset.

Biologic therapies are highly effective in JIA. This study demonstrated that even in real-life settings, inactive arthritis can be achieved in more than 80% of children across all nonsystemic disease subtypes. These findings mirror findings in patients with sJIA, in whom early start of the interleukin-1 inhibitor anakinra resulted in rapid and sustained control of disease activity.[18] In contrast, delayed starters of biologics in sJIA had an eight-fold higher risk of being a nonresponder.[18] Achieving these results

Table 4. Coefficients table of the multivariate logistic regression model to estimate the effect of time to biologic start in months on having active arthritis at six months of treatment in Juvenile Idiopathic Arthritis corrected for active joint count and the physician global assessment*

Variables at the start of biologic therapy	Crude/adjusted	Estimate	SE	*p*-value	OR	CI (95%)
Time to biologic start (per month)	Crude	0.069	0.032	0.032	1.07	1.01–1.14
Time to biologic start per month)	Adjusted	0.090	0.035	0.0089	1.09	1.02–1.17
AJC	Adjusted	0.021	0.027	0.438	1.02	0.97–1.08
PhGA	Adjusted	0.185	0.109	0.089	1.20	0.97–1.49

* Odds ratio for failing to reach clinical inactive arthritis at six months after the start of biologic treatment.
Abbreviations: AJC, Active Joint Count; CI, Confidence Interval; OR, Odds Ratio; PhGA, Physician Global Assessment; SE, Standard Error.

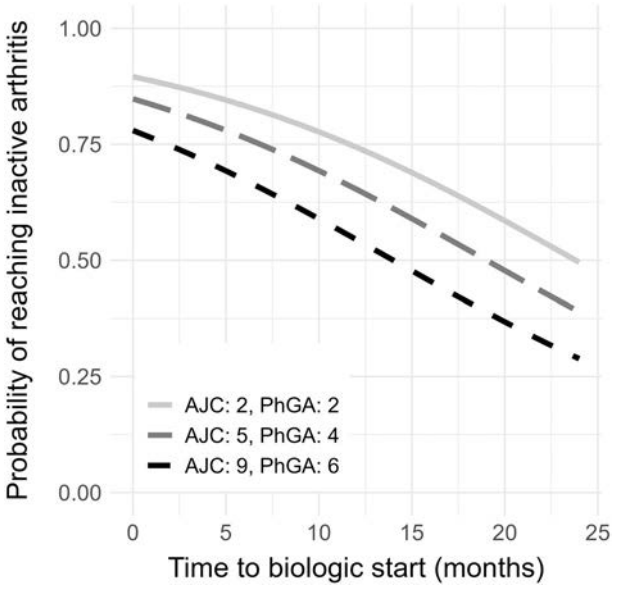


Figure 2. Visualization of the impact of AJC, PhGA at the start of biologic therapy, and time from symptom onset to biologic start on the probability of reaching inactive arthritis (AJC = 0) after six months of biologic treatment for patients with JIA. Effect of time to biologic start on the probability of reaching inactive arthritis in patients with JIA after six months of biologic therapy. Values for AJC and PhGA at the start of biologic therapy were based on the more extreme phenotypes within the interquartile range of both groups. Values for AJC, PhGA, and time were entered into the described model for visualization of the effect of these variables on the probability of response. The calculated probability of response is depicted. AJC, Active Joint Count; JIA, Juvenile Idiopathic Arthritis; PhGA, Physician Global Assessment.

required a paradigm change—a switch from step-up therapy approaches from conventional nonsteroidal anti-inflammatory drugs to csDMARDs to finally biologics for those with polyarticular disease course of sJIA, commonly also viewed as trial-and-error strategies—to the early use of biologic therapies for patients with sJIA, the most severe subtype of JIA.

The potential existence of a window of opportunity in patients with rheumatic arthritis has been described previously.[13,19,20] According to this theory, the disease becomes more difficult to treat when the disease duration is longer because the disease progresses irreversibly. In adults with rheumatic arthritis, early adalimumab start was more successful than the treatment of patients with long disease duration.[21] In a randomized controlled trial in 85 patients with polyarticular JIA, the chance of reaching clinically inactive disease at six months increased by 32% for each month earlier that aggressive treatment (MTX with etanercept and prednisolone) was started after disease onset.[22] A three-arm, single-blinded, treat-to-target study in patients with nonsystemic JIA did not find different remission rates between the treatment groups that started with etanercept directly after diagnosis compared to the two groups starting on csDMARD with or without prednisolone. However, the time from symptom onset to biologic was not investigated. Moreover, this study showed that the majority of patients in all three groups ended up with biologic treatment during the 24-month study because the goal of inactive disease was not reached without it.[23] In addition, patients with nonsystemic JIA with an early biologic treatment are more likely to be in drug-free remission in adulthood compared to patients with a late biologic treatment start (more than five years).[24,25]

Our findings support the theory of a window of opportunity in patients with nonsystemic JIA as well, showing that for patients who will start a biologic, an early start of the therapy is beneficial for the patient's treatment response at least in the first six months after the start of biologic treatment. The study provides evidence that the treatment of patients with nonsystemic JIA may need a similar paradigm shift.

There are several limitations to the study. Our study (with prospectively collected data) cannot rule out that the more severe cases of JIA are overrepresented in the early biologic treatment group. It is not unlikely that at the first outpatient clinic visit of a patient with severe JIA, a physician decides to ignore csDMARDs and start a biologic immediately, whereas less severe patients will start on csDMARDs and will only receive a biologic after several months. To test whether the patients differed at the moment of first therapy choice, a subanalysis was conducted on patients from whom a treatment-naive (both csDMARDs and biologics) diagnosis visit was available. This analysis showed that the early, intermediate, and late treatment groups were comparable in disease activity (AJC, PhGA, PPGA, and cJADAS10), JIA subtype, age, and biologic sex when the first treatment was initiated at the hospital. This suggests that patients in all three groups had a comparable probability for early start of biologic treatment.

Secondly, we did correct for prior csDMARD use in our analysis, but we did not consider the exact duration of csDMARD therapy or the time between csDMARD therapy start and escalation to biologic treatment. According to the theory regarding the window of opportunity, early treatment with any anti-inflammatory agent is important to prevent further disease progress. More research is needed to further study the effect of prior csDMARD use on the efficacy of biologic therapy in patients with nonsystemic JIA.

In conclusion, this study shows that the time from symptom onset to biologic treatment negatively affects the treatment response in patients with nonsystemic JIA. These findings support the theory of a window of opportunity for successful timing of biologic treatment in patients with active JIA. The treatment of patients with JIA will be improved when patients who are in need of biologics could start early with biologic treatment without losing time on other treatment strategies first.

ACKNOWLEDGMENTS

The authors thank all participating patients and parents. In addition, the authors thank all the people who are part of the UCAN CAN-DU and UCAN CURE consortia.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, de Jonge confirms that all authors have provided the final approval of the version to be published and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. Ravelli A, Martini A. Juvenile idiopathic arthritis. Lancet 2007; 369(9563):767–778.
2. Chhabra A, Robinson C, Houghton K, et al. Long-term outcomes and disease course of children with juvenile idiopathic arthritis in the ReACCh-Out cohort: a two-centre experience. Rheumatology (Oxford) 2020;59(12):3727–3730.
3. Petty RE, Southwood TR, Manners P, et al; International League of Associations for Rheumatology. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. J Rheumatol 2004;31(2):390–392.
4. Foster HE, Eltringham MS, Kay LJ, et al. Delay in access to appropriate care for children presenting with musculoskeletal symptoms and ultimately diagnosed with juvenile idiopathic arthritis. Arthritis Rheum 2007;57(6):921–927.
5. Beukelman T, Patkar NM, Saag KG, et al. 2011 American College of Rheumatology recommendations for the treatment of juvenile idiopathic arthritis: initiation and safety monitoring of therapeutic agents for the treatment of arthritis and systemic features. Arthritis Care Res (Hoboken) 2011;63(4):465–482.
6. Raab A, Kallinich T, Huscher D, et al. Outcome of children with oligoarticular juvenile idiopathic arthritis compared to polyarthritis on methotrexate- data of the German BIKER registry. Pediatr Rheumatol Online J 2021;19(1):41.
7. Bulatović M, Heijstek MW, Verkaaik M, et al. High prevalence of methotrexate intolerance in juvenile idiopathic arthritis: development and validation of a methotrexate intolerance severity score. Arthritis Rheum 2011;63(7):2007–2013.
8. Tiller G, Hernandez BL, Buckle J, et al. Three- and five-year outcomes of an inception cohort of Australian children with juvenile idiopathic arthritis. Int J Rheum Dis 2024;27(5):e15189.
9. Ong MS, Ringold S, Kimura Y, et al; CARRA Registry Investigators. Improved disease course associated with early initiation of biologics in polyarticular juvenile idiopathic arthritis: trajectory analysis of a childhood arthritis and rheumatology research alliance consensus treatment plans study. Arthritis Rheumatol 2021;73(10):1910–1920.
10. Kimura Y, Schanberg LE, Tomlinson GA, et al; CARRA STOP-JIA Investigators. Optimizing the start time of biologics in polyarticular juvenile idiopathic arthritis: a comparative effectiveness study of childhood arthritis and rheumatology research alliance consensus treatment plans. Arthritis Rheumatol 2021;73(10):1898–1909.
11. Ringold S, Ong MS, Tomlinson G, et al; CARRA Registry STOP-JIA Investigators. Three-year outcomes and latent class trajectory analysis of the childhood arthritis and rheumatology research alliance polyarticular JIA consensus treatment plans study. Arthritis Rheumatol 2025;77(10):1433–1441.
12. Ter Haar NM, van Dijkhuizen EHP, Swart JF, et al. Treatment to target using recombinant Interleukin-1 receptor antagonist as first-line monotherapy in new-onset systemic juvenile idiopathic arthritis: results from a five-year follow-up study. Arthritis Rheumatol 2019; 71(7):1163–1173.
13. Chang MH, Fuhlbrigge RC, Nigrovic PA. Joint-specific memory, resident memory T cells and the rolling window of opportunity in arthritis. Nat Rev Rheumatol 2024;20(5):258–271.
14. Consolaro A, Ruperto N, Bazso A, et al; Paediatric Rheumatology International Trials Organisation. Development and validation of a composite disease activity score for juvenile idiopathic arthritis. Arthritis Rheum 2009;61(5):658–666.
15. R. Version 4.4.1. R Project for Statistical Computing; 2024. Accessed December 2, 2024. https://www.R-project.org/
16. Wickham H, Averick M, Bryan J, et al. Welcome to the tidyverse. J Open Source Softw 2019;4(43):1686.
17. Bates D, Mächler M, Bolker B, et al. Fitting linear mixed-effects models using lme4. J Stat Softw 2015;67(1):1–48.
18. Pardeo M, Rossi MN, Pires Marafon D, et al. Early treatment and IL1RN single-nucleotide polymorphisms affect response to anakinra in systemic juvenile idiopathic arthritis. Arthritis Rheumatol 2021; 73(6):1053–1061.
19. Quinn MA, Emery P. Window of opportunity in early rheumatoid arthritis: possibility of altering the disease process with early intervention. Clin Exp Rheumatol 2003;21(5 suppl 31):S154–S157.
20. Boers M. Understanding the window of opportunity concept in early rheumatoid arthritis. Arthritis Rheum 2003;48(7):1771–1774.
21. Aletaha D, Maa JF, Chen S, et al. Effect of disease duration and prior disease-modifying antirheumatic drug use on treatment outcomes in patients with rheumatoid arthritis. Ann Rheum Dis 2019;78(12): 1609–1615.
22. Wallace CA, Giannini EH, Spalding SJ, et al; Childhood Arthritis and Rheumatology Research Alliance. Trial of early aggressive therapy in polyarticular juvenile idiopathic arthritis. Arthritis Rheum 2012;64(6): 2012–2021.
23. Hissink Muller P, Brinkman DMC, Schonenberg-Meinema D, et al. Treat to target (drug-free) inactive disease in DMARD-naive juvenile idiopathic arthritis: 24-month clinical outcomes of a three-armed randomised trial. Ann Rheum Dis 2019;78(1):51–59.
24. Minden K, Horneff G, Niewerth M, et al. Time of disease-modifying antirheumatic drug start in juvenile idiopathic arthritis and the likelihood of a drug-free remission in young adulthood. Arthritis Care Res (Hoboken) 2019;71(4):471–481.
25. Oliveira Ramos F, Rodrigues AM, Melo AT, et al. Influence of the timing of biological treatment initiation on juvenile idiopathic arthritis long-term outcomes. Arthritis Res Ther 2023;25(1):177.

APPENDIX A: MEMBERS OF THE UCAN CAN-DU AND UCAN CURE CONSORTIA

Members of the UCAN CAN-DU and UCAN CURE consortia are as follows: Adam Huber, Bianca Lang, Chelsea DeCoste, Elizabeth Stringer, and Suzanne Ramsey (IWK Health Centre); Alan Rosenberg, Kate Neufeld, Mehul Jariwala, and Tristan Kerr (University of Saskatchewan); Alexander Mosoiu, Alisa Rachlis, Amy Xu, Arthur Cheng, Brenleigh Jebb, Brian Feldman, Bruno Pereira, Deborah Levy, Dilan Dissanayake, Elizaveta Limenis, Evelyn Rozenblyum, Harper Cheng, Jennifer Ji Young Lee, Lynn Spiegel, Rayfel Schneider, Ronald Laxer, Ruud Verstegen, Shirley Tse, and Trang Duong (The Hospital for Sick Children); Andrea Human, David Cabral, Herman Tam, Jaime Guzman, Kim Morishita, Kristin Houghton, Lori Tucker, Mercedes Chan, Ross Petty, and Tommy Gerschman (British Columbia Children's Hospital); Annet van Royen-Kerkhof, Berent Prakken, Erika Van Nieuwenhove, Marc Jansen, and Nico Wulffraat (Wilhelmina Children's Hospital, University Medical Center Utrecht); Ciarán Duffy, Nadia Luca, Roman Jurencak, and Tala El Tal (Children's Hospital

of Eastern Ontario); Claire LeBlanc, Gaëlle Chédeville, Piya Lahiry, Rosie Scuccimarri, and Sarah Campillo (Montreal Children's Hospital); Clare Hutchinson (North York General Hospital); Daniah Basodan, Dax G. Rumsey, Hon Yan Ng, Jeanine McColl, Lillian Lim, and Tara McGrath (Stollery Children's Hospital); Danielle Brinkman and Petra Hissink Muller (Willem-Alexander Children's Hospital, Leiden University Medical Center); Elizabeth Legger and Wineke Armbrust (Beatrix Children's Hospital, University Medical Center Groningen); Ellen Schatorje and Esther Hoppenreijs (Amalia Children's Hospital, Radboudumc and Sint Maartenskinderkliniek); Elodie Boudes, Gillian Currie, Heinrike Schmeling, Muhammed Dhalla, Nicole Johnson, Paivi Miettunen, Ravneet Sran, and Rebeka Stevenson (University of Calgary); Erkan Demirkaya, Jonathan Park, and Roberta Berard (Children's Hospital, London Health Sciences Centre); Giske Biesbroek and Mariken Gruppen (Emma Children's Hospital, Amsterdam University Medical Centers); Gordon Soon (North York General Hospital/Health Sciences North); Joseph Cafazzo (Centre for Global eHealth Innovation, University Health Network, University of Toronto); Liane Heale, Michelle Batthish, and Tania Cellucci (McMaster Children's Hospital); Lily Lim (Children's Hospital Health Centre Winnipeg); Maarten IJzerman (Erasmus University and University of Melbourne); Marinka Twilt (University of Calgary and Alberta Children's Hospital); Marleen Verkaaik, Philomine van Pelt, and Sylvia Kamphuis (Sophia Children's Hospital, Erasmus University Medical Center); Michelle Kip (University of Twente); Nickolas Blanchette (Trillium Health Partners); Paul Dancey (Janeway Children's/Memorial University); and Regina de Geus (University Medical Center Utrecht).

Arthritis & Rheumatology
Vol. 78, No. 3, March 2026, pp 752–762
DOI 10.1002/art.43428

American College of Rheumatology
Empowering Rheumatology Professionals

Tofacitinib for the Treatment of Juvenile Idiopathic Arthritis: Patient-Reported Outcomes in a Phase 3, Randomized, Double-Blind, Placebo-Controlled Withdrawal Trial

Hermine I. Brunner,[1] Ekaterina Alexeeva,[2] Marcia Bandeira,[3] Ruy Carrasco,[4] Jeffrey Chaitow,[5] Peter Chiraseveenuprapund,[6] Tracy V. Ting,[1] Muferet Erguven,[7] Beth Gottlieb,[8] Sheila Knupp Oliveira,[9] Pooja N. Patel,[10] Seza Ozen,[11] Claudia Saad Magalhaes,[12] Inmaculada Calvo Penadés,[13] David Gruben,[14] Rebecca Germino,[15] Holly B. Posner,[15] Lori Stockert,[16] Alberto Martini,[17] Daniel J. Lovell,[1] and Nicolino Ruperto,[18,19] for the Paediatric Rheumatology International Trials Organisation (PRINTO) and Pediatric Rheumatology Collaborative Study Group (PRCSG)

Objective. Juvenile idiopathic arthritis (JIA) is associated with impaired overall health-related quality of life (HRQoL). We evaluated the impact of tofacitinib on patient-reported outcomes (PROs) in patients with JIA.

Methods. This was a post hoc analysis of a phase 3, randomized, double-blind, placebo-controlled withdrawal trial (NCT02592434) in patients with JIA. In the open-label phase (part 1; weeks 0–18), patients received body weight–based doses of tofacitinib. During the double-blind phase (part 2; weeks 18–44), responders (per JIA-American College of Rheumatology 30 response criteria) were randomized 1:1 to continue tofacitinib or switch to placebo for up to 26 weeks. Assessed PROs included the validated parent and/or legal guardian versions of the Childhood Health Assessment Questionnaire for evaluation of disability, arthritis pain, overall well-being, and the Child Health Questionnaire (CHQ).

Results. Overall, 225 patients were enrolled and received open-label tofacitinib in part 1, and 173 patients were randomized in part 2. During part 1, least-squares (LS) mean (SE) disability, arthritis pain, and overall well-being scores numerically improved from mean 1.04 (SE 0.05), mean 5.53 (SE 0.20), and mean 5.07 (SE 0.20) at baseline to mean 0.57 (SE 0.05), mean 2.46 (SE 0.18), and mean 2.47 (SE 0.18) at week 18, respectively. LS mean (SE) CHQ physical summary and psychological summary scores numerically improved from mean 29.51 (SE 1.15) and mean 47.24 (SE 0.85) at baseline to mean 42.70 (SE 0.98) and mean 51.53 (SE 0.80) at week 18, respectively. Improvements were generally maintained to week 44 in part 2.

Conclusion. Tofacitinib improved a range of PROs in patients with JIA, suggesting potential HRQoL benefits.

INTRODUCTION

Juvenile idiopathic arthritis (JIA) is associated with pain, poor physical functioning, and considerable impairment of overall health-related quality of life (HRQoL).[1–12] Besides improvement of articular involvement and other core measures of JIA, key goals of treatment are the improvement of patient HRQoL, including improving physical function, physical and psychological well-

ClinicalTrials.gov identifier NCT02592434.
Supported by Pfizer Inc.
[1]Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio; [2]National Medical Research Center for Children's Health Federal State Autonomous Institution of the Russian Federation Ministry of Health and I. M. Sechenov First Moscow State Medical University (Sechenovskiy University), Moscow, Russian Federation; [3]Hospital Pequeno Principe, Curitiba, Brazil; [4]Dell Children's Medical Center of Central Texas, Austin; [5]The Sydney Children's Hospital Network, Westmead, New South Wales, Australia; [6]Division of Pediatric Rheumatology, University of Utah, Salt Lake City; [7]Department of Children Health and Diseases, Istanbul Medeniyet University Medical Faculty, Istanbul, Turkey; [8]Pediatric Rheumatology, Cohen Children's Medical Center, New Hyde Park, New York; [9]Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil; [10]Ann & Robert H. Lurie Children's Hospital of Chicago and Northwestern University Feinberg School of Medicine, Chicago, Illinois; [11]Department of Pediatric Rheumatology, Hacettepe University Medical Faculty, Ankara, Turkey; [12]Pediatric Rheumatology Unit, Universidade Estadual Paulista, Botucatu, Brazil; [13]Hospital Universitario y Politecnico La Fe, Valencia, Spain; [14]Pfizer Inc, Groton, Connecticut; [15]Pfizer Inc, New York, New York; [16]Pfizer Inc, Collegeville, Pennsylvania; [17]University of Genova, Genova, Italy; [18]Università degli Studi di Milano-Bicocca, Milan, Italy; [19]Fondazione IRCCS San Gerardo dei Tintori, Reumatologia Pediatrica/Paediatric Rheumatology International Trials Organisation, Monza, Italy.

Upon request, and subject to review, Pfizer will provide the data that support the findings of this study. Subject to certain criteria, conditions, and exceptions, Pfizer may also provide access to the related individual deidentified participant data. See https://www.pfizer.com/science/clinical-trials/trial-data-and-results for more information.

being, and reducing pain.[13–15] Such outcomes are generally referred to as patient-reported outcomes (PROs) and provide unique information on the impact of a medical condition and its treatment that are important to the patient.

In 1997, a core set of six JIA outcome variables was identified, which included two PROs (parent and patient assessment of overall well-being and functional ability). From this core set, a definition of clinically important improvement was derived with the aim of promoting a single efficacy measure in JIA trials,[14] which was later endorsed by the American College of Rheumatology (ACR; JIA-ACR response criteria)[14] and regulatory agencies. In 2016, this definition was re-evaluated and updated with input from patients and parents in which pain was added to the core set.[15] Encouraged by regulatory authorities,[16] the inclusion of PROs in clinical trials has increased over time,[17] and assessment of HRQoL and other PRO measures is considered essential to fully assess the benefits of JIA treatments in both research and routine clinical care.[18,19] PROs in children and adolescents can be collected from either the patient or the parent and/or legal guardian (parent-proxy report).[19,20]

Tofacitinib is an oral JAK inhibitor for the treatment of patients with polyarticular course JIA (pcJIA) and juvenile psoriatic arthritis (jPsA)[21,22] and is currently being investigated for systemic JIA (sJIA).[23] In this post hoc analysis, we report a comprehensive assessment of the impact of tofacitinib on a range of PROs in a phase 3, 44-week clinical trial conducted in patients with JIA.

PATIENTS AND METHODS

Trial design and patients. Key details of the trial design and patient inclusion and exclusion criteria were reported previously.[24] Briefly, this was a phase 3, randomized, double-blind, placebo-controlled withdrawal trial (ClinicalTrials.gov NCT02592434). In the open-label run-in phase (part 1; weeks 0–18), patients received body weight–based doses of tofacitinib (5 mg tablets twice daily or oral formulations as appropriate).[24] During the double-blind phase (part 2; weeks 18–44), patients meeting the JIA-ACR 30 response criteria were randomized 1:1 to continue tofacitinib or switch to placebo for up to an additional 26 weeks. Patients who experienced a JIA flare during the open-label or double-blind phases were discontinued from the trial. Following completion of this trial (or discontinuation for reasons other than treatment-related serious adverse events), patients could enroll in a long-term extension (LTE) study of open-label tofacitinib (NCT01500551).

Eligible patients were aged 2 to 18 years old, met the International League of Associations for Rheumatology JIA classification criteria[25] for any of the following JIA categories grouped under the functional concept of pcJIA (≥5 active joints at enrollment): extended oligoarthritis, rheumatoid factor (RF)+ or RF– polyarthritis, sJIA without systemic features 6 months or more before enrollment, jPsA or enthesitis-related arthritis (ERA; ≥3 active joints at enrollment), and had an inadequate response to one or more disease-modifying antirheumatic drug (DMARD; methotrexate [MTX], or biologic DMARDs [bDMARDs]). The study was sponsored by Pfizer, which manufactures tofacitinib. The trial was designed jointly by Paediatric Rheumatology International Trials Organisation (PRINTO)/Pediatric Rheumatology Collaborative Study Group (PRCSG) investigators and the sponsor, and data were collected by the PRINTO/PRCSG investigators. Data from this trial were used by the sponsor to support the marketing approval of tofacitinib.

This trial was conducted in accordance with the International Council on Harmonization Guidelines for Good Clinical Practice, the Declaration of Helsinki, and local regulatory requirements and laws. The final protocol, any amendments, and informed consent documents were reviewed and approved by the institutional review board or independent ethics committee at every center. Parents, legal guardians, and patients provided written informed consent and assent when required by local institutional review boards and independent ethics committees before performing any study-related activities.

Assessment of PROs. PRO assessments were completed by the parent and/or legal guardian but could be completed by the patient if they were aged ≥14 years and able to complete the assessment correctly and consistently for the duration of the study. Disability, arthritis pain, and patient overall well-being were assessed as part of the Childhood Health Assessment Questionnaire (CHAQ), administered at each visit (baseline and weeks 2, 4, 8, 12, 18, 20, 24, 28, 32, 36, 40, and 44). Disability was assessed by measuring physical function impairment using the CHAQ-Disability Index (CHAQ-DI); scores could range from 0 to 3, with higher scores indicating more disability. Arthritis pain (CHAQ-Discomfort Index) and patient overall well-being were assessed on a 21-circle visual analog scale (VAS); scores could range from 0 to 10, with higher scores indicating more arthritis pain or worse well-being, as appropriate. Additionally, duration of morning stiffness (minutes) was collected at each visit. HRQoL was assessed at baseline and weeks 18 and 44 by the Child Health Questionnaire (CHQ), which includes ratings for 15 CHQ health concepts that can be summarized in physical summary (PhS) and psychosocial summary (PsS) scores; each of the 15 health concept scores could range from 0 to 100

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/art.43428).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/art.43428.

Address correspondence via email to Hermine I. Brunner, MD, MSc, MBA, at Hermine.Brunner@cchmc.org.

Submitted for publication August 8, 2024; accepted in revised form July 25, 2025.

(with the exception of change in health, ranging from 1–5), with higher scores indicating better physical function or mental health, as appropriate. CHQ PhS and PsS scores were standardized to have a mean (SD) of 50 (±10) in a normative population of US children.[5,26] The cross-culturally adapted and validated national language versions of the CHAQ (parent and/or legal guardian version) and CHQ-parent form 50 were used.[11] For patients with ERA, overall back pain and nocturnal back pain were also measured at each visit.

Rates of patients achieving minimal clinically important differences (MCID) in CHAQ-DI (score reduction of ≥0.188),[27] minimal arthritis pain (CHAQ-Discomfort Index score of ≤0.35), or no disability status (CHAQ-DI = 0) were also assessed. Additional details on the PROs assessed in this analysis are summarized in the Supplementary Materials.

Physician-reported outcome assessments. In addition to the assessment of PROs, Physician Global Assessment (PGA) and the number of joints with active arthritis were also assessed at each visit. PGA was assessed on a 10-cm VAS with scores ranging from 0 (no disease activity) to 10 (maximum disease activity). The number of joints with active arthritis was identified using the JIA-ACR definition: any joint with swelling or, in the absence of swelling, limitation of motion accompanied by either pain on motion or tenderness not due to deformity (range 0–71).

Statistical analysis. Descriptive statistics are provided for selected characteristics of the part 1 baseline values. Mixed-effects linear models were fit using PROC MIXED in SAS version 9.4. For part 1 of the study, the model included the effects of the week (postbaseline, at visits where the dependent variable was measured), baseline C-reactive protein (categorized as normal or above normal), patient population and/or group (overall, pcJIA, jPsA, and ERA), and baseline value of the dependent variable; patients were a random effect. Least-squares (LS) means for each outcome were assessed at each timepoint and are reported. For part 2 of the study, the model was the same as part 1, except with the addition of treatment group and an interaction term of treatment with the effect of week. LS means were analyzed with a longitudinal mixed-effects model. *P* values for the difference between the treatment arms in part 2 at each timepoint were provided for all PROs, except CHQ health concepts. The rate of patients achieving MCID in CHAQ-DI, minimal arthritis pain, or no disability status were analyzed by visit using the normal approximation approach for binomial populations. A Kruskal-Wallis test was performed by visit to assess whether distributions of the data between treatment arms were significantly different (medians, interquartile ranges, and *P* values for these tests are reported in the Supplementary Materials). Because these are post hoc analyses, *P* values were provided to indicate the strength in the degree of change in the outcomes and are not for hypothesis testing. Analyses of PROs were based on observed data without imputation for missing values; patients who discontinued from the study due to JIA flares or other reasons were not analyzed at subsequent timepoints. The number of patients with available data at each timepoint is reported for each PRO. Analyses of PGA and the number of joints with active arthritis were descriptive.

RESULTS

Patients. Overall, 225 patients were enrolled and received open-label tofacitinib in part 1 (pcJIA, N = 184; jPsA, N = 20; ERA, N = 21). A total of 173 patients who enrolled in part 1 were subsequently randomized in part 2, including 142 patients with pcJIA (tofacitinib, n = 72; placebo, n = 70), 15 patients with jPsA (tofacitinib, n = 7; placebo, n = 8), and 16 patients with ERA (tofacitinib, n = 9; placebo, n = 7).

The majority of patients enrolled in part 1 of the study were girls (75.1%), aged 12 to 18 years (61.8%), White (87.1%), and weighed ≥40 kg (62.7%). At baseline, the scores (25th percentile [Q1], 75th percentile [Q3]) were median 0.9 (Q1 0.3, Q3 1.5) for CHAQ-DI, median 6.0 (Q1 3.5, Q3 7.0) for arthritis pain, and median 5.0 (Q1 3.0, Q3 7.0) for patient overall well-being (Table 1).

Baseline scores of PROs for individual pcJIA, jPsA, and ERA groups are reported within Supplementary Table 1. Part 1 baseline PRO scores for patients randomized to part 2, overall and by JIA category, are summarized in Supplementary Table 2.

Changes in PROs during part 1. During part 1 in the overall population, continued improvements in patient disability, pain, and overall well-being were observed through to week 18. LS mean CHAQ-DI scores (Figure 1), arthritis pain scores (Figure 2), and patient overall well-being scores (Figure 3) numerically improved from baseline to week 18. At week 18, the rate of achieving MCID in CHAQ-DI was 48.0% (SE 3.33; Supplementary Figure 1). The proportion of patients with no disability (CHAQ-DI = 0) generally increased over time from 12.9% (SE 2.2) at baseline to 28.0% (SE 3.0) at week 18 (Supplementary Figure 2A). The proportion of patients achieving minimal arthritis pain (CHAQ-Discomfort Index score of ≤0.35) generally increased during part 1 from 4.4% (SE 1.4) at baseline to 15.6% (SE 2.4) at week 18 (Supplementary Figure 3A).

LS mean scores for patient disability, arthritis pain, and overall well-being in pcJIA, jPsA, and ERA groups are shown in Supplementary Figures 4 to 6. The rates of achieving MCID in CHAQ-DI, proportion of patients achieving no disability, and proportion of patients achieving minimal arthritis pain in the pcJIA, jPsA, and ERA groups are reported in Supplementary Figures 1, 2B to 2D, and 3B to 3D, respectively.

In the overall population, numerical improvements from baseline in both CHQ PhS and PsS scores were observed at week 18 (Figure 4A). The greatest improvements were observed

Table 1. Patient characteristics and baseline patient-reported outcomes for patients with juvenile idiopathic arthritis receiving tofacitinib in part 1 (overall population)*

Characteristics	Overall (N = 225)
Patient characteristics	
Female sex, n (%)	169 (75.1)
Age, median (Q1, Q3), years	13.0 (9.0, 15.0)
Disease duration, median (Q1, Q3), years	2.5 (1.0, 5.6)
Patient-reported outcomes, median (Q1, Q3)	
CHAQ-DI score	0.9 (0.3, 1.5)
Arthritis pain	6.0 (3.5, 7.0)
Patient overall well-being	5.0 (3.0, 7.0)
CHQ physical summary score	30.6 (19.6, 42.6)
CHQ psychosocial summary score	48.2 (40.6, 56.6)
CHQ health concepts	
Global health	60.0 (30.0, 60.0)
Physical function	66.7 (38.9, 88.9)
Social limitations–physical	66.7 (33.3, 100.0)
Social limitations–emotional	88.9 (44.4, 100.0)
Bodily pain	40.0 (20.0, 60.0)
Behavior	79.2 (64.2, 89.2)
Global behavior	85.0 (60.0, 85.0)
Mental health	70.0 (55.0, 80.0)
Self-esteem	75.0 (62.5, 87.5)
General health	51.7 (42.5, 60.0)
Change in health	3.0 (2.0, 4.0)
Parental emotional impact	58.3 (33.3, 75.0)
Parental time impact	77.8 (55.6, 100.0)
Family activity	79.2 (62.5, 95.8)
Family cohesion	85.0 (60.0, 85.0)
Duration of morning stiffness, minutes	30.0 (15.0, 60.0)
Physician-reported outcomes, median (Q1, Q3)	
PGA	6.0 (4.5, 7.5)
Number of joints with active arthritis	10.0 (6.0, 15.0)

* CHAQ-DI, Childhood Health Assessment Questionnaire-Disability Index; CHQ, Child Health Questionnaire; PGA, Physician Global Assessment of Overall Disease Activity; Q1, 25th percentile; Q3, 75th percentile.

in CHQ PhS, which improved from approximately 2 SDs below the standardized mean score of 50 for a normative population to within the range of the normative population (Figure 4A); mean PsS scores remained within the range of the normative population at baseline and week 18. LS mean scores for CHQ PhS and PsS for the pcJIA, jPsA, and ERA groups are shown in Supplementary Figure 7.

The LS mean values of the 15 CHQ health concepts scores at baseline and week 18 for the overall population are shown in Figure 4B. At the end of part 1, numerical improvements from baseline were generally observed with tofacitinib across all CHQ health concepts. The greatest improvement was observed in bodily pain, which improved from 40.05 at baseline to 70.54 at week 18. Substantial improvements from baseline were also observed in global health (baseline 47.32, week 18 68.09), physical function (baseline 58.33, week 18 79.30), and role and social limitations–physical (baseline 62.75, week 18 83.44). LS mean values of the 15 CHQ health concepts for pcJIA, jPsA, and ERA groups are shown in Supplementary Figure 8A to 8C.

A numerical reduction in the duration of morning stiffness was observed during part 1 in the overall population (Figure 5). Changes in the duration of morning stiffness for pcJIA, jPsA, and ERA groups, as well as changes in average back pain and nocturnal back pain in the ERA group, are reported in Supplementary Figure 9 and Supplementary Figure 10, respectively.

Changes in PROs during part 2. In the overall population, improvements in PROs achieved at the end of part 1 were generally maintained throughout part 2, although some differences were observed between the two treatment arms. Disability scores (LS mean) were generally maintained throughout part 2, consistent with part 2 baseline (week 18) values for both treatment arms (Figure 1). Throughout part 2, arthritis pain scores generally remained stable and consistent with part 2 baseline values, although patients receiving tofacitinib demonstrated lower ($P < 0.05$) arthritis pain scores compared with patients receiving placebo at week 28 (Figure 2). Patient overall well-being scores remained consistent with part 2 baseline values throughout part 2 in patients receiving tofacitinib (Figure 3). Patients receiving tofacitinib generally demonstrated similar overall well-being scores compared with those receiving placebo; however, scores were lower (indicating greater well-being) in the tofacitinib group at week 24 and week 32 (both $P < 0.05$).

In general, the proportion of patients achieving no disability (CHAQ = 0) in the overall population numerically improved throughout part 2 in patients receiving tofacitinib (week 18, proportion 33.0%, SE 5.0; week 44, proportion 38.6%, SE 5.2) but numerically declined in patients receiving placebo (week 18, proportion 40.0%, SE 5.3; week 44, proportion 23.5%, SE 4.6; Supplementary Figure 2A). A higher ($P < 0.05$) proportion of patients receiving tofacitinib achieved no disability compared with patients receiving placebo at week 36 (proportion 37.5%, SE 5.2 vs proportion 23.5%, SE 4.6) and week 44 (proportion 38.6%, SE 5.2 vs proportion 23.5%, SE 4.6). The proportion of patients with minimal arthritis pain generally remained stable throughout part 2 in patients receiving tofacitinib, whereas a numerically higher proportion of patients achieved minimal arthritis pain at week 44 (proportion 26.1%, SE 4.7) compared with part 2 baseline (proportion 20.5%, SE 4.3; Supplementary Figure 3A). The proportion of patients receiving placebo and achieving minimal arthritis pain generally declined during part 2 from 20.0% (SE 4.3) at baseline to 11.8% (SE 3.5) at week 44 (Supplementary Figure 3A). A higher proportion of patients achieved minimal arthritis pain in the tofacitinib group than in the placebo group at week 28 (proportion 23.9, SE 4.5 vs proportion 8.2, SE 3.0; $P < 0.01$), week 32 (proportion 22.7, SE 4.5 vs proportion 7.1, SE 2.8; $P < 0.01$), week 40 (proportion 28.4, SE 4.8 vs proportion 12.9, SE 3.6; $P < 0.05$), and week 44 (proportion 26.1, SE 4.7 vs proportion 11.8, SE 3.5; $P < 0.05$).

Improvements in CHQ PhS and PsS scores (LS means) in the overall population were maintained at part 2 baseline

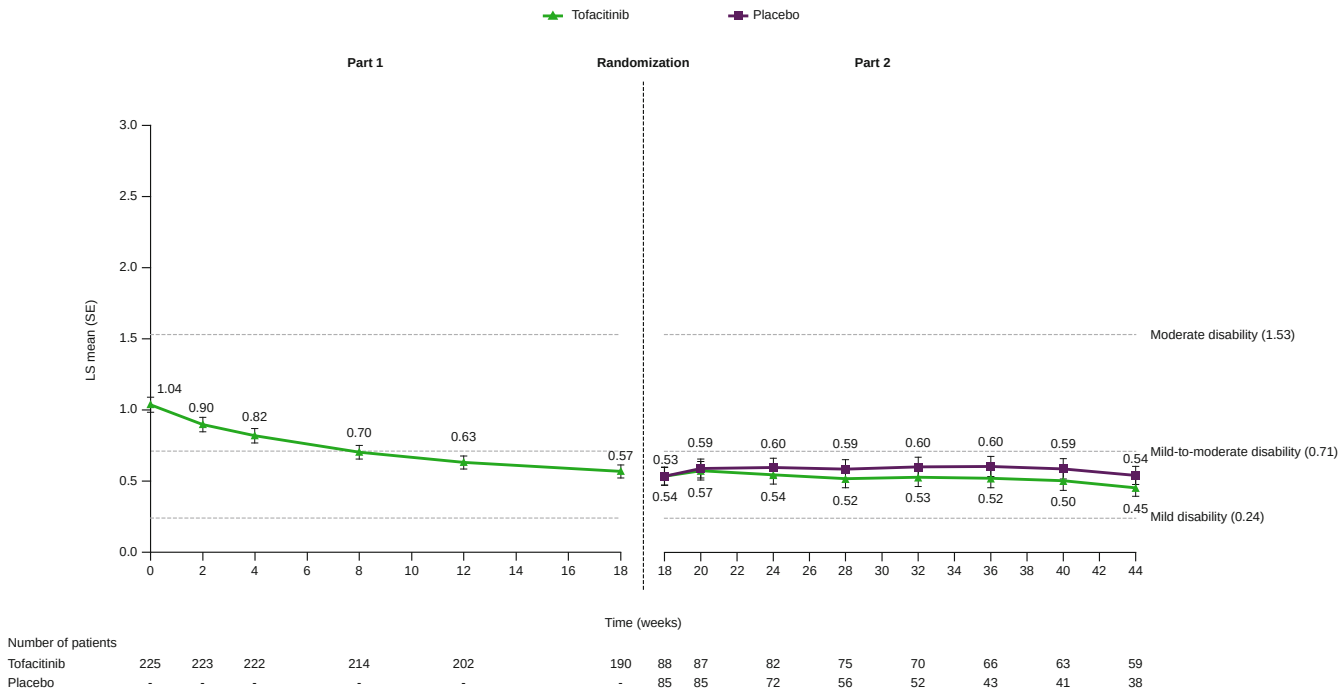


Figure 1. LS mean (±SE) Childhood Health Assessment Questionnaire-Disability Index scores over time in patients with juvenile idiopathic arthritis treated with tofacitinib in part 1 and tofacitinib or placebo in part 2 (overall population). Analyses were based on observed data; missing values were not imputed. Horizontal dashed lines represent mean Childhood Health Assessment Questionnaire-Disability Index scores for patients with mild, mild-to-moderate, and moderate disability, which have been previously reported.[28] Vertical dashed line represents randomization (patients meeting the juvenile idiopathic arthritis American College of Rheumatology 30 response criteria randomized 1:1 to continue tofacitinib or switch to placebo for up to an additional 26 weeks). Median and interquartile range values for Childhood Health Assessment Questionnaire-Disability Index are summarized in Supplementary Table 3; no significant differences in the distributions between the treatment arms were observed during part 2. LS, least-squares.

levels through to week 44 and were similar between patients receiving tofacitinib and those receiving placebo (Figure 4). At week 44, most CHQ health concept scores (LS means) were either maintained at part 2 baseline values or showed further numerical improvement regardless of treatment arm (Figure 4B).

There was a small numerical increase in the duration of morning stiffness during part 2 for patients receiving both tofacitinib and placebo (Figure 5). Patients receiving tofacitinib demonstrated a shorter duration of morning stiffness at week 28 compared with those receiving placebo ($P < 0.05$). Part 2 data for pcJIA, jPsA, and ERA groups are presented within the Supplementary Materials.

Physician-reported outcomes in parts 1 and 2. In the overall population, both the PGA scores and the number of joints with active arthritis demonstrated a similar pattern of improvement to that observed with the PROs. During part 1, the mean PGA score decreased from 6.18 (SD 1.88) at baseline to 1.68 (SD 1.66) at week 18 (Supplementary Figure 11A). Similarly, the mean number of joints with active arthritis decreased from 12.23 (SD 8.11) at baseline to 1.94 (SD 3.73) at week 18 (Supplementary Figure 12A).

During part 2 of the study, PGA scores and the number of joints with active arthritis were similar between both treatment arms and remained consistent with part 2 baseline values throughout the observation period (Supplementary Figures 11A and 12A). Data for pcJIA, jPsA, and ERA groups are presented within Supplementary Figures 11B to 11D and 12B to 12D.

DISCUSSION

The efficacy of tofacitinib versus placebo to improve JIA signs and symptoms has previously been demonstrated in a phase 3 study in patients with JIA.[24] HRQoL assessments highlight aspects of disease that are important to the patient but may not be captured by standard clinical assessments. In this post hoc analysis, we observed improvements in a range of PROs and two physician-reported effectiveness outcomes in patients with JIA receiving open-label tofacitinib during part 1 of the study.

When patients were randomized to tofacitinib or placebo in part 2, both treatment arms generally maintained these improvements or showed further numerical improvements with tofacitinib versus placebo across several PROs. Improvements in PROs were also observed in all JIA subgroups studied. Although the general trends of improvement observed in the jPsA and ERA

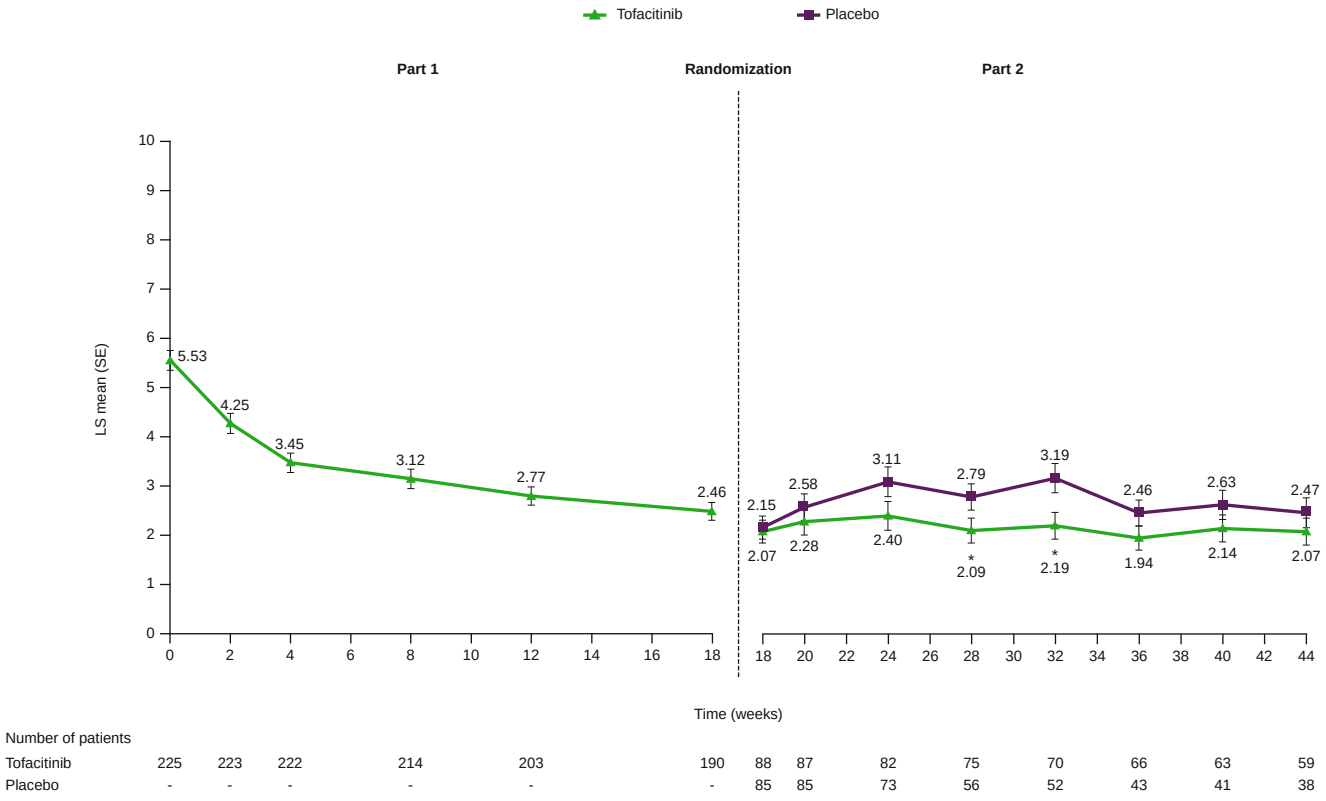


Figure 2. LS mean (±SE) arthritis pain (Childhood Health Assessment Questionnaire-Discomfort Index) scores over time in patients with juvenile idiopathic arthritis treated with tofacitinib in part 1 and tofacitinib or placebo in part 2 (overall population). *$P < 0.05$ vs placebo. Analyses were based on observed data; missing values were not imputed. Dashed line represents randomization (patients meeting the juvenile idiopathic arthritis American College of Rheumatology 30 response criteria randomized 1:1 to continue tofacitinib or switch to placebo for up to an additional 26 weeks). Median and interquartile range values for arthritis pain (Childhood Health Assessment Questionnaire-Discomfort Index) scores are summarized in Supplementary Table 3; significant differences ($P < 0.05$) in the distributions were observed between the treatment arms at week 32. LS, least-squares. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.43428/abstract.

groups were typically consistent with those in the overall population, which mostly consisted of patients with pcJIA, there were fluctuations of improvement that were largely driven by the small sample sizes of these two patient groups.

The observation that improvements at the end of part 1 with tofacitinib (week 18) were generally maintained at week 44 in the placebo group may suggest a prolonged biologic effect of tofacitinib on patient-reported and perceived impacts of JIA that persisted once treatment was switched to placebo. As previously reported, more patients receiving placebo discontinued during part 2 of the study compared with patients receiving tofacitinib (55% vs 31%, respectively), with the majority of patients discontinuing due to flares (52% vs 25%, respectively).[24] Analysis of long-term follow-up data could inform on whether any minimal deterioration in PROs reported due to flare may later be reversed.

The functional limitation and disability resulting from JIA can contribute to reduced HRQoL by restricting a patient's mobility and activities. Indeed, it has previously been observed that a CHAQ score of >1 is the strongest determinant of poorer HRQoL in CHQ PhS, indicating that functional impairment has a profound impact on physical well-being.[3] In the current study, CHAQ scores improved from baseline during part 1, and these improvements were generally maintained with scores between 0.5 to 0.6 wthrough to week 44, regardless of treatment arm. Furthermore, we observed substantial improvements in "physical function" and "role/social limitations–physical" CHQ health concepts; improvement in physical function may therefore be associated with improved physical well-being and allow patients greater mobility to resume their daily activities.

Consistent with previous studies in patients with JIA, CHQ PhS scores were considerably lower than CHQ PsS scores at baseline[4,5] and were approximately 2 SD below the standardized mean score of a normative population (mean 50 ± SD 10).[5] Substantial increases from baseline were observed in CHQ PhS scores, which reached the normative population values at week 18 and were generally maintained at week 44 in both the tofacitinib and the placebo groups. Although CHQ PsS scores demonstrated slight numerical improvement from baseline at week 18, which was generally maintained at week 44, mean scores were within normative population values throughout the study irrespective of treatment group. Notably, the differences in scores observed are unlikely to be clinically meaningful. Furthermore, the

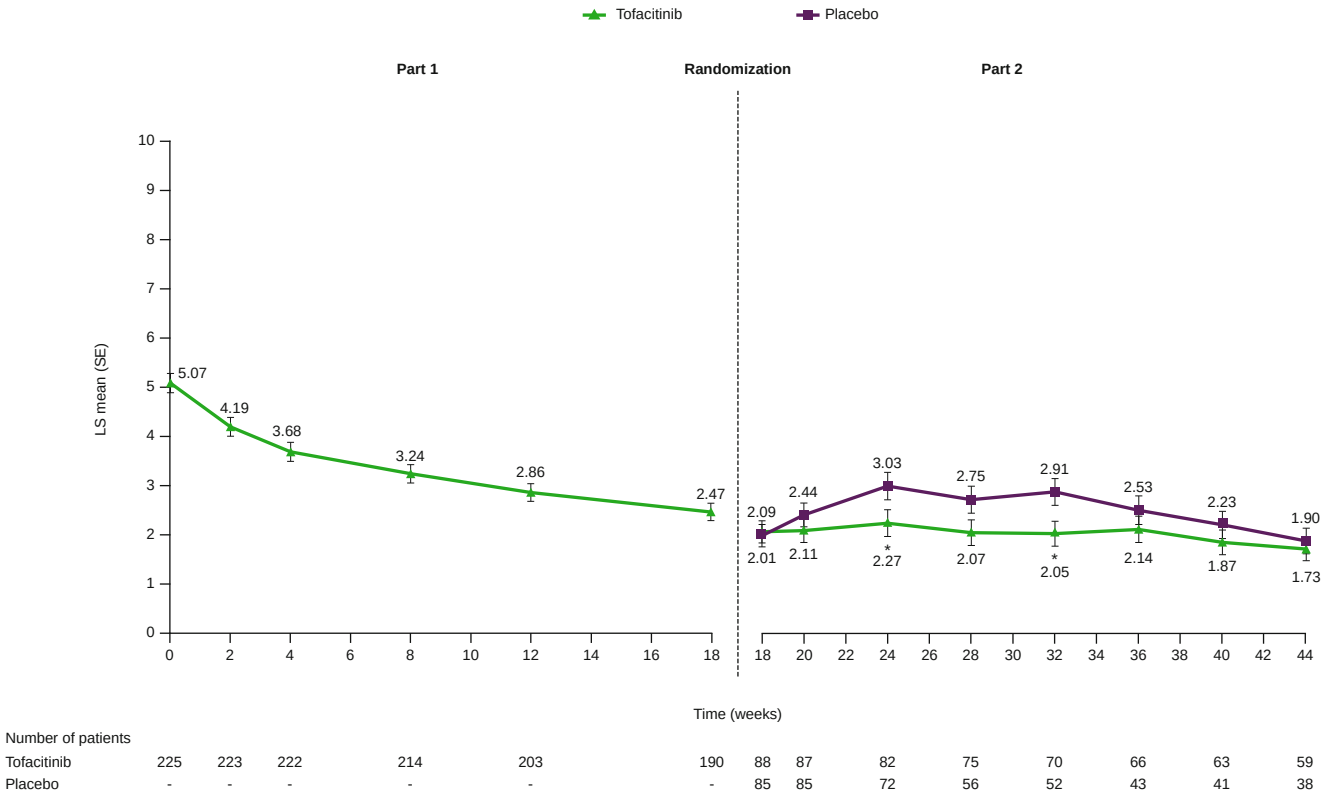


Figure 3. LS mean (±SE) patient overall well-being scores over time in patients with juvenile idiopathic arthritis treated with tofacitinib in part 1 and tofacitinib or placebo in part 2 (overall population). *$P < 0.05$ vs placebo. Analyses were based on observed data; missing values were not imputed. Dashed line represents randomization (patients meeting the juvenile idiopathic arthritis American College of Rheumatology 30 response criteria randomized 1:1 to continue tofacitinib or switch to placebo for up to an additional 26 weeks). Median and interquartile range values for patient overall well-being are summarized in Supplementary Table 3; no significant differences in the distributions were observed between the treatment arms during part 2. LS, least-squares. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.43428/abstract.

patterns observed with treatment in both the CHQ PhS and PsS are consistent with those observed previously in patients with JIA treated with abatacept[5] or MTX.[4]

Duration of morning stiffness was substantially reduced during part 1 with tofacitinib treatment. Although there was a slight numerical increase in the duration of morning stiffness for both treatment arms in part 2, patients receiving placebo demonstrated numerically higher durations of stiffness throughout compared with patients receiving tofacitinib. This trend in improvement is generally similar to that observed in adults with rheumatoid arthritis who received 12 months of tofacitinib monotherapy, although adult patients demonstrated longer durations of morning stiffness versus the current population.[29]

The disease burden associated with JIA includes disability, joint pain, and reduced HRQoL, as previously reported by patients and their caregivers.[7,30,31] Indeed, pain has previously been identified as a core outcome by patients with JIA and their caregivers when evaluating the effects of treatments, highlighting it as an important outcome for those living with JIA.[15] Pain can contribute to impaired HRQoL; previous studies have observed that persistence of poor physical well-being following treatment with MTX was associated with several factors that included greater disability and pain[4] and that the intensity of a child's pain is a major predictor of poor psychosocial well-being.[3] Additionally, patients with JIA and higher levels of pain may experience more problems with their physical, emotional, and social functioning.[32] In the current analysis, tofacitinib improved arthritis pain during part 1, and these improvements were generally maintained at the end of part 2, regardless of treatment arm. Consistent with this observation, the CHQ health concept of "bodily pain" showed the largest improvement among the 15 CHQ health concepts at the end of part 1, with further, although more modest, improvement observed at the end of part 2 in patients receiving tofacitinib. The proportion of patients achieving minimal pain also generally increased during part 1 and was subsequently maintained in part 2 in patients receiving tofacitinib. Such improvements in pain may enable patients to resume day-to-day activities and, in school-age children, minimize time they are away from school, which can subsequently reduce stress placed on parents and/or caregivers.[33] Thus, alleviation of pain may contribute to a patient's improved physical, mental, emotional, and social functioning.

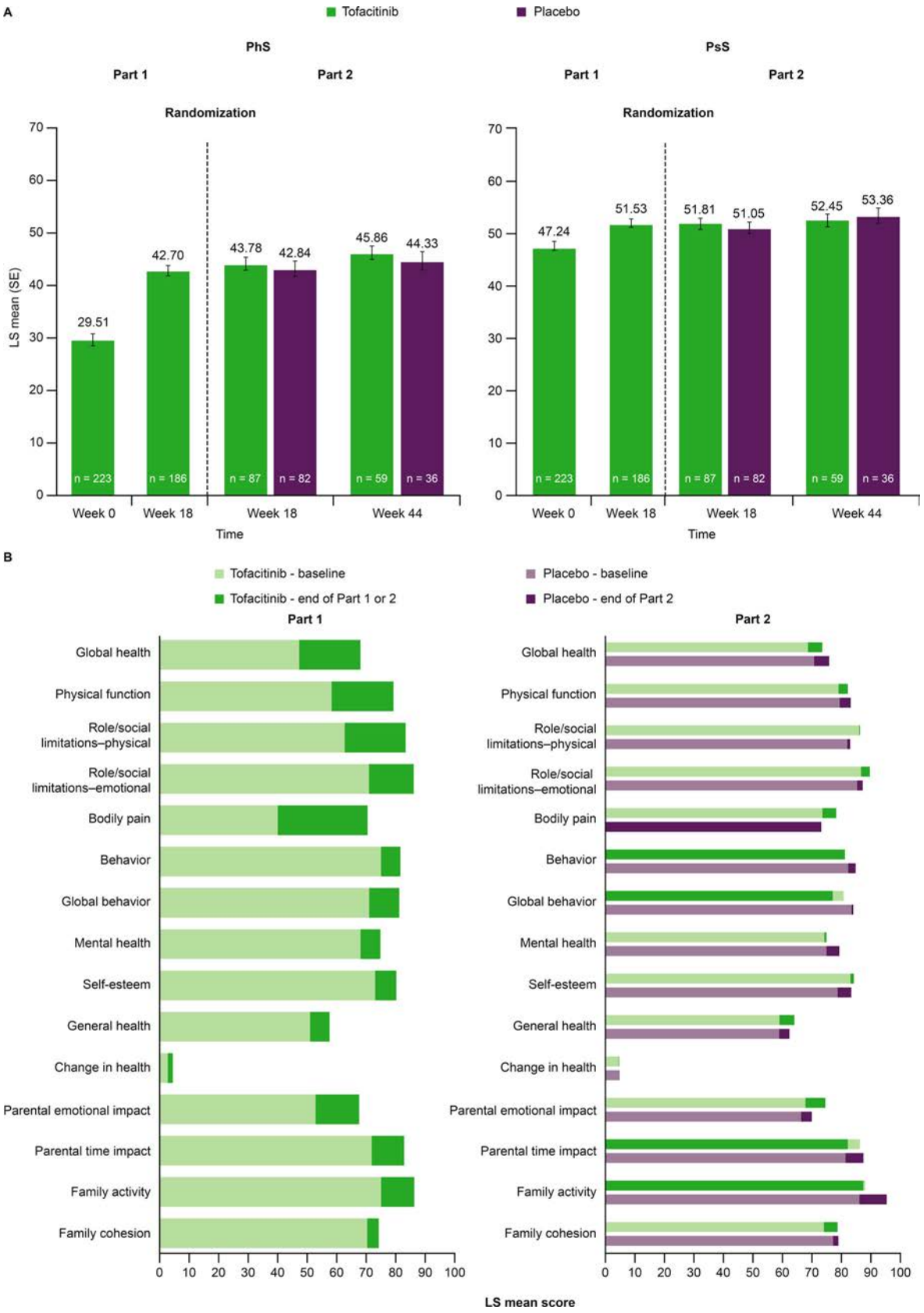


Figure 4. Legend on next page.

Figure 4. (A) LS mean (±SE) Child Health Questionnaire physical summary and Child Health Questionnaire psychosocial summary scores and (B) LS mean Child Health Questionnaire health concept scores at baseline and at the end of part 1 and part 2 in patients with juvenile idiopathic arthritis treated with tofacitinib in part 1 and tofacitinib or placebo in part 2 (overall population). Analyses were based on observed data; missing values were not imputed. Dashed line represents randomization (patients meeting the juvenile idiopathic arthritis American College of Rheumatology 30 response criteria randomized 1:1 to continue tofacitinib or switch to placebo for up to an additional 26 weeks). Median and interquartile range values for Child Health Questionnaire physical summary and psychosocial summary are summarized in Supplementary Table 3; no significant differences in the distributions between the treatment arms was observed during part 2. N values, medians, and interquartile range values for Child Health Questionnaire health concept scores are summarized in Supplementary Table 4. LS, least-squares; PhS, physical summary; PsS, psychosocial summary. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.43428/abstract.

The effect of bDMARDs on pain, physical function, and HRQoL in patients with JIA has been investigated previously. Improvement in functional ability, pain, and well-being was observed over 104 weeks following biweekly intravenous infusion of tocilizumab in patients with sJIA or every 4 weeks for patients with pcJIA.[7] Furthermore, in a phase 3 withdrawal trial in pcJIA, substantial improvements in all CHQ health concepts, PhS, and PsS scores were observed during the four-month open-label period in patients receiving abatacept, with further improvements versus placebo in the six-month double-blind period.[5] Similar trends were observed for CHQ in a randomized trial with standard or higher doses of MTX in JIA.[4] Although comparisons between these studies and the current post hoc analysis are limited due to the different methodologies used, collectively, our observations highlight that tofacitinib, an oral treatment,[21,22] can improve HRQoL in patients with JIA, akin to other advanced therapies that are administered intravenously or subcutaneously.

There are a number of limitations to be considered when interpreting the findings of this analysis. The phase 3 study was powered for the analysis of the primary outcome, whereas the current analysis was post hoc in nature and therefore insufficiently powered for the assessment of additional outcomes. Because of the large number of outcomes assessed, the study findings should be interpreted with caution, given that correction for multiple testing was not performed. Although data for individual patient groups are provided in the Supplementary Materials, these data should be interpreted within the context of the reduced sample sizes, particularly for the jPsA or ERA groups. Additionally, per

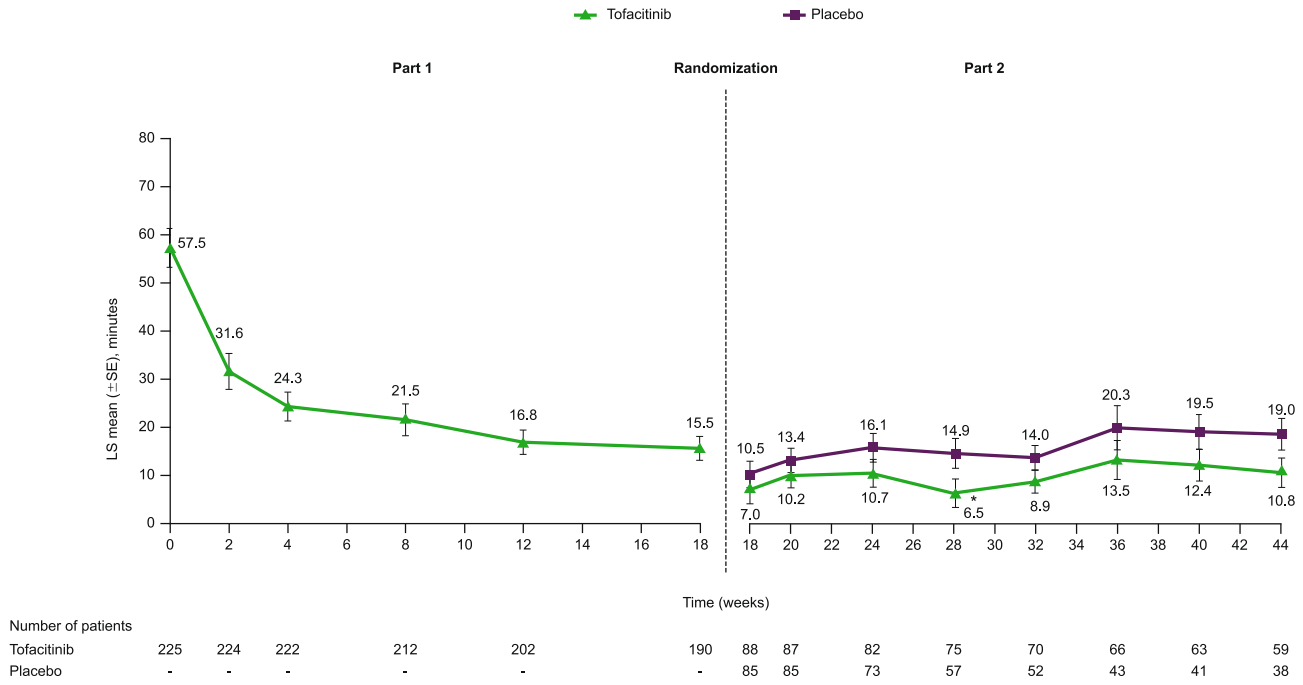


Figure 5. LS mean (±SE) duration of morning stiffness (minutes) in patients with juvenile idiopathic arthritis treated with tofacitinib in part 1 and tofacitinib or placebo in part 2 (overall population). *$P < 0.05$ vs placebo. Analyses were based on observed data; missing values were not imputed. Dashed line represents randomization (patients meeting the juvenile idiopathic arthritis American College of Rheumatology 30 response criteria randomized 1:1 to continue tofacitinib or switch to placebo for up to an additional 26 weeks). Median and interquartile range values for the duration of morning stiffness are summarized in Supplementary Table 3; significant differences ($P < 0.05$) in the distributions were observed between the treatment arms at weeks 28 and 40. LS, least-squares. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/art.43428/abstract.

the trial protocol, the CHAQ assessments were administered to the patient's parent and/or legal guardian at each visit. Patients aged ≥14 years old could complete the assessment themselves; however, the responses used in the analysis were from the parent and/or legal guardian version of the CHAQ, rather than the patient version. Finally, the results may be influenced by the discontinuation of patients who experienced JIA flares, as per the trial protocol (data for these patients were not imputed for these post hoc analyses). It would be expected that JIA flare would be accompanied by worsening of PROs, which would not be captured at postdiscontinuation timepoints in our analyses. Patients who discontinued due to JIA flares could enter the LTE study (NCT01500551) and continue and/or reinitiate treatment with open-label tofacitinib. Although it was beyond the scope of this analysis to separately evaluate those patients who restarted tofacitinib in the LTE study, an interim analysis of the full LTE population found that improvements in CHAQ-DI scores, arthritis pain, and well-being were generally sustained through <48 months of open-label treatment with tofacitinib.[34]

In conclusion, among patients with pcJIA, jPsA, or ERA who had inadequate response to one or more DMARD, tofacitinib use was associated with sustained improvements in a range of PROs that encompassed disability and/or physical function, arthritis pain, well-being, and HRQoL up to approximately one year. These benefits were often maintained even in patients randomized to placebo in the double-blind phase of the study after receiving open-label tofacitinib in part 1. Collectively, our findings demonstrate that tofacitinib may provide substantial improvement in HRQoL for patients with JIA.

ACKNOWLEDGMENTS

We thank the participating centers, the families of the children with JIA, and the research assistants at the PRINTO and PRCSG coordinating centers for their remote and independent evaluation of the primary and secondary endpoints. We thank the late Dr Ricardo M. Suehiro for his contributions to the study and to the development of the article and express our condolences to his family and colleagues.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr Brunner confirms that all authors have provided the final approval of the version to be published and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

ROLE OF THE STUDY SPONSOR

This study was sponsored by Pfizer. The phase 3 trial was designed jointly by PRINTO/PRCSG investigators (NR, HIB, AM, and DJL) and the sponsor. Data were collected by the PRINTO/PRCSG investigators. The sponsor was involved in the overall management of the trial, data collection, and data analysis. The authors meet criteria for authorship as recommended by the International Committee of Medical Journal Editors and did not receive payment for development of this article. Authors from Pfizer were involved in the study design, data analysis, interpretation of data, and writing of the manuscript, and approved the content of the submitted manuscript. Medical writing support, under the direction of the authors, was provided by Robert Morgan, PhD, of CMC Connect, a division of IPG Health Medical Communications, and was funded by Pfizer Inc, New York, NY, USA, in accordance with Good Publication Practice (GPP 2022) guidelines (*Ann Intern Med* 2022;175:1298–1304). Pfizer was given the opportunity to review the manuscript for medical and scientific accuracy as well as intellectual property considerations. Publication of this article was not contingent upon approval by Pfizer.

REFERENCES

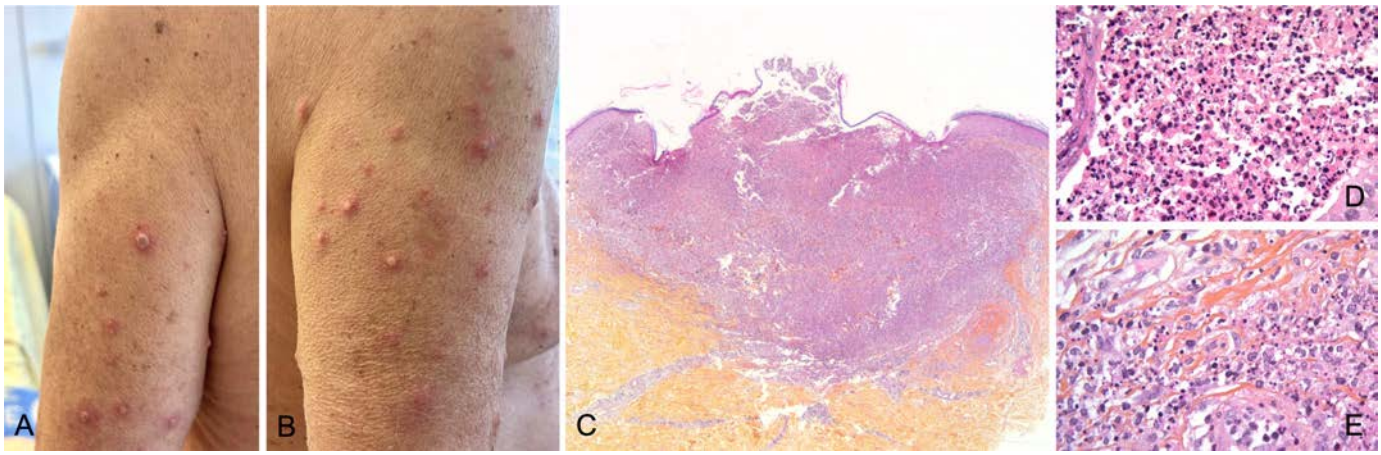
1. Taxter AJ, Wileyto EP, Behrens EM, et al. Patient-reported outcomes across categories of juvenile idiopathic arthritis. J Rheumatol 2015; 42(10):1914–1921.
2. Gutiérrez-Suárez R, Pistorio A, Cespedes Cruz A, et al; Pediatric Rheumatology International Trials Organisation (PRINTO). Health-related quality of life of patients with juvenile idiopathic arthritis coming from 3 different geographic areas. The PRINTO multinational quality of life cohort study. Rheumatology (Oxford) 2007;46(2):314–320.
3. Oliveira S, Ravelli A, Pistorio A, et al; Pediatric Rheumatology International Trials Organization (PRINTO). Proxy-reported health-related quality of life of patients with juvenile idiopathic arthritis: the Pediatric Rheumatology International Trials Organization multinational quality of life cohort study. Arthritis Rheum 2007;57(1):35–43.
4. Céspedes-Cruz A, Gutiérrez-Suárez R, Pistorio A, et al; Pediatric Rheumatology International Trials Organization (PRINTO). Methotrexate improves the health-related quality of life of children with juvenile idiopathic arthritis. Ann Rheum Dis 2008;67(3):309–314.
5. Ruperto N, Lovell DJ, Li T, et al; Paediatric Rheumatology International Trials Organisation (PRINTO); Pediatric Rheumatology Collaborative Study Group (PRCSG). Abatacept improves health-related quality of life, pain, sleep quality, and daily participation in subjects with juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2010; 62(11):1542–1551.
6. Lovell DJ, Ruperto N, Mouy R, et al; Pediatric Rheumatology Collaborative Study Group and the Paediatric Rheumatology International Trials Organisation. Long-term safety, efficacy, and quality of life in patients with juvenile idiopathic arthritis treated with intravenous abatacept for up to seven years. Arthritis Rheumatol 2015;67(10):2759–2770.
7. Brunner HI, Chen C, Bovis F, et al; Paediatric Rheumatology International Trials Organisation and the Pediatric Rheumatology Collaborative Study Group. Functional ability and health-related quality of life in randomized controlled trials of tocilizumab in patients with juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2021;73(9):1264–1274.
8. Filocamo G, Consolaro A, Schiappapietra B, et al. A new approach to clinical care of juvenile idiopathic arthritis: the Juvenile Arthritis Multidimensional Assessment Report. J Rheumatol 2011;38(5):938–953.
9. Ruperto N, Levinson JE, Ravelli A, et al. Long-term health outcomes and quality of life in American and Italian inception cohorts of patients with juvenile rheumatoid arthritis. I. Outcome status. J Rheumatol 1997;24(5):945–951.
10. Ruperto N, Ravelli A, Levinson JE, et al. Long-term health outcomes and quality of life in American and Italian inception cohorts of patients with juvenile rheumatoid arthritis. II. Early predictors of outcome. J Rheumatol 1997;24(5):952–958.

11. Ruperto N, Ravelli A, Pistorio A, et al; Paediatric Rheumatology International Trials Organisation. Cross-cultural adaptation and psychometric evaluation of the Childhood Health Assessment Questionnaire (CHAQ) and the Child Health Questionnaire (CHQ) in 32 countries. Review of the general methodology. Clin Exp Rheumatol 2001; 19(4 Suppl 23):S1–S9.
12. Bovis F, Consolaro A, Pistorio A, et al; Paediatric Rheumatology International Trials Organisation (PRINTO). Cross-cultural adaptation and psychometric evaluation of the Juvenile Arthritis Multidimensional Assessment Report (JAMAR) in 54 languages across 52 countries: review of the general methodology. Rheumatol Int 2018;38(Suppl 1): 5–17.
13. Onel KB, Horton DB, Lovell DJ, et al. 2021 American College of Rheumatology guideline for the treatment of juvenile idiopathic arthritis: recommendations for nonpharmacologic therapies, medication monitoring, immunizations, and imaging. Arthritis Rheumatol 2022;74(4):570–585.
14. Giannini EH, Ruperto N, Ravelli A, et al. Preliminary definition of improvement in juvenile arthritis. Arthritis Rheum 1997;40(7):1202–1209.
15. Morgan EM, Munro JE, Horonjeff J, et al. Establishing an updated core domain set for studies in juvenile idiopathic arthritis: a report from the OMERACT 2018 JIA workshop. J Rheumatol 2019;46(8):1006–1013.
16. US Food and Drug Administration. Guidance for Industry. Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. December 2009. Accessed March 13, 2024. https://www.fda.gov/downloads/drugs/guidances/ucm193282.pdf
17. Mercieca-Bebber R, King MT, Calvert MJ, et al. The importance of patient-reported outcomes in clinical trials and strategies for future optimization. Patient Relat Outcome Meas 2018;9:353–367.
18. Lovell DJ, Passo MH, Beukelman T, et al. Measuring process of arthritis care: a proposed set of quality measures for the process of care in juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2011;63(1): 10–16.
19. Hersh AO, Salimian PK, Weitzman ER. Using patient-reported outcome measures to capture the patient's voice in research and care of juvenile idiopathic arthritis. Rheum Dis Clin North Am 2016;42(2): 333–346.
20. van Dijkhuizen EHP, Ridella F, Naddei R, et al; Paediatric Rheumatology International Trials Organisation (PRINTO). Validity and reliability of four parent/patient-reported outcome measures for juvenile idiopathic arthritis remote monitoring. Arthritis Care Res (Hoboken) 2023;75(2):391–400.
21. Xeljanz. Summary of product characteristics. European Medicines Agency. Accessed January 25, 2023. https://www.ema.europa.eu/en/documents/product-information/xeljanz-epar-product-information_en.pdf
22. Xeljanz. Highlights of prescribing information. Pfizer; 2024. Accessed January 23, 2025. https://labeling.pfizer.com/showlabeling.aspx?id=959
23. A safety, efficacy and pharmacokinetics study of tofacitinib in pediatric patients with sJIA. ClinicalTrials.gov identifier: NCT03000439. Updated June 3, 2025. Accessed February 22, 2023. https://clinicaltrials.gov/ct2/show/NCT03000439
24. Ruperto N, Brunner HI, Synoverska O, et al; Paediatric Rheumatology International Trials Organisation (PRINTO) and Pediatric Rheumatology Collaborative Study Group (PRCSG). Tofacitinib in juvenile idiopathic arthritis: a double-blind, placebo-controlled, withdrawal phase 3 randomised trial. Lancet 2021;398(10315):1984–1996.
25. Petty RE, Southwood TR, Manners P, et al; International League of Associations for Rheumatology. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. J Rheumatol 2004;31(2):390–392.
26. Landgraf JM. Child Health Questionnaire (CHQ). In: Michalos AC, ed. Encyclopedia of Quality of Life and Well-Being Research. 2nd ed. Springer; 2014:698–702.
27. Brunner HI, Klein-Gitelman MS, Miller MJ, et al. Minimal clinically important differences of the childhood health assessment questionnaire. J Rheumatol 2005;32(1):150–161.
28. Dempster H, Porepa M, Young N, et al. The clinical meaning of functional outcome scores in children with juvenile arthritis. Arthritis Rheum 2001;44(8):1768–1774.
29. Behrens F, Prothmann U, Klopsch T, et al. Effectiveness, safety and quality of life with tofacitinib treatment in adult patients with rheumatoid arthritis under routine clinical care: first interim results from a German non-interventional, prospective, multi-center study. Abstract presented at: ACR Convergence; November 8, 2020. Arthritis Rheumatol 2020; 72(suppl 10). https://acrabstracts.org/abstract/effectiveness-safety-and-quality-of-life-with-tofacitinib-treatment-in-adult-patients-with-rheumatoid-arthritis-under-routine-clinical-care-first-interim-results-from-a-german-non-interventional-pr/
30. Shenoi S, Horneff G, Cidon M, et al. The burden of systemic juvenile idiopathic arthritis for patients and caregivers: an international survey and retrospective chart review. Clin Exp Rheumatol 2018;36(5): 920–928.
31. Lundberg V, Lindh V, Eriksson C, et al. Health-related quality of life in girls and boys with juvenile idiopathic arthritis: self- and parental reports in a cross-sectional study. Pediatr Rheumatol Online J 2012; 10(1):33.
32. Sawyer MG, Whitham JN, Roberton DM, et al. The relationship between health-related quality of life, pain and coping strategies in juvenile idiopathic arthritis. Rheumatology (Oxford) 2004;43(3): 325–330.
33. Sanzo M. The child with arthritis in the school setting. J Sch Nurs 2008;24(4):190–196.
34. Brunner HI, Akikusa JD, Al-Abadi E, et al; Paediatric Rheumatology International Trials Organisation (PRINTO) and Pediatric Rheumatology Collaborative Study Group (PRCSG). Safety and efficacy of tofacitinib for the treatment of patients with juvenile idiopathic arthritis: preliminary results of an open-label, long-term extension study. Ann Rheum Dis 2024;83(11):1561–1571.

Arthritis & Rheumatology
Vol. 78, No. 3, March 2026, pp 763

DOI 10.1002/art.43361

Clinical Images: Vacuoles, E1 enzyme, X-linked, autoinflammatory, somatic syndrome presenting as aseptic pustular dermatosis: A rare clinical manifestation



The patient, a 79-year-old man, presented with acute-onset fever, bilateral anterior nodular scleritis, and a pustular skin rash. He had been diagnosed with vacuoles, E1 enzyme, X-linked, autoinflammatory, somatic syndrome (VEXAS) eight months earlier, with concomitant myelodysplastic syndrome. He was initially treated with prednisone and tocilizumab, with ruxolitinib introduced five months prior. Lastly, 5-azacytidine was added two months before presentation because of uncontrolled inflammatory manifestations. (A and B) On clinical examination, large, widespread, nonfollicular pustules were observed. (C, D, and E) Skin biopsies from two separate lesions revealed a dense neutrophilic infiltrate in the superficial dermis containing poorly differentiated neutrophils similar to band neutrophils and prominent leukocytoclasia, with no evidence of vasculitis. No cytopathic effect of the herpesvirus group was observed, and the immunohistochemical staining result for herpes simplex virus was negative. Swab cultures from pustular lesions were sterile. Laboratory investigations showed elevated C-reactive protein (69.0 mg/L; reference, <5 mg/L), anemia (8 g/dL; reference, 13–17 g/dL), thrombocytopenia (70 × 10^9/L; reference, 150–450 × 10^9/L), and monocytopenia (0.16 × 10^9/L; reference, 0.2–1 × 10^9/L). Because of negative infectious studies and consistent histopathological findings, a diagnosis of VEXAS-associated pustular neutrophilic dermatosis was made. High-dose prednisone and intravenous tocilizumab led to rapid resolution of systemic inflammation and cutaneous lesions. Neutrophilic dermatoses are a hallmark of VEXAS, being the most commonly reported histopathological patterns in VEXAS skin lesions.[1] Aseptic pustules, uncommon manifestations of neutrophilic dermatoses, have been described in 7% to 13% of VEXAS cases.[2,3] Clinicians should consider underlying VEXAS in patients presenting with atypical neutrophilic dermatosis, particularly when suggestive pathologic features are present.

1. Tan IJ, Ferrada MA, Ahmad S, et al. Skin manifestations of VEXAS syndrome and associated genotypes. JAMA Dermatol 2024;160(8):822–829.
2. Zakine È, Papageorgiou L, Bourguiba R, et al; National French VEXAS Study Group (NFVS). Clinical and pathological features of cutaneous manifestations in VEXAS syndrome: A multicenter retrospective study of 59 cases. J Am Acad Dermatol 2023;88(4):917–920.
3. Hines AS, Mohandesi NA, Lehman JS, et al. Cutaneous involvement in VEXAS syndrome: Clinical and histopathologic findings. Int J Dermatol 2023;62(7):938–945.

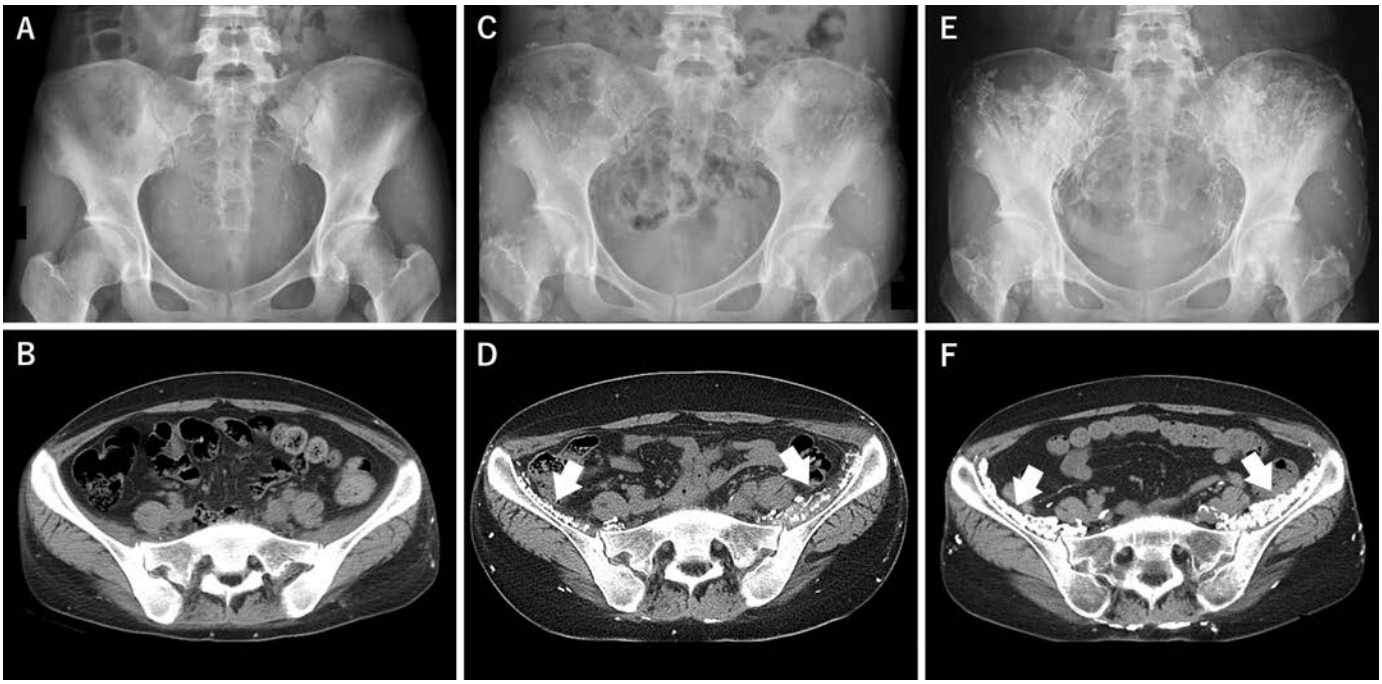
Peter Chen, MSc
peter.chen@aphp.fr
Héloïse Briançon, MSc
Deborah Eshagh, MD
Tali-Anne Szwebel, MD
Benjamin Terrier, MD, PhD
benjamin.terrier@aphp.fr
Pierre Sohier, MD, PhD
Rudy Birsen, MD, PhD
Cochin Hospital, AP-HP
Paris, France

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis & Rheumatology
Vol. 78, No. 3, March 2026, pp 764–765

DOI 10.1002/art.43368

Clinical Images: Progressive calcinosis cutis universalis involving the intrapelvic muscles in systemic lupus erythematosus



The patient, a 43-year-old woman with a 30-year history of systemic lupus erythematosus (SLE), presented with progressive systemic calcification. She was diagnosed with SLE in 1994 after symptoms of fever, malar rash, photosensitivity, arthralgia, and positive serologic markers, including antinuclear and anti–double-stranded DNA antibodies. (A and B) A pelvic radiograph and computed tomography (CT) scan incidentally revealed subcutaneous calcifications in 2006. Subsequently, these calcifications spread to the buttocks and lower extremities. (C and D) Additional calcifications were detected (indicated by arrows) in the iliacus muscle in 2013. Despite immunosuppressive therapy, she had experienced six SLE flares by 2018, including neuropsychiatric SLE and lupus panniculitis. Because of buttock pain, surgical excision of some subcutaneous calcifications was performed. Subsequent histologic examination revealed fat necrosis with calcification. Serum calcium, phosphorus, parathyroid hormone, and vitamin D levels were normal. No evidence of myositis or systemic sclerosis was detected. Calcinosis cutis universalis complicated with SLE was diagnosed. (E and F) By 2024, the intrapelvic and subcutaneous calcifications (indicated by arrows) had further progressed. Calcinosis cutis is typically associated with systemic sclerosis or myositis but is less associated with SLE.[1] Although calcinosis cutis universalis is extremely rare in patients with SLE, it can be associated with a history of chronic SLE activity.[2] However, no cases of SLE and calcinosis cutis involving the intrapelvic muscles have been reported. Although plain radiographs are effective in detecting calcinosis cutis,[3] this case highlights that CT can precisely detect and monitor progressive calcification in deeper tissues than on the skin and subcutaneous tissues. Currently, no established treatment exists for calcinosis cutis. No new data were generated or analyzed in support of this research. We obtained written informed consent from the patient for publication.

No specific funding was received from any bodies in the public, commercial, or not-for-profit sectors to carry out the work described in this article. The author would like to thank Editage (www.editage.com) for English language editing.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/art.43368.

1. Balin SJ, Wetter DA, Andersen LK, et al. Calcinosis cutis occurring in association with autoimmune connective tissue disease: The Mayo Clinic experience with 78 patients, 1996–2009. Arch Dermatol 2012;148(4):455–462.
2. Khudadah M, Jawad A, Pyne D. Calcinosis cutis universalis in a patient with systemic lupus erythematosus: A case report. Lupus 2020;29: 1630–1632.
3. Shahi V, Wetter DA, Howe BM, et al. Plain radiography is effective for the detection of calcinosis cutis occurring in association with autoimmune connective tissue disease. Br J Dermatol 2014;170: 1073–1079.

Takeshi Zoshima, MD, PhD
takeshiz1983@yahoo.co.jp
Yasunori Iwata, MD, PhD
Department of Nephrology and Rheumatology
Kanazawa University Hospital
Kanazawa, Japan

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis & Rheumatology
Vol. 78, No. 3, March 2026, pp 766–780

LETTER

DOI 10.1002/art.43148

Concordance between ultrasound and physical examination in enthesitis assessment: a call for further investigation: comment on the article by Di Matteo et al

To the Editor:

We read with great interest the article by Di Matteo et al titled "Relationship Between Ultrasound and Physical Examination in the Assessment of Enthesitis in Patients With Spondyloarthritis: Results From the DEUS Multicenter Study."[1] This pioneering multicenter study provides crucial insights into the concordance between physical examination and ultrasound for enthesitis evaluation. However, some limitations warrant further discussion.

Firstly, implementing dual-observer assessment for both physical examination and ultrasound evaluation of enthesitis, with independent evaluations by two trained clinicians, could substantially improve diagnostic precision. This approach would help to minimize subjectivity and enhance the precision of the findings. Regarding cross-center variability, centralized image interpretation using standardized scoring protocols might better mitigate intersite discrepancies compared to web-based consensus exercises alone.

Secondly, the article points out that more than two-thirds of patients with spondyloarthritis (SpA) exhibit clinical signs of enthesitis but do not show evidence of "active enthesitis" on ultrasound. Currently, the existing literature[2] does not provide strong evidence for a consistent correlation between magnetic resonance imaging (MRI) and ultrasound in the evaluation of enthesitis, primarily because of the limited research in this area. We propose that further investigation using MRI in this specific patient subgroup could be instrumental in definitively ruling out the presence of active enthesitis.

Thirdly, enthesitis is a dynamic pathologic process, and the correlation between physical examination and ultrasound findings may vary over time. The cross-sectional design of the current study restricts the ability to explore this temporal aspect. Future research should prioritize longitudinal studies to assess whether this consistency changes over time.

Lastly, the study possesses significant clinical implications, as ultrasound examination of peripheral entheses in patients with SpA is both convenient and readily integrated into routine clinical practice for widespread screening. However, the study does not address the assessment of axial entheses. Given the anatomic complexity of the axial spine, it is plausible that the agreement between ultrasound and physical examination for axial enthesitis may be less robust. Therefore, future investigations should aim to develop comprehensive assessment tools and further evaluate the consistency between imaging modalities and physical examination.

In summary, this study enhances our understanding of the consistency between physical examination and ultrasound in assessing enthesitis. We advocate for further research to explore this consistency.

Supported by Weifang Health Commission's scientific research program (grant WFWSJK-2023-222 and WFWSJK-2023-240) and the Weifang Youth Medical Talent lift project.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/art.43148.

Danxiang Shen and Di Jin were dual first authors and contributed equally to this work.

Danxiang Shen, MMed
Changxing Third People's Hospital
Huzhou, Zhejiang, China
Zhejiang Provincial People's Hospital, Affiliated People's Hospital, Hangzhou Medical College
Hangzhou, Zhejiang, China
Di Jin, MMed
Weifang People's Hospital
Shandong, China
Hongye Xu, MBBS
Yang Jia Bu Community Health Service Center
Huzhou, Zhejiang, China
James Cheng-Chung Wei, MD, PhD
jccwei@gmail.com
Chung Shan Medical University Hospital
Taichung, Taiwan
Zhenhua Ying, MD
yingzh2021@163.com
Zhejiang Provincial People's Hospital, Affiliated People's Hospital, Hangzhou Medical College
Hangzhou, Zhejiang, China

1. Di Matteo A, Di Donato S, Smerilli G, et al. Relationship between ultrasound and physical examination in the assessment of enthesitis in patients with spondyloarthritis: results from the DEUS multicenter study. Arthritis Rheumatol 2025;77(1):22–33.
2. Aguila Maldonado R, Ruta S, Valuntas ML, et al. Ultrasonography assessment of heel entheses in patients with spondyloarthritis: a comparative study with magnetic resonance imaging and conventional radiography. Clin Rheumatol 2017;36(8):1811–1817.

DOI 10.1002/art.43150

Reply

To the Editor:

We are grateful to Danxiang Shen et al for their interest in our study and for providing valuable insights into assessing enthesitis in patients with spondyloarthritis (SpA). In their letter, they clearly highlight key areas for improvement and propose potential solutions. In essence, they outline a research agenda with which we fully agree.

Increasing the number of clinicians to reduce subjective assessment, centralizing image interpretation to mitigate intersite discrepancies, incorporating magnetic resonance imaging as an independent "third eye" alongside clinical and ultrasound evaluations, and using longitudinal data to better interpret baseline findings are all essential steps to enhance our clinical knowledge and performance.

We believe that all available tools (ie, clinical and imaging examinations) that provide information on entheseal status should be considered, and an index integrating the best aspects of each tool should be developed. Although clinical examination cannot visualize the different pathologies causing tenderness, imaging cannot assess symptoms such as pain. In our study, the greatest discrepancy between ultrasound and physical examination findings was observed at the Achilles tendon.[1] This tendon is traditionally a key site for evaluating enthesitis in SpA.[2] Furthermore, a recent study by our group identified the Achilles tendon enthesis as crucial in distinguishing SpA-related inflammatory enthesitis from mechanical enthesitis (ie, noninflammatory) via ultrasound.[3]

In conclusion, we appreciate the insights of Danxiang Shen et al and fully support their research agenda. Further studies should explore the relationship between physical examination and ultrasound, particularly at the Achilles tendon. Developing a comprehensive index that combines clinical and imaging findings may improve enthesitis assessment.

Andrea Di Matteo, MD, PhD
andrea.dimatteo@hotmail.com
"Carlo Urbani" Hospital
and Polytechnic University of Marche
Ancona, Italy
and University of Leeds
Leeds, United Kingdom
Leeds Institute of Rheumatic
and Musculoskeletal Medicine
University of Leeds
Leeds, United Kingdom
Stefano Di Donato, MD
Leeds Institute of Rheumatic
and Musculoskeletal Medicine
University of Leeds
Leeds, United Kingdom
Emilio Filippucci, MD, PhD
"Carlo Urbani" Hospital
and Polytechnic University of Marche
Ancona, Italy
and University of Leeds
Leeds, United Kingdom

1. Di Matteo A, Di Donato S, Smerilli G, et al. Relationship between ultrasound and physical examination in the assessment of enthesitis in patients with spondyloarthritis: results from the DEUS multicenter study. Arthritis Rheumatol 2025;77(1):22–33.
2. Zabotti A, Piga M, Canzoni M, et al; UPSTREAM study group. Ultrasonography in psoriatic arthritis: which sites should we scan? Ann Rheum Dis 2018;77(10):1537–1538.
3. Di Matteo A, Smerilli G, Di Donato S, et al. Power Doppler signal at the enthesis and bone erosions are the most discriminative OMERACT ultrasound lesions for SpA: results from the DEUS (Defining Enthesitis on Ultrasound in Spondyloarthritis) multicentre study. Ann Rheum Dis 2024;83(7):847–857.

DOI 10.1002/art.43369

JAK inhibitors did not increase the risk of major adverse cardiovascular events in patients with rheumatoid arthritis within the first two years of treatment: comment on the article by Aymon et al

To the Editor:

We read with great interest the article by Aymon et al, a multinational study that provides valuable real-world evidence on the cardiovascular safety of JAK inhibitors (JAKi) compared to biologic disease-modifying antirheumatic drugs to treat rheumatoid arthritis (RA).[1] Although the multinational cohort design and robust data are commendable, several limitations warrant discussion to guide future research and clinical practice; we would like to raise several critical points for further discussion.

First, the study did not address the use of nonsteroidal anti-inflammatory drugs (NSAIDs), which are widely prescribed for pain management in RA. It was reported that all NSAIDs increase the risk of cardiovascular diseases, and their cardiovascular risks depend on the degree of inhibition of cyclooxygenase 1 (COX-1) and COX-2: the smaller the degree of COX-1 inhibition, the greater the degree of COX-2 inhibition, and the greater the cardiovascular risk.[2] The study lacked granularity on NSAID classes, dosing, or duration. Stratified analyses incorporating these variables could elucidate whether residual confounding explains these results.

Second, although the study adjusts for prevalent comorbidities (eg, hypertension, diabetes, and hyperlipidemia), it omits key contemporary therapies such as antihypertensive drugs, glucagon-like peptide 1 (GLP1) receptor agonists,[3] or statins,[4] which independently modify cardiovascular risk. Propensity-score matching or adjustment for medication adherence could mitigate this residual confounding. The study only recorded the first event and did not consider the impact of recurrent events or multiple events of major adverse cardiovascular events, which may have underestimated the long-term risk.

Lastly, the study population was predominantly European (Switzerland, Germany, and Scandinavia) and Canadian, with no

representation from Asia, Africa, or South America. Regional disparities in health care access (eg, JAKi availability), genetic susceptibility (eg, ethnic differences in drug metabolism), and cardiovascular risk profiles may limit generalizability. Multiregional collaborations, such as inclusion of Asian registries, are critical to validate these findings globally.

In conclusion, this work underscores the short-term cardiovascular safety of JAKi in routine care. However, addressing these limitations would strengthen the robustness and applicability of future studies. We look forward to the authors' feedback and their efforts to enhance our understanding of this important topic.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/art.43369.

Hao Xing, MD
Shanxi Bethune Hospital, Shanxi Academy of Medical Sciences, Third Hospital of Shanxi Medical University
Tongji Shanxi Hospital
Taiyuan, China
Ya-zhen Su, MD, PhD
suyazhen@sxbqeh.com.cn
Shanxi Province Clinical Research Center for Dermatologic and Immunologic Diseases (Rheumatic Diseases)
Shanxi Bethune Hospital
Taiyuan, China

1. Aymon R, Mongin D, Guemara R, et al. Incidence of major adverse cardiovascular events in patients with rheumatoid arthritis treated with JAK inhibitors compared to biologic disease-modifying antirheumatic drugs: data from an international collaboration of registries. Arthritis Rheumatol 2025;77(9):1194–1204.
2. Roubille C, Richer V, Starnino T, et al. The effects of tumor necrosis factor inhibitors, methotrexate, non-steroidal anti-inflammatory drugs and corticosteroids on cardiovascular events in rheumatoid arthritis, psoriasis and psoriatic arthritis: a systematic review and meta-analysis. Ann Rheum Dis 2015;74(3):480–489.
3. Marso SP, Daniels GH, Brown-Frandsen K, et al; LEADER Steering Committee; LEADER Trial Investigators. Liraglutide and cardiovascular outcomes in Type 2 diabetes. N Engl J Med 2016;375(4):311–322.
4. Oesterle A, Laufs U, Liao JK. Pleiotropic effects of statins on the cardiovascular system. Circ Res 2017;120(1):229–243.

DOI 10.1002/art.43372

Reply

To the Editor:

We thank Hao Xing and Ya-Zhen Su for their interest in our study[1] and for the opportunity to clarify several methodologic aspects. Regarding exposure to nonsteroidal anti-inflammatory drugs (NSAIDs), we agree that these drugs may influence the cardiovascular risk. However, NSAIDs are often obtained over the counter, and their use is not consistently recorded, even when registries attempt to capture their use. As a result, routine capture of NSAIDs, dosage, and frequency of use was incomplete across the 15 contributing registries, and thus adjustment was not feasible. Importantly, because NSAID use in routine care is unrelated to the choice of JAK inhibitor (JAKi) versus biologic disease-modifying antirheumatic drug (bDMARD), any misclassification should be nondifferential, resulting in an attenuation of potential differences between bDMARDs, rather than creating treatment differences.

Xing et al also raise the important issue of concomitant cardioprotective medications, such as statins, renin–angiotensin system inhibitors, sodium-dependent d-glucose cotransporter 2 inhibitors, or glucagon-like peptide 1 receptor agonists. Because our study relies on rheumatology-specific registries, data on nonrheumatologic medications are often incomplete, particularly in countries without linkage to nationwide prescription databases. Only a minority of countries contributing to our study have access to such comprehensive data. Therefore, residual confounding is possible, but again, misclassification should be nondifferential because their use is unlikely to be systematically associated with the choice of JAKi versus bDMARDs. That said, a key strength of using rheumatoid arthritis–specific registries is the availability of clinical data not typically present in claims datasets, including measures of disease activity, disease duration and severity, and lifestyle factors, such as smoking status, all of which are important confounders for cardiovascular outcomes.

With respect to recurrent major adverse cardiovascular events (MACEs), all registries record event dates, allowing distinction between the first and subsequent MACEs. Only one patient experienced a second event during follow-up in the combined analysis (0.4% of 282 total MACEs). The scarcity of recurrences within two years supports our focus on the first MACE.

We agree that inclusion of data from underrepresented regions such as Asia, Africa, or South America would enhance geographic generalizability. Although our study currently includes predominantly European and one Canadian (Québec) registries, this already entails considerable differences in health care systems, coding systems, and baseline patient demographics. One of the major strengths of the JAK-pot collaboration is the use of several methods to account for this heterogeneity. Extending the collaboration to Asia, Africa, or South America would undoubtedly enhance external validity and help contextualize the Oral Rheumatoid Arthritis Trial (ORAL) Surveillance signal but would also introduce additional between-registry variability that must be handled with care. Early real-world data from Japan,[2] China,[3] and Korea[4] have also suggested no excess risk of MACE with baricitinib or tofacitinib. We strongly encourage similar collaborative efforts in these regions.

In conclusion, despite unavoidable limitations inherent to real-world pharmacovigilance, the robustness of our findings, consistent across two analytical frameworks and 15 health care systems and after accounting for confounding, supports the cardiovascular neutrality of JAKi as they are currently used in routine clinical care in countries with universal health coverage during the first two years

of treatment. Ongoing follow-up and broader geographic inclusion will further refine long-term safety estimates.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/art.43372.

On behalf of the JAK-pot collaboration:
Romain Aymon, MSc
romain.aymon@hug.ch
Delphine S. Courvoisier, PhD
Axel Finckh, MD, PhD
Kim Lauper, MD, PhD
Rheumatology Division
Geneva University Hospitals
Geneva, Switzerland

1. Aymon R, Mongin D, Guemara R, et al. Incidence of major adverse cardiovascular events in patients with rheumatoid arthritis treated with JAK inhibitors compared to biologic disease-modifying antirheumatic drugs: data from an international collaboration of registries. Arthritis Rheumatol 2025;77(9):1194–1204.
2. Okamoto N, Atsumi T, Takagi M, et al. Safety of baricitinib in Japanese patients with rheumatoid arthritis in clinical use: 3-year data of all-case postmarketing surveillance study. Mod Rheumatol 2025;35(2):215–224.
3. Wu CY, Wang Q, Shi J, et al. Safety and effectiveness of baricitinib in Chinese patients with moderate-to-severe rheumatoid arthritis: 24-week results from a post-marketing safety study. Rheumatol Ther 2023;10(6):1609–1622.
4. Song YK, Lee G, Hwang J, et al. Cardiovascular risk of Janus kinase inhibitors compared with biologic disease-modifying antirheumatic drugs in patients with rheumatoid arthritis without underlying cardiovascular diseases: a nationwide cohort study. Front Pharmacol 2023;14:1165711.

DOI 10.1002/art.43373

Trimethoprim-sulfamethoxazole prophylaxis in antineutrophil cytoplasmic antibody-associated vasculitis: watch out for pseudoacute kidney injury. Comment on the article by Kim et al

To the Editor:

We read with great interest the article by Kim et al describing the beneficial effect of trimethoprim-sulfamethoxazole (T/S) on the prevention of infectious events in antineutrophil cytoplasmic antibody-associated vasculitis (AAV).[1] Cohen Tervaert commented on this interesting multicenter cohort study,[2] highlighting the significant role of T/S use not only for the prevention of *Pneumocystis jirovecii* and no-*Pneumocystis jirovecii* pneumonia infections, but also for the prevention of vasculitis relapses.[3]

Renal adverse effects of T/S treatment are also remarkable: (1) hyperkalemia, which is caused by the amiloride-like inhibition of renal outer medullary potassium channels, mediated by the S component, leading to the inhibition of sodium absorption in the distal tubule and the reduction of potassium tubular secretion;[4] (2) acute interstitial nephritis, documented by biopsy and clinically occurring with fever, malaise, eosinophilia, eosinophiluria, leucocyte casts, and progressive renal tubular dysfunction;[5] (3) nephrotoxicity, likely due to the S component when used at high doses or at doses inappropriate for glomerular filtration rate (GFR); and (4) pseudoacute kidney injury (pseudo-AKI), likely due to the T component effect which, even at recommended doses, competes with the organic cation transporter 2, inhibiting tubular creatinine secretion, leading to a rapid but reversible increase in serum creatinine independent from any changes in GFR levels. The latter effect can cause a false GFR level underestimation when creatinine-based equations are used, potentially precluding the use of therapies, currently not recommended for GFR levels less than 15 mL/min (ie, avacopan).[6] In fact, as recently showed, the simultaneous use of creatinine and cystatin C estimated GFR levels distinguishes between a true nephrotoxicity or a S/T related pseudo-AKI.[7]

Given these considerations and the latest Kidney Disease: Improving Global Outcomes guidelines,[8] we strongly recommend for all patients affected by AAV receiving T/S prophylaxis to closely monitor serum ions and urine dipstick and use both formulas for the GFR equations (ie, cystatin C and creatinine-based formulas), to estimate the renal function in order to promptly recognize and adequately manage the renal adverse effect of T/S prophylaxis and to distinguish actual kidney involvement from pseudo-AKI.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/art.43373.

Ivana Capuano, MD, PhD
ivana.capuano@unina.it
Eleonora Riccio, MD, PhD
Antonio Pisani, MD, PhD
Department of Public Health,
University of Naples Federico II
Naples, Italy
Pasquale Buonanno, MD, PhD
Department of Neurosciences, Reproductive and Odontostomatological Sciences,
University of Naples Federico II
Naples, Italy

1. Kim YK, Curtis JR, Choi SR, et al. Time-dependent effect of prophylactic trimethoprim-sulfamethoxazole on the incidence of serious infections in antineutrophil cytoplasmic antibody-associated vasculitis: a target trial emulation study. Arthritis Rheumatol 2025;77(10):1407–1415.
2. Cohen Tervaert JW. Prophylactic trimethoprim-sulfamethoxazole in antineutrophil cytoplasmic autoantibody-associated vasculitis: comment on the article by Kim et al. Arthritis Rheumatol 2026;78(1):260–261. https://doi.org/10.1002/art.43296
3. Stegeman CA, Tervaert JW, de Jong PE, et al; Dutch Co-Trimoxazole Wegener Study Group. Trimethoprim-sulfamethoxazole (co-trimoxazole) for the prevention of relapses of Wegener's granulomatosis. N Engl J Med 1996;335(1):16–20.
4. Yokoyama S, Nakagawa J, Aiuchi N, et al. Impact of trimethoprim on serum creatinine, sodium, and potassium concentrations in patients taking trimethoprim-sulfamethoxazole without changes in glomerular filtration rate. J Clin Pharm Ther 2022;47(9):1409–1417.

5. Cryst C, Hammar SP. Acute granulomatous interstitial nephritis due to co-trimoxazole. Am J Nephrol 1988;8(6):483–488.
6. Delanaye P, Mariat C, Cavalier E, et al. Trimethoprim, creatinine and creatinine-based equations. Nephron Clin Pract 2011;119(3):c187–c193.
7. Urakami T, Hamada Y, Oka Y, et al. Is trimethoprim/sulfamethoxazole-associated increase in serum creatinine a pseudo-elevation or true nephrotoxicity? J Infect Chemother 2021;27(8):1193–1197.
8. Kidney Disease: Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2024 clinical practice guideline for the evaluation and management of chronic kidney disease. Kidney Int 2024;105(4S): S117–S314.

DOI 10.1002/art.43374

Reply

To the Editor:

We appreciate the valuable comment by Capuano et al on our recent article examining the time-dependent effect of prophylactic trimethoprim-sulfamethoxazole (TMP-SMX) on the incidence of serious infections in antineutrophil cytoplasmic antibody-associated vasculitis (AAV).[1] They also pointed out the potential renal adverse effects of TMP-SMX and emphasized the clinical importance of monitoring both cystatin C and creatinine.

We agree on the importance of close monitoring of renal function and serum electrolyte levels to minimize the occurrence of TMP-SMX-related renal adverse drug reactions. A metabolite of sulfamethoxazole could be crystalized in acidic urine, which leads to tubular damage. In addition, trimethoprim can impair electrolyte transport and inhibit creatinine secretion in the renal tubule.[2,3] These effects are main mechanisms that cause renal adverse reactions of TMP-SMX and are dose-dependent. In fact, several previous studies have suggested that prophylactic doses of TMP-SMX are associated with lower risk of renal toxicities compared with therapeutic doses.[4,5] A recent study from our group also showed that incidences of hyperkalemia and azotemia directly related to prophylactic dose of TMP-SMX are uncommon (1.2 and 2.0 per 100 person-years, respectively).[6] Therefore, we emphasize that excessive concern about its relatively benign renal safety profile may cause physicians to hesitate in using it, which could potentially worsen patient outcomes.

Accumulating evidence suggests that cystatin C-based glomerular filtration rate (GFR) estimation is more accurate than creatinine-based methods. However, in many real-world settings, the use of cystatin C testing is still low, and test availability may be limited in some countries.[7] Furthermore, although cystatin C levels are less influenced by muscle mass, race and functional status, they can be elevated by steroid use and inflammation, potentially leading to underestimation of GFR in patients with AAV receiving induction therapy.[8–10] Therefore, we believe that the accuracy of cystatin C-based GFR estimation in these patients should be further validated in future studies.

Establishing an optimal prophylactic strategy requires accurate risk-benefit assessment. Our study showed that prophylactic TMP-SMX has additional benefit beyond the prevention of *Pneumocystis jirovecii* pneumonia (PJP) in patients with AAV. Vigilant monitoring for adverse drug reactions during the prophylaxis ensures a more favorable safety profile. However, the optimal strategy for such monitoring remains an important area for future research.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/art.43374.

Yun Kyu Kim, MD
Jun Won Park, MD
mpersonality@gmail.com
Division of Rheumatology, Department of Internal Medicine
Seoul National University College of Medicine
Seoul, Republic of Korea
Jeffrey R. Curtis, MD, MS, MPH
Division of Clinical Immunology & Rheumatology
University of Alabama at Birmingham
Birmingham, Alabama
Eun Bong Lee, MD, PhD
Division of Rheumatology, Department of Internal Medicine,
Seoul National University College of Medicine
and Department of Molecular Medicine and
Biopharmaceutical Sciences, Graduate School of
Convergence Science and Technology, Seoul National University
Seoul, Republic of Korea

1. Kim YK, Curtis JR, Choi SR, et al. Time-dependent effect of prophylactic trimethoprim-sulfamethoxazole on the incidence of serious infections in antineutrophil cytoplasmic antibody-associated vasculitis: a target trial emulation study. Arthritis Rheumatol 2025;77(10):1407–1415.
2. Gentry CA, Nguyen AT. An evaluation of hyperkalemia and serum creatinine elevation associated with different dosage levels of outpatient trimethoprim-sulfamethoxazole with and without concomitant medications. Ann Pharmacother 2013;47(12):1618–1626.
3. Tsapepas D, Chiles M, Babayev R, et al. Incidence of hyponatremia with high-dose trimethoprim-sulfamethoxazole exposure. Am J Med 2016;129(12):1322–1328.
4. Kawato R, Rokutanda R, Okada M, et al. Effect of prophylactic dose of trimethoprim-sulfamethoxazole on serum creatinine in Japanese patients with connective tissue diseases. Clin Med Insights Arthritis Musculoskelet Disord 2022;15:11795441221085141.
5. Park JW, Curtis JR, Choi SR, et al. Risk-benefit analysis of primary prophylaxis against *Pneumocystis jirovecii* pneumonia in patients with rheumatic diseases receiving rituximab. Arthritis Rheumatol 2023; 75(11):2036–2044.
6. Park JW, Curtis JR, Jun KI, et al. Primary prophylaxis for *Pneumocystis jirovecii* pneumonia in patients receiving rituximab. Chest 2022; 161(5):1201–1210.
7. Chen DC, Potok OA, Rifkin D, et al. Advantages, limitations, and clinical considerations in using cystatin C to estimate GFR. Kidney360 2022;3(10):1807–1814.
8. Stevens LA, Schmid CH, Greene T, et al. Factors other than glomerular filtration rate affect serum cystatin C levels. Kidney Int 2009;75(6):652–660.
9. Knight EL, Verhave JC, Spiegelman D, et al. Factors influencing serum cystatin C levels other than renal function and the impact on renal function measurement. Kidney Int 2004;65(4):1416–1421.

10. Kleeman SO, Thakir TM, Demestichas B, et al. Cystatin C is glucocorticoid responsive, directs recruitment of Trem2+ macrophages, and predicts failure of cancer immunotherapy. Cell Genom 2023;3(8):100347.

DOI 10.1002/art.43385

"Negative" finding supports two-hit model: comment on the article by Kroese et al

To the Editor:

The recent communication by Kroese et al[1] was of great interest. They report on a cohort of 50 individuals who were at risk for rheumatoid arthritis (RA; defined as anticitrullinated peptide antibody [ACPA] and/or rheumatoid factor [RF]-positive arthralgia) and in whom they had previously shown differences in the composition of the oral microbiome when compared to healthy individuals.[2] They now inform us that eight individuals out of this group developed RA, and these were compared with the 33 who did not (for the remainder, follow-up was limited); in this comparison, no differences in the oral microbiome were seen.

Although these results can be viewed as "negative," it will certainly not have escaped the investigators' notice that these results lend strong support to the "second hit" hypothesis of RA pathogenesis[3]: the view that the development of autoimmunity as reflected by ACPA and/or RF, which is strongly linked to the shared epitope, is the result of a process at the mucosal surfaces, whereas the development of the clinical disease—sometimes years later—is dependent on a second hit, the nature of which has been strikingly elusive. Having all but ruled out genetic, epigenetic, bacterial, common viral, environmental, dietary, or neuropsychological mechanisms for this second hit, one is left to speculate that we are dealing with an unknown infectious trigger or, perhaps least satisfactorily, a purely stochastic event.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/art.43385.

Ronald van Vollenhoven, MD, PhD
r.vanvollenhoven@amsterdamumc.nl
Amsterdam University Medical Center
Amsterdam, the Netherlands

1. Kroese JM, Brandt BW, Buijs MJ, et al. The oral microbiota in persons at risk for rheumatoid arthritis: follow-up data for the article by Kroese et al. Arthritis Rheumatol 2025;77(6):778–780.
2. Kroese JM, Brandt BW, Buijs MJ, et al. Differences in the oral microbiome in patients with early rheumatoid arthritis and individuals at risk of rheumatoid arthritis compared to healthy individuals. Arthritis Rheumatol 2021;73(11):1986–1993.
3. Nielen MM, van Schaardenburg D, Reesink HW, et al. Specific autoantibodies precede the symptoms of rheumatoid arthritis: a study of serial measurements in blood donors. Arthritis Rheum 2004;50(2):380–386.

DOI 10.1002/art.43381

Reply

To the Editor:

We thank Ronald van Vollenhoven for bringing up the two-hit model, that is, the notion that two separate events are necessary to trigger the development of clinically apparent rheumatoid arthritis (RA). He states this hypothesis is supported by our finding[1] that the oral bacterial composition was not predictive of future RA development, even though we had previously shown that the bacterial composition in groups of persons at risk of RA and of persons with early RA was unfavorable compared to healthy controls.[2] If we understand the author of the letter correctly, the fact that oral dysbiosis was present at presentation in the at-risk phase, but was not predictive of progression to RA, would lend support to the idea that dysbiosis may be a necessary step but not a sufficient one by itself to induce progression to clinically apparent RA. Although this may be true, we feel that such a conclusion cannot be drawn from the present data.

The main reason for this is that the group of persons at risk we were able to follow was too small, which leaves the possibility open that the presence of certain bacterial species may be associated with the development of arthritis. At a more general level, the second-hit hypothesis supposes a sequence of two events, in which the first one is often thought of as the occurrence of autoantibodies, regardless if induced at mucosal surfaces. A second hit suggests a specific event such as an infection that would bring a previously subclinical inflammation to the surface. Although there are indications that (a range of) infections occur with increased frequency in the year preceding the diagnosis of RA,[3,4] the established genetic and lifestyle risk factors for RA involve long-term exposures, consistent with a gradually increasing subclinical inflammation that eventually manifests as clinical RA.[5]

We agree with van Vollenhoven that a second hit has not yet been specified. Although it is an attractive concept, the existing evidence points more to a gradually increasing effect of a combination of multiple risk factors. As such, dysbiosis may well act as an intermediate between the environment and the immunologic disorder. Finally, we need more data before we can determine whether bacterial compositions are predictive of RA.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/art.43381.

Johanna M. Kroese, DDS, PhD
Departments of Orofacial Pain and Dysfunction, Preventive Dentistry, and Periodontology, Academic Center for Dentistry Amsterdam, University of Amsterdam and Vrije Universiteit Amsterdam
Amsterdam, the Netherlands

Bernd W. Brandt, PhD
Mark J. Buijs, BE
Wim Crielaard, PhD
Egija Zaura, DDS, PhD
Department of Preventive Dentistry, Academic Center for Dentistry Amsterdam, University of Amsterdam and Vrije Universiteit Amsterdam
Amsterdam, the Netherlands
Frank Lobbezoo, DDS, PhD
Department of Orofacial pain and Dysfunction, Academic Center for Dentistry Amsterdam, University of Amsterdam and Vrije Universiteit Amsterdam
Amsterdam, the Netherlands
Bruno G. Loos, DDS, MSc, PhD
Department of Periodontology, Academic Center for Dentistry Amsterdam, University of Amsterdam and Vrije Universiteit Amsterdam
Amsterdam, the Netherlands
Laurette van Boheemen, MD, PhD
Dirkjan van Schaardenburg, MD, PhD
d.v.schaardenburg@reade.nl
Amsterdam Rheumatology and Immunology Center, Reade and Amsterdam University Medical Center
Amsterdam, the Netherlands
Catherine M.C. Volgenant, DDS, PhD
Department of Orofacial Pain and Dysfunction and Department of Preventive Dentistry, Academic Center for Dentistry Amsterdam, University of Amsterdam and Vrije Universiteit Amsterdam
Amsterdam, the Netherlands

1. Kroese JM, Brandt BW, Buijs MJ, et al. The oral microbiota in persons at risk for rheumatoid arthritis: follow-up data for the article by Kroese et al. Arthritis Rheumatol 2025;77(6):778–780.
2. Kroese JM, Brandt BW, Buijs MJ, et al. Differences in the oral microbiome in patients with early rheumatoid arthritis and individuals at risk of rheumatoid arthritis compared to healthy individuals. Arthritis Rheumatol 2021;73(11):1986–1993.
3. van Beers-Tas M, Nielen MM, Twisk JWR, et al. Increased primary care use for musculoskeletal symptoms, infections and comorbidities in the years before the diagnosis of inflammatory arthritis. RMD Open 2020;6:e001163.
4. Porter D, Jain S, Qian E, et al. Time-varying risks of infection in patients as they proceed through the phases of 'pre-RA': results from the Scottish Early RA inception cohort. RMD Open 2023;9:e003224.
5. Hahn J, Malspeis S, Choi MY, et al. Association of healthy lifestyle behaviors and the risk of developing rheumatoid arthritis among women. Arthritis Care Res (Hoboken) 2023;75(2):272–276.

DOI 10.1002/art.43147

Minimal tofacitinib dose for maintenance in axial spondyloarthritis: real-world data from a resource-limited setting

To the Editor:

Over the past decade, numerous studies have investigated the tapering of biologic disease-modifying antirheumatic drugs in axial spondyloarthritis (AxSpA),[1] and this is supported in recent recommendations for AxSpA.[2] After positive results in trials,[2] JAK inhibitors are being used as an alternative to biologics, especially in resource-poor settings, considering the lower cost of therapy. However, considering the risks of infections, cardiovascular events, cancer, and venous thromboembolism,[3] we often attempt to maintain patients with AxSpA on minimum possible doses of tofacitinib after informed consent.

This retrospective study aimed to assess the risk of flares, flare-free survival, and factors associated with flares on tapering tofacitinib in AxSpA. It was a medical records review of patients with AxSpA (Assessment of SpondyloArthritis International Society criteria)[4] on tofacitinib in remission or low disease activity (LDA) (Ankylosing Spondylitis Disease Activity Score with C-reactive protein [ASDAS-CRP] <2.1) for at least 3 months and were gradually tapered off tofacitinib. It was tapered at the treating physician's discretion (by 25%–50% of the weekly dose of 70 mg at each visit), accompanied by an evaluation of the disease activity. Tofacitinib was resumed at the previous effective dose in case of flares, defined as an increase in ASDAS-CRP[5] by 0.9.

In our cohort, among 240 patients with AxSpA who were on tofacitinib,142 were in remission on stable doses of tofacitinib for the previous 3 months. The 77 patients who met our inclusion criteria had a median age of 38 years (interquartile range [IQR] 31–46), and 20 (28.6%) were women. They had a median disease duration of 10 years (IQR 6–19). Radiographic sacroiliitis was present in 61 patients (87%), HLA-B27 positivity in 45 patients (58%), and uveitis in 11 patients (14.3%). Median ASDAS-CRP was 2.9 (IQR 2.3–3.5) at initial presentation (Supplementary Table 1). The tapering was started after they had received the standard dose of tofacitinib (10 mg/day) for a median of 4.6 months.

During the median follow-up of 259 days (IQR 98–570) since the start of tapering, out of 77 patients, 10 flared at a dose of 7.5 mg/day, another 16 flared at the dose of 5 mg/day, and another seven flared at doses less than 5 mg/day (Figure 1). All the patients who had flared reattained remission within 4 weeks on hiking the dose of tofacitinib immediately to prior effective dose. The median time to flare was 189 days (IQR 98–341). However, 44 patients (57%) remained in remission despite the tapering protocol. In univariate survival analysis, prior exposure to biologics was associated with flares (log-rank test $P = 0.035$; Supplementary Figure 1), whereas radiographic sacroiliitis (log-rank test $P = 0.076$; Supplementary Figure 2) and the presence of HLA-B27 was not (log-rank test $P = 0.75$; Supplementary Figure 3).

Tapering studies of JAK inhibitors in rheumatoid arthritis are available with lower rates of LDA, remission, nonserious infection, and higher relapse in the tapering group in a long-term extension of baricitinib phase 3 trial. However, two-thirds of the patients who lost response to the 2-mg dose regained LDA after dose escalation to 4 mg.[6] Similar results were obtained in a Japanese

Bernd W. Brandt, PhD
Mark J. Buijs, BE
Wim Crielaard, PhD
Egija Zaura, DDS, PhD
Department of Preventive Dentistry, Academic Center for Dentistry Amsterdam, University of Amsterdam and Vrije Universiteit Amsterdam
Amsterdam, the Netherlands
Frank Lobbezoo, DDS, PhD
Department of Orofacial pain and Dysfunction, Academic Center for Dentistry Amsterdam, University of Amsterdam and Vrije Universiteit Amsterdam
Amsterdam, the Netherlands
Bruno G. Loos, DDS, MSc, PhD
Department of Periodontology, Academic Center for Dentistry Amsterdam, University of Amsterdam and Vrije Universiteit Amsterdam
Amsterdam, the Netherlands
Laurette van Boheemen, MD, PhD
Dirkjan van Schaardenburg, MD, PhD
d.v.schaardenburg@reade.nl
Amsterdam Rheumatology and Immunology Center, Reade and Amsterdam University Medical Center
Amsterdam, the Netherlands
Catherine M.C. Volgenant, DDS, PhD
Department of Orofacial Pain and Dysfunction and Department of Preventive Dentistry, Academic Center for Dentistry Amsterdam, University of Amsterdam and Vrije Universiteit Amsterdam
Amsterdam, the Netherlands

1. Kroese JM, Brandt BW, Buijs MJ, et al. The oral microbiota in persons at risk for rheumatoid arthritis: follow-up data for the article by Kroese et al. Arthritis Rheumatol 2025;77(6):778–780.
2. Kroese JM, Brandt BW, Buijs MJ, et al. Differences in the oral microbiome in patients with early rheumatoid arthritis and individuals at risk of rheumatoid arthritis compared to healthy individuals. Arthritis Rheumatol 2021;73(11):1986–1993.
3. van Beers-Tas M, Nielen MM, Twisk JWR, et al. Increased primary care use for musculoskeletal symptoms, infections and comorbidities in the years before the diagnosis of inflammatory arthritis. RMD Open 2020;6:e001163.
4. Porter D, Jain S, Qian E, et al. Time-varying risks of infection in patients as they proceed through the phases of 'pre-RA': results from the Scottish Early RA inception cohort. RMD Open 2023;9:e003224.
5. Hahn J, Malspeis S, Choi MY, et al. Association of healthy lifestyle behaviors and the risk of developing rheumatoid arthritis among women. Arthritis Care Res (Hoboken) 2023;75(2):272–276.

DOI 10.1002/art.43147

Minimal tofacitinib dose for maintenance in axial spondyloarthritis: real-world data from a resource-limited setting

To the Editor:

Over the past decade, numerous studies have investigated the tapering of biologic disease-modifying antirheumatic drugs in axial spondyloarthritis (AxSpA),[1] and this is supported in recent recommendations for AxSpA.[2] After positive results in trials,[2] JAK inhibitors are being used as an alternative to biologics, especially in resource-poor settings, considering the lower cost of therapy. However, considering the risks of infections, cardiovascular events, cancer, and venous thromboembolism,[3] we often attempt to maintain patients with AxSpA on minimum possible doses of tofacitinib after informed consent.

This retrospective study aimed to assess the risk of flares, flare-free survival, and factors associated with flares on tapering tofacitinib in AxSpA. It was a medical records review of patients with AxSpA (Assessment of SpondyloArthritis International Society criteria)[4] on tofacitinib in remission or low disease activity (LDA) (Ankylosing Spondylitis Disease Activity Score with C-reactive protein [ASDAS-CRP] <2.1) for at least 3 months and were gradually tapered off tofacitinib. It was tapered at the treating physician's discretion (by 25%–50% of the weekly dose of 70 mg at each visit), accompanied by an evaluation of the disease activity. Tofacitinib was resumed at the previous effective dose in case of flares, defined as an increase in ASDAS-CRP[5] by 0.9.

In our cohort, among 240 patients with AxSpA who were on tofacitinib,142 were in remission on stable doses of tofacitinib for the previous 3 months. The 77 patients who met our inclusion criteria had a median age of 38 years (interquartile range [IQR] 31–46), and 20 (28.6%) were women. They had a median disease duration of 10 years (IQR 6–19). Radiographic sacroiliitis was present in 61 patients (87%), HLA-B27 positivity in 45 patients (58%), and uveitis in 11 patients (14.3%). Median ASDAS-CRP was 2.9 (IQR 2.3–3.5) at initial presentation (Supplementary Table 1). The tapering was started after they had received the standard dose of tofacitinib (10 mg/day) for a median of 4.6 months.

During the median follow-up of 259 days (IQR 98–570) since the start of tapering, out of 77 patients, 10 flared at a dose of 7.5 mg/day, another 16 flared at the dose of 5 mg/day, and another seven flared at doses less than 5 mg/day (Figure 1). All the patients who had flared reattained remission within 4 weeks on hiking the dose of tofacitinib immediately to prior effective dose. The median time to flare was 189 days (IQR 98–341). However, 44 patients (57%) remained in remission despite the tapering protocol. In univariate survival analysis, prior exposure to biologics was associated with flares (log-rank test $P = 0.035$; Supplementary Figure 1), whereas radiographic sacroiliitis (log-rank test $P = 0.076$; Supplementary Figure 2) and the presence of HLA-B27 was not (log-rank test $P = 0.75$; Supplementary Figure 3).

Tapering studies of JAK inhibitors in rheumatoid arthritis are available with lower rates of LDA, remission, nonserious infection, and higher relapse in the tapering group in a long-term extension of baricitinib phase 3 trial. However, two-thirds of the patients who lost response to the 2-mg dose regained LDA after dose escalation to 4 mg.[6] Similar results were obtained in a Japanese

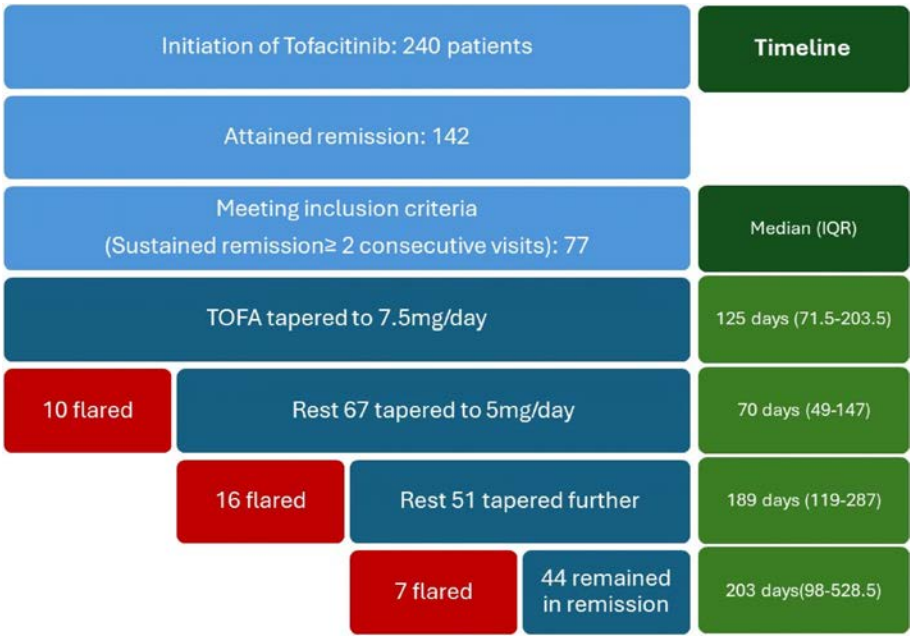


Figure 1. Study flowchart and summary. IQR, interquartile range; TOFA, tofacitinib.

study with higher rates of relapse on immediate complete withdrawal of tofacitinib compared with gradual dose reduction, hence advocating the latter.[7]

This is the first-of-its-kind data on effectiveness of tofacitinib tapering in AxSpA, proving that tofacitinib taper is viable for patients with AxSpA in remission, minimizing drug exposure. All patients who flared reattained remission after restoring the tofacitinib to the previous effective dose, showing that this approach does not lead to resistance to the effects of tofacitinib. The main limitations are the retrospective study design and absence of a control group. It sets the stage for larger clinical trials to explore the effect of tapering on the rate of adverse effects (primarily infections and cardiovascular risk) and radiographic progression in AxSpA.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/art.43147).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/art.43147.

Anu Sreekanth, MD, DrNB
Justin George, Pharm D
Sanjo Saijan, Pharm D
Aparna R. Menon, Pharm D
Padmanabha Shenoy D., MD, DM
drdpshenoy@gmail.com
Centre for Arthritis and Rheumatism Excellence
Kochi, Kerala, India
Sakir Ahmed, MD, DM
Kalinga Institute of Medical Sciences
Bhubaneswar, Orissa, India

Pankti Mehta, MD, DM
University of Toronto
Toronto, ON, Canada
Bhowmik Meghnathi, MD, DrNB
SVP Hospital
and Smt NHL Municipal Medical College
and RheumaCARE Ahmedabad
and Marengo CIMS Hospital
Ahmedabad, India

1. Webers C, Nikiphorou E, Boonen A, et al. Tapering or discontinuation of biological disease-modifying antirheumatic drugs in axial spondyloarthritis: a review of the literature and discussion on current practice. Joint Bone Spine 2023;90(1):105482.
2. Ramiro S, Nikiphorou E, Sepriano A, et al. ASAS-EULAR recommendations for the management of axial spondyloarthritis: 2022 update. Ann Rheum Dis 2023;82(1):19–34.
3. Ytterberg SR, Bhatt DL, Mikuls TR, et al. Cardiovascular and cancer risk with tofacitinib in rheumatoid arthritis. N Engl J Med 2022:386(4):316–326.
4. Sieper J, Rudwaleit M, Baraliakos X, et al. The assessment of Spondylo Arthritis International Society (ASAS) handbook: a guide to assess spondyloarthritis. Annals of the Rheumatic Diseases 2009;68(Suppl 2):ii1–44.
5. Gossec L, Portier A, Landewé R, et al. Preliminary definitions of 'flare' in axial spondyloarthritis, based on pain, BASDAI and ASDAS-CRP: an ASAS initiative. Ann Rheum Dis 2016;75(6):991–996.
6. Takeuchi T, Genovese MC, Haraoui B, et al. Dose reduction of baricitinib in patients with rheumatoid arthritis achieving sustained disease control: results of a prospective study. Ann Rheum Dis 2019;78(2):171–178.
7. Mori S, Ueki Y. Outcomes of dose reduction, withdrawal, and restart of tofacitinib in patients with rheumatoid arthritis: a prospective observational study. Clin Rheumatol 2019;38(12):3391–3400.

DOI 10.1002/art.43352

Secukinumab after tocilizumab failure in giant cell arteritis

To the Editor:

Secukinumab has been shown to be of potential interest in the management of patients with giant cell arteritis (GCA).[1] We report five patients with GCA who relapsed during tocilizumab treatment and were subsequently treated with secukinumab. These patients met the American College of Rheumatology/EULAR diagnostic criteria for GCA.[2] Patients' characteristics are shown in Table 1.

Before secukinumab treatment, all patients had received prednisone at 0.7 to 1 mg/kg/day, four patients had received methotrexate at 0.3 mg/kg/wk for a period of 3 to 48 months, and all patients had been treated with tocilizumab at 162 mg/wk subcutaneously for a period of 6 to 21 months. Secukinumab was initiated with five subcutaneous injections at 300 mg weekly, then 300 mg monthly. Methotrexate was discontinued in all patients at introduction of tocilizumab or secukinumab. Just before secukinumab therapy, all patients were taking tocilizumab, and three were still being treated with prednisone. Concomitantly to secukinumab initiation, prednisone was increased to 20 mg/day for each patient, with progressive withdrawal in three to six months.

Secukinumab was introduced for GCA relapse, with clinical symptoms including polymyalgia rheumatica (PMR) in all patients and recurrence of headache in three patients (one with halo sign on sonography). Four patients had worsening vascular lesions on positron emission tomography (PET).

At introduction of secukinumab, only one patient's C-reactive protein (CRP) level increased to 16 mg/L just after stopping tocilizumab. The CRP level tended to rise for a few weeks and then normalized.

All patients had a PET scan just before secukinumab initiation. PET scans were performed during secukinumab treatment, at 6 months for four patients, and at 12 months for two patients. Evolution of PET in these patients is shown in Table 1.

At last follow-up, two patients were still in clinical and biologic remission without glucocorticoids, one patient was in clinical and biologic remission during prednisone treatment of 7.5 mg/day, and two patients relapsed after 6 and 12 months of treatment with secukinumab. These two patients were subsequently switched to upadacitinib, with good clinical and biologic efficacy. None of the patients had any infectious complications during the secukinumab treatment period.

In these five cases of GCA with large vessel vasculitis (LVV) refractory to tocilizumab, treatment with secukinumab induced clinical and biologic remission in three patients. Only two of these three patients achieved improvement of LVV or PMR on PET.

Secukinumab appears to be an interesting treatment for GCA, as interleukin-17 is strongly implicated in the pathogenesis of this disease. Its mechanisms of action are different from those of tocilizumab and methotrexate.[3,4] In this case series, tocilizumab failed to achieve sustained remission, and secukinumab contributed to clinical and biologic remission. Only one patient had a negative PET scan result after secukinumab initiation; however, vascular remission on imaging is difficult to achieve in these patients with refractory disease. Although our experience suggests the potential use of secukinumab for patients with refractory GCA, that efficacy needs to be confirmed in larger randomized controlled trials.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/art.43352.

[Correction added on 15 December 2025, after first online publication: The order of authors has been changed.]

Olivier Espitia, MD, PhD
Olivier.espitia@chu-nantes.fr
Marie Morard, MD
Department of Internal and Vascular Medicine, Centre Hospitalier Universitaire de Nantes, Nantes Université and L'Institut du Thorax Team III Vascular and Pulmonary Diseases INSERM UMR 1087, CNRS UMR 6291, Centre Hospitalier Universitaire de Nantes,
Nantes Université, Nantes, France
Bastien Jamet, MD
Médecine Nucléaire, CRCI2NA, CNRS, INSERM, Centre Hospitalier Universitaire de Nantes,
Nantes Université, University of Angers
Nantes, France
Antoine Benichou, MD
Victor Genin, MD
Christian Agard, MD, PhD
Department of Internal and Vascular Medicine, Centre Hospitalier Universitaire de Nantes,
Nantes Université
Nantes, France

Table 1. Patient characteristics*

	Case 1	Case 2	Case 3	Case 4	Case 5
Age at GCA diagnosis, y	53	67	64	66	68
Sex	Male	Female	Female	Female	Female
Characteristics at time of GCA diagnosis	Fever, headache, jaw claudication, chest pain	PMR, weight loss, chest pain	Headache, jaw claudication, PMR	Headache, jaw claudication, scalp tenderness	PMR, weight loss
	Aortitis Negative TAB result	Aortitis Positive TAB result	Aortitis during follow-up Positive TAB result	Aortitis Positive TAB result	Aortitis Positive TAB result
Treatments before secukinumab (duration of treatments)	Prednisone (initially 1 mg/kg/day)	Prednisone (initially 1 mg/kg/day)	Prednisone (initially 0.7 mg/kg/day)	Prednisone (initially 0.7 mg/kg/day)	Prednisone (initially 0.7 mg/kg/day)
	MTX 15 mg/wk (3 mo) TCZ 162 mg/wk (8 mo)	MTX 22.5 mg/wk (27 mo) TCZ 162 mg/wk (6 mo)	MTX 15 mg/wk (47 mo) TCZ 162 mg/wk (21 mo)	TCZ 162 mg/wk (11 mo) –	MTX 20 mg/wk (3 mo) TCZ 162 mg/wk (12 mo)
No. of relapses before introduction of secukinumab	3	4	4	1	3
Time from diagnosis to introduction of secukinumab, mo	89	94	70	13	15
Clinical signs at the start of secukinumab treatment	PMR	PMR, headache	PMR, headache	PMR, headache	PMR
CRP level at secukinumab start, mg/L	0.6	5	16	1.9	3.5
PET result at secukinumab start	LVV: no	LVV: PETVAS, 11/27; G-PETVAS, 16/51; appearance of 3 arterial segments with grade III uptakes	LVV: PETVAS, 20/27; G-PETVAS, 40/51; 6 arterial segments with grade III uptakes	LVV: PETVAS, 27/27; G-PETVAS, 40/51; 12 arterial segments with grade III uptakes	LVV: PETVAS, 7/27; G-PETVAS, 19/51; appearance of 1 arterial segments with grade III uptakes
	PMR: bilateral subacromial bursitis uptakes	PMR: cervical interspinous bursitis uptakes	PMR: no	PMR: lumbar interspinous bursitis and ischial tuberosities and posteromedial knee uptakes	PMR: no
Prednisone dosage just before secukinumab start, mg/day	0	5	5	0	7.5
Duration of treatment with secukinumab, mo	12	12	13	6	6
Last CRP level while taking secukinumab, mg/L	1.1	12	1	2.5	9
Last fibrinogen level while taking secukinumab, g/L	2.6	5.1	3.3	5.4	NA
Last prednisone dosage while taking secukinumab, mg/day	0	0	0	0	7.5

(Continued)

Table 1. (*Cont'd*)

	Case 1	Case 2	Case 3	Case 4	Case 5
PET result after 6 mo of secukinumab treatment	LVV: no PMR: no	LVV: PETVAS, 9/27; G-PETVAS, 9/51; 2 arterial segments with grade III uptakes PMR: no	LVV: PETVAS, 12/27; G-PETVAS, 20/51; 1 arterial segments with grade III uptakes PMR: no	LVV: PETVAS, 27/27; G-PETVAS, 41/51; 13 arterial segments with grade III uptakes PMR: lumbar interspinous bursitis and ischial tuberosities and posteromedial knee uptakes	NA
PET result after 12 mo of secukinumab treatment	NA	LVV: PETVAS, 10/27; G-PETVAS, 10/51; 3 arterial segments with grade III uptakes PMR: spinal bursitis, ischial tuberosities, pelvic and scapular girdle synovitis uptakes	LVV: PETVAS, 10/27; G-PETVAS, 14/51; 1 arterial segments with grade III uptakes PMR: no	NA	NA
GCA relapse during secukinumab treatment	No	Yes	No	Yes	No
Biologic response	Yes	No	Yes	No	Yes
Clinical response	Yes	No	Yes	No	Yes
PET response	Yes	No	Partial	No	NA

* PETVAS is the sum of the fluorodeoxyglucose uptake grades of the ascending aorta, the aortic arch, the descending thoracic aorta, the abdominal aorta, the carotid arteries, the subclavian arteries, and the brachiocephalic arterial trunk. G-PETVAS analyzes the same arterial segments as PETVAS, with the addition of vertebral, axillary, iliac, and femoral arteries; for paired arteries, each side is evaluated independently. CRP, C-reactive protein; GCA, giant cell arteritis; G-PETVAS, global PETVAS; LVV, large vessel vasculitis; MTX, methotrexate; NA, not available; PET, positron emission tomography; PETVAS, Positron Emission Tomography Vascular Activity Score; PMR, polymyalgia rheumatica; TAB, temporal artery biopsy; TCZ, tocilizumab.

1. Venhoff N, Schmidt WA, Bergner R, et al. Safety and efficacy of secukinumab in patients with giant cell arteritis (TitAIN): a randomised, double-blind, placebo-controlled, phase 2 trial. Lancet Rheumatol 2023;5(6):e341–e350.
2. Ponte C, Grayson PC, Robson JC, et al; DCVAS Study Group. 2022 American College of Rheumatology/EULAR classification criteria for giant cell arteritis. Arthritis Rheumatol 2022;74(12):1881–1889.
3. Adriawan IR, Atschekzei F, Dittrich-Breiholz O, et al. Novel aspects of regulatory T cell dysfunction as a therapeutic target in giant cell arteritis. Ann Rheum Dis 2022;81(1):124–131.
4. Miyabe C, Miyabe Y, Strle K, et al. An expanded population of pathogenic regulatory T cells in giant cell arteritis is abrogated by IL-6 blockade therapy. Ann Rheum Dis 2017;76(5):898–905.

DOI 10.1002/art.43390

Less is more: hematologic safety and sustained efficacy of a modified low-dose etoposide regimen in macrophage activation syndrome

To the Editor:

Macrophage activation syndrome (MAS), as a potentially fatal complication of autoimmune inflammatory rheumatic diseases and a subtype of secondary hemophagocytic lymphohistiocytosis (HLH), is characterized by dysregulated immune activation culminating in cytokine storm.[1] Patients typically present with persistent high fever, cytopenia, hepatosplenomegaly, and multiorgan dysfunction, with a high mortality rate.[2,3] Because of the lack of specific diagnostic criteria for MAS, HLH-2004 criteria and the H-score system are commonly used for MAS recognition, with reasonable diagnostic accuracy.[4] Glucocorticoid pulse therapy (adopted from pediatric regimens) is the first-line treatment for MAS, yet approximately half of adults exhibit refractoriness.[5] For patients with refractory MAS, severe organ involvement or those failing first-line therapy, early initiation of an etoposide-based protocol should be considered.[6] However, given that MAS is frequently associated with cytopenia and etoposide is a known cause of severe myelosuppression, the combination of these factors poses significant risk of severe infections, which remain a significant clinical challenge and limit the application of etoposide in MAS. We have previously reported, in a MAS case report, that low-dose etoposide may reduce the risk of infection while conferring therapeutic effects on uncontrolled inflammation.[7] To compare the efficacy and safety of this optimized low-dose etoposide regimen with conventional-dose etoposide treatment, we retrospectively analyzed the outcomes of all patients with MAS treated with low-dose etoposide in our hospital. Twenty-two patients with MAS hospitalized in Tongji Hospital between August 2015 and July 2024 were included, with 10 of them receiving a modified low-dose etoposide regimen (100 mg, twice weekly, for a total of three or four doses) and 12 patients were treated with conventional-dose etoposide (twice weekly with a dose of 165–200 mg, determined based on patient's body surface area and organ function[8,9]). Therapeutic responses, defined as complete remission, partial remission, or no remission using criteria established by Wang et al,[10] were evaluated one month after treatment initiation and again at the last follow-up. Safety was assessed by monitoring hematologic toxicity, particularly the degree of myelosuppression during treatment and within seven days after the final dose.

We observed no significant differences in primary disease type, baseline disease activity, and antirheumatic medications between the conventional-dose therapy group and the modified therapy group ($P > 0.05$) (Table 1). Throughout the follow-up period (median follow-up duration: 16 [interquartile range (IQR) 3.5–29] vs 10 [IQR 6.75–14.5] months), no differences in treatment efficacy were observed between the two groups, either in the short term (1 month) or long term (up to the final follow-up: $P > 0.05$ for treatment response at week four or last follow-up, H-score at week 4, and recurrence). It is noteworthy that the modified treatment group demonstrated improved hematologic safety. Compared to the conventional-dose treatment group, the modified treatment group exhibited a significantly higher minimum leukocyte count (2.38 [IQR 1.98–2.68] vs 1.43 [IQR 0.88–2.09], $P < 0.05$; Table 1 and Figure 1), a lower magnitude of leukocyte decline (0.11 [IQR -0.10 to 0.59] vs 0.81 [IQR 0.72–0.88], $P < 0.05$; Table 1 and Figure 1), and shorter duration required for leukocyte counts to return to normal (11 [IQR 7–13] vs 2 [IQR 2–4], $P < 0.05$; Table 1). Similar protective effects were also observed for neutrophil subset, reflected by higher minimum count (1.19 [IQR 1.16–1.44] vs 0.40 [IQR 0.14–0.75], $P <$

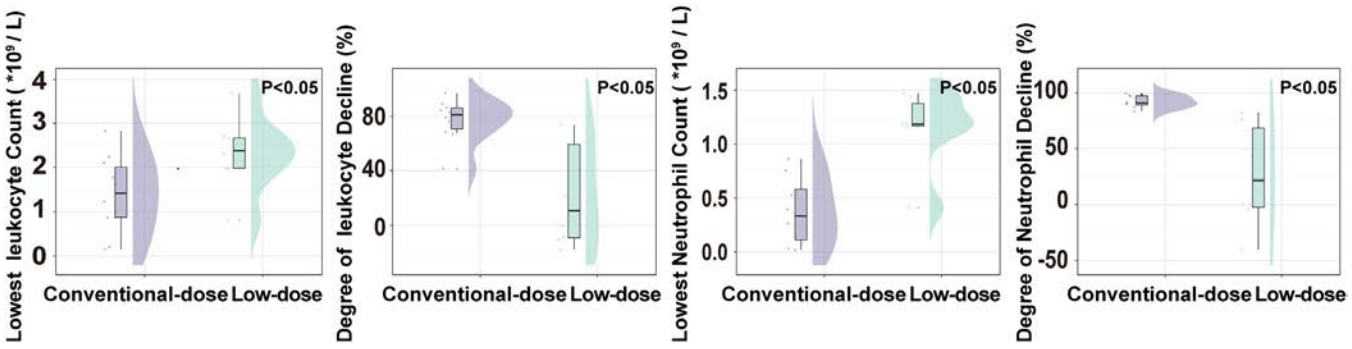


Figure 1. Comparison of leukocyte and neutriphil nadirs and declines between conventional-dose and low-dose etoposide.

1. Venhoff N, Schmidt WA, Bergner R, et al. Safety and efficacy of secukinumab in patients with giant cell arteritis (TitAIN): a randomised, double-blind, placebo-controlled, phase 2 trial. Lancet Rheumatol 2023;5(6):e341–e350.
2. Ponte C, Grayson PC, Robson JC, et al; DCVAS Study Group. 2022 American College of Rheumatology/EULAR classification criteria for giant cell arteritis. Arthritis Rheumatol 2022;74(12):1881–1889.
3. Adriawan IR, Atschekzei F, Dittrich-Breiholz O, et al. Novel aspects of regulatory T cell dysfunction as a therapeutic target in giant cell arteritis. Ann Rheum Dis 2022;81(1):124–131.
4. Miyabe C, Miyabe Y, Strle K, et al. An expanded population of pathogenic regulatory T cells in giant cell arteritis is abrogated by IL-6 blockade therapy. Ann Rheum Dis 2017;76(5):898–905.

DOI 10.1002/art.43390

Less is more: hematologic safety and sustained efficacy of a modified low-dose etoposide regimen in macrophage activation syndrome

To the Editor:

Macrophage activation syndrome (MAS), as a potentially fatal complication of autoimmune inflammatory rheumatic diseases and a subtype of secondary hemophagocytic lymphohistiocytosis (HLH), is characterized by dysregulated immune activation culminating in cytokine storm.[1] Patients typically present with persistent high fever, cytopenia, hepatosplenomegaly, and multiorgan dysfunction, with a high mortality rate.[2,3] Because of the lack of specific diagnostic criteria for MAS, HLH-2004 criteria and the H-score system are commonly used for MAS recognition, with reasonable diagnostic accuracy.[4] Glucocorticoid pulse therapy (adopted from pediatric regimens) is the first-line treatment for MAS, yet approximately half of adults exhibit refractoriness.[5] For patients with refractory MAS, severe organ involvement or those failing first-line therapy, early initiation of an etoposide-based protocol should be considered.[6] However, given that MAS is frequently associated with cytopenia and etoposide is a known cause of severe myelosuppression, the combination of these factors poses significant risk of severe infections, which remain a significant clinical challenge and limit the application of etoposide in MAS. We have previously reported, in a MAS case report, that low-dose etoposide may reduce the risk of infection while conferring therapeutic effects on uncontrolled inflammation.[7] To compare the efficacy and safety of this optimized low-dose etoposide regimen with conventional-dose etoposide treatment, we retrospectively analyzed the outcomes of all patients with MAS treated with low-dose etoposide in our hospital. Twenty-two patients with MAS hospitalized in Tongji Hospital between August 2015 and July 2024 were included, with 10 of them receiving a modified low-dose etoposide regimen (100 mg, twice weekly, for a total of three or four doses) and 12 patients were treated with conventional-dose etoposide (twice weekly with a dose of 165–200 mg, determined based on patient's body surface area and organ function[8,9]). Therapeutic responses, defined as complete remission, partial remission, or no remission using criteria established by Wang et al,[10] were evaluated one month after treatment initiation and again at the last follow-up. Safety was assessed by monitoring hematologic toxicity, particularly the degree of myelosuppression during treatment and within seven days after the final dose.

We observed no significant differences in primary disease type, baseline disease activity, and antirheumatic medications between the conventional-dose therapy group and the modified therapy group ($P > 0.05$) (Table 1). Throughout the follow-up period (median follow-up duration: 16 [interquartile range (IQR) 3.5–29] vs 10 [IQR 6.75–14.5] months), no differences in treatment efficacy were observed between the two groups, either in the short term (1 month) or long term (up to the final follow-up: $P > 0.05$ for treatment response at week four or last follow-up, H-score at week 4, and recurrence). It is noteworthy that the modified treatment group demonstrated improved hematologic safety. Compared to the conventional-dose treatment group, the modified treatment group exhibited a significantly higher minimum leukocyte count (2.38 [IQR 1.98–2.68] vs 1.43 [IQR 0.88–2.09], $P < 0.05$; Table 1 and Figure 1), a lower magnitude of leukocyte decline (0.11 [IQR -0.10 to 0.59] vs 0.81 [IQR 0.72–0.88], $P < 0.05$; Table 1 and Figure 1), and shorter duration required for leukocyte counts to return to normal (11 [IQR 7–13] vs 2 [IQR 2–4], $P < 0.05$; Table 1). Similar protective effects were also observed for neutrophil subset, reflected by higher minimum count (1.19 [IQR 1.16–1.44] vs 0.40 [IQR 0.14–0.75], $P <$

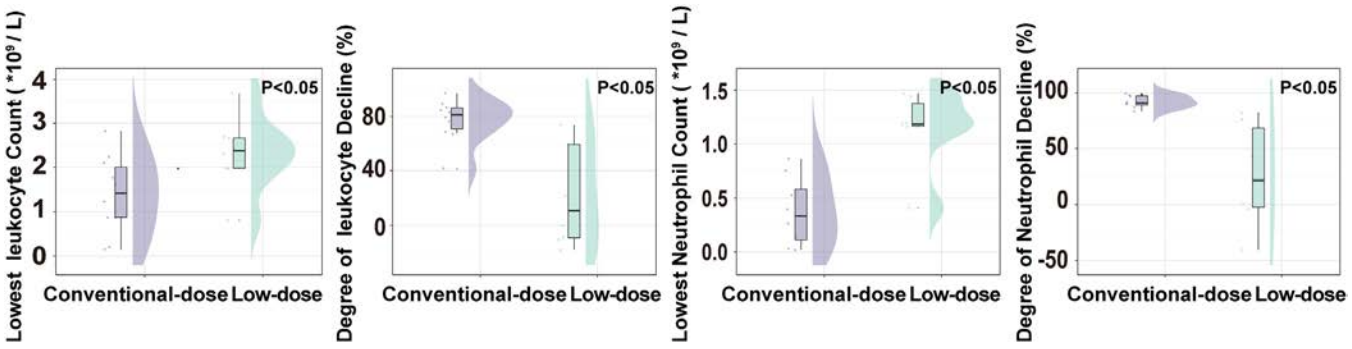


Figure 1. Comparison of leukocyte and neutriphil nadirs and declines between conventional-dose and low-dose etoposide.

Table 1. Baseline characteristics, treatments, and clinical outcomes in patients with MAS treated with conventional-dose or low-dose etoposide.

	Variant	Conventional-dose Etoposide (n = 12)[1]	Low-dose Etoposide (n = 10)[1]	p
Baseline characteristics	Gender (Male)	3 (25.00%)	1 (10.00%)	0.59
	Age (years)	35 (27, 43)	40 (25, 57)	0.51
	Etiology/trigger			0.82
	Systemic Lupus Erythematosus (Yes)	3 (25.00%)	5 (50.00%)	
	Adult-Onset Still's Disease (Yes)	6 (50.00%)	3 (30.00%)	
	Dermatomyositis (Yes)	1 (8.30%)	1 (10.00%)	
	Undifferentiated Connective Tissue Disease (Yes)	1 (8.30%)	1 (10.00%)	
	Systemic Juvenile Idiopathic Arthritis (Yes)	1 (8.30%)	0 (0.00%)	
	Fever (%)	12 (100.00%)	10 (100.00%)	1.00
	Hepatomegaly (Yes)	0 (0.00%)	0 (0.00%)	1.00
	Splenomegaly (Yes)	1 (8.30%)	0 (0.00%)	1.00
	Lymphadenopathy (Yes)	1 (8.30%)	2 (20.00%)	0.57
	White cell count (10^9/L)[†]	7.18(5.27, 9.86)	4.08 (2.09, 7.55)	0.16
	Neutrophil count (10^9/L)[†]	5.03 (4.13, 7.00)	2.92 (1.43, 6.54)	0.20
	Haemoglobin (g/L)[†]	110 (100, 119)	98 (89, 110)	0.23
	Platelet count (10^9/L)[†]	206 (152, 270)	128 (90, 191)	0.14
	Alanine Aminotransferase (U/L)[‡]	80 (27, 150)	50 (30, 58)	0.57
	Aspartate transferase (U/L)[‡]	87 (45, 213)	74 (43, 154)	0.87
	Triglyceride (mmol/L)[†]	2.45 (1.99, 3.73)	2.83 (2.06, 3.12)	0.97
	Fibrinogen (g/L)[‡]	2.14 (1.59, 3.53)	2.81 (2.31, 3.08)	0.46
	Albumin (g/L)[‡]	31.15 (29.33 32.20)	29.85 (25.78, 35.80)	0.54
	Serum creatinine (umol/L)[‡]	49.50 (40.00, 70.00)	59.00 (42.50, 75.25)	0.82
	Total bilirubin (umol/L)[‡]	7.25 (6.60, 10.00)	8.05 (5.53, 13.55)	0.67
	Prothrombin time (s)[‡]	14.70 (12.80, 19.10)	14.50 (14.00, 18.60)	0.60
	INR[‡]	1.01 (0.97, 1.09)	1.11 (1.02, 1.18)	0.15
	Ferritin (ng/mL)[‡]	7383 (4559, 11467)	3222 (898, 14171)	0.16
	Bone marrow haemophagocytosis (Yes)	9 (75.00%)	8 (80.00%)	1.00
	Low NK cell activity (Yes)	11 (92.00%)	6 (75.00%)	0.54
	Soluble CD25 (U/mL)[‡]	1637 (1060, 2795)	2,875 (1493, 4734)	0.31
	HLH-2004 criteria[*,‡]	5.00 (5.00, 7.00)	6.00 (5.25, 6.00)	0.14
	H score[*,†]	203 (177, 228)	183 (179, 194)	0.57
	Follow-up (Months)[‡]	16.00 (3.50, 29.00)	10.00 (6.75, 14.50)	0.17
Medications used during hospitalization	**Etoposide**			
	Dose per Administration (mg)[†]	200.00 (165.00, 200.00)	100.00 (100.00, 100.0)	0.00
	Dose per Administration (mg/m^2)[†]	122.70 (102.50, 146.30)	62.33 (56.86, 69.02)	0.00
	Cumulative Inpatient Dose (mg)[‡]	550.00 (305.00, 780.00)	300.00 (300.00, 400.00)	0.02
	Cumulative Inpatient Dose (mg/m^2)[‡]	318.20 (191.50, 495.70)	189.80 (178.10, 238.60)	0.01
	Other medications			
	Average Daily Dose, Prednisone Equivalent (mg/kg/day)[‡]	1.56 (1.21, 2.47)	1.54 (1.01, 1.98)	0.55
	Cyclosporine A (Yes)	9 (75.00%)	9 (90.00%)	0.59
	Mycophenolate Mofetil (Yes)	1 (8.30%)	0 (0.00%)	1.00
	Intravenous Immunoglobulin (Yes)	8 (67.00%)	3 (30.00%)	0.20
	Cyclophosphamide (Yes)	3 (25.00%)	4 (60.00%)	0.19
	Plasma Exchange (Yes)	4 (33.00%)	3 (30.00%)	1.00
	Other Myelosuppressive drugs (Yes)	4 (33.00%)	6 (60.00%)	0.39
Efficacy and hematologic safety	**Patient's response at 4 weeks**			
	Complete response	10 (83.33%)	7 (70.00%)	0.62
	Partial response	2 (16.67%)	3 (30.00%)	
	Patient's response at last follow-up			
	Complete response	3 (25.00%)	6 (60.00%)	0.19
	Partial response	9 (75.00%)	4 (40.00%)	
	H score[*] at 4 weeks[†]	54 (24, 104)	62 (44, 69)	0.99
	Recurrence (Yes)	5 (42.00%)	1 (10.00%)	0.16
	Leukopenia (< 4*10^9/L) (Yes)	10 (83.30%)	8 (80.00%)	1.00
	Leukopenia (< 2*10^9/L) (Yes)	7 (58.30%)	3 (30.00%)	0.23
	Lowest White Blood Cell Count (10^9/L)[†]	1.43 (0.88, 2.09)	2.38 (1.98, 2.68)	< 0.05
	Degree of White Blood Cell Decline (%)[†]	81 (72, 88)	11 (−10, 59)	< 0.05
	Neutropenia (< 1.5*10^9/L) (Yes)	9 (75.00%)	4 (40.00%)	0.19

(Continued)

Table 1. (*Cont'd*)

Variant	Conventional-dose Etoposide (n = 12)[1]	Low-dose Etoposide (n = 10)[1]	p
Neutropenia (< 0.5*10^9/L) (Yes)	5 (41.70%)	1 (10.00%)	0.16
Lowest Neutrophil Count (10^9/L)‡	0.40 (0.14, 0.75)	1.19 (1.16, 1.44)	< 0.05
Degree of Neutrophil Decline (%)‡	91 (90, 97)	22 (−3, 77)	< 0.05
Use of Granulocyte colony-stimulating factor (Yes)	9 (75.00%)	4 (40.00%)	0.12
Frequency of Granulocyte colony-stimulating factor Use (μg)†	600 (100, 2100)	250 (100, 900)	0.48
Anemia (Yes)	10 (83.00%)	10 (100.00%)	0.48
Lowest Hemoglobin Level (/L)†	88 (64, 97)	82 (65, 94)	0.69
Degree of Hemoglobin Decline (%)†	21 (12, 29)	19 (14, 24)	0.31
Thrombocytopenia (Yes)	9 (75.00%)	7 (70.00%)	1.00
Lowest Platelet Count (/L)†	46 (41, 76)	94 (47, 110)	0.48
Severity of Thrombocytopenia (%)†	72 (28, 85)	25 (8, 60)	0.11
Blood Recovery Time (day)‡	11 (7, 13)	2 (2, 4)	< 0.05

Note: Glucocoticoids were calculated as prednisone-equivalent using, dexamethasone x6, methylprednisolone x1.25, prednisone x1; Other myelosuppressive drugs, including concomitant medications like cyclophosphamide and tacrolimus may contribute to hematologic toxicity. Recurrence (Yes) was defined as readmission for macrophage activation syndrome; Since no bone marrow aspiration cytology was conducted in any group at 4 weeks post-treatment, all data were recorded as 0.
* HLH criteria (at least five criteria should be met for MAS diagnosis): temperature ≥ 38.5°C for at least 7 days; splenomegaly; cytopenia in two or more cell lines (haemoglobin < 90 g/L, PLT < 100 × 10^9/L, or leukocyte < 4 × 10^9/L); hypertriglyceridaemia (> 3 mmol/L) and/or hypofibrinogenaemia (< 150 mg/dL); ferritin ≥ 500 μg/L; increased sCD25; decreased or absent NK cell activity; haemophagocytic cells in bone marrow, spleen or lymph nodes. HScore (http://saintantoine.aphp.fr/score/) estimates hemophagocytic syndrome risk.
† Welch Two Sample t-test.
‡ Wilcoxon rank sum exact test; Fisher's exact test was used for categorical variables.
[1] n (%); Median (Q1, Q3).

0.05) and reduced magnitude of decline (0.22 [IQR -0.03 to 0.77] vs 0.91 [IQR 0.90–0.97], P < 0.05; Table 1 and Figure 1). This hematologic protection was accompanied by a trend of lower frequency of granulocyte colony-stimulating factor (G-CSF) (75.0% vs 40.0%, P = 0.12) and lower accumulative dose of G-CSF (600 [IQR 100–2,100] vs 250 [IQR 100–900], P = 0.48) required for the treatment of leukopenia, although no statistical significance was observed (Table 1). It nevertheless indicates less need for supportive care. No significant differences were found in frequencies of anemia and thrombocytopenia, minimum hemoglobin levels, and the magnitude of hemoglobin decline (Table 1).

Overall, these data suggest that the modified low-dose etoposide regimen significantly reduces the severity of myelosuppression while maintaining therapeutic efficacy in MAS treatment, particularly in terms of leukocyte and neutrophil suppression. Although not first-line therapy, partially because of its myelosuppressive toxicity, etoposide usage is increasingly prevalent in MAS, especially in glucocorticoid-refractory cases. Given the retrospective nature of this research, future prospective studies are required to confirm its wider applicability and long-term safety for treating MAS.

Supported by the National Natural Science Foundation of China (grants 82270903, 82370465, and 82401588), National Key Research and Development Program of China (grants 2023YFC2507900 and 2022YFA1105303), China Postdoctoral Science Foundation (grant 2024M751019), Hubei Provincial Key Research and Development Program (grant 2023BCB013), and Hubei Provincial Natural Science Foundation (grant 2023AFB771).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/art.43390.

Zijun Ma, MD
Yingying Wei, MD, PhD
Department of Rheumatology, Fujian Medical University Union Hospital
Fuzhou, China
and Department of Rheumatology and Immunology, Tongji Hospital, Tongji Medical College,
Huazhong University of Science and Technology
Wuhan, China
Xiaoquan Rao, MD, PhD
xqrao@tjh.tjmu.edu.cn
Division of Cardiology, Department of Internal Medicine, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology
Wuhan, China
Wei Tu, MD, PhD
tuwei@tjh.tjmu.edu.cn
Department of Rheumatology and Immunology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology
Wuhan, China
Jixin Zhong, MD, PhD
zhongjixin620@163.com
Department of Rheumatology, Fujian Medical University Union Hospital
Fuzhou, China
and Key Laboratory of Vascular Aging (HUST), Ministry of Education
Wuhan, China
and Fujian Institute of Clinical Immunology
Fuzhou, China

1. Nigrovic PA. Macrophage activation syndrome. Arthritis Rheumatol 2025;77(4):367–379.
2. Hines MR, von Bahr Greenwood T, Beutel G, et al. Consensus-based guidelines for the recognition, diagnosis, and management of hemophagocytic lymphohistiocytosis in critically ill children and adults. Crit Care Med 2022;50(5):860–872.
3. Shakoory B, Geerlinks A, Wilejto M, et al; HLH/MAS task force. The 2022 EULAR/ACR points to consider at the early stages of diagnosis and management of suspected haemophagocytic lymphohistiocytosis/macrophage activation syndrome (HLH/MAS). Ann Rheum Dis 2023;82(10):1271–1285.
4. Lachmann G, Heeren P, Schuster FS, et al. Multicenter validation of secondary hemophagocytic lymphohistiocytosis diagnostic criteria. J Intern Med 2025;297(3):312–327.
5. Fukaya S, Yasuda S, Hashimoto T, et al. Clinical features of haemophagocytic syndrome in patients with systemic autoimmune diseases: analysis of 30 cases. Rheumatology (Oxford) 2008;47(11):1686–1691.
6. Gavand PE, Serio I, Arnaud L, et al. Clinical spectrum and therapeutic management of systemic lupus erythematosus-associated macrophage activation syndrome: A study of 103 episodes in 89 adult patients. Autoimmun Rev 2017;16(7):743–749.
7. Gao SP, Luo XF, Kosari M, et al. Successful management of infection and macrophage activation syndrome patient using low-dose etoposide: A case report. World J Radiol 2024;16(10):579–585.
8. Arbuck SG, Douglass HO, Crom WR, et al. Etoposide pharmacokinetics in patients with normal and abnormal organ function. J Clin Oncol 1986;4(11):1690–1695.
9. Joel SP, Shah R, Clark PI, et al. Predicting etoposide toxicity: relationship to organ function and protein binding. J Clin Oncol 1996;14(1):257–267.
10. Wang Y, Huang W, Hu L, et al. Multicenter study of combination DEP regimen as a salvage therapy for adult refractory hemophagocytic lymphohistiocytosis. Blood 2015;126(19):2186–2192.

Arthritis & Rheumatology
Vol. 78, No. 3, March 2026, pp 781
DOI 10.1002/art.43178

American College of Rheumatology
Empowering Rheumatology Professionals

CORRECTION

Correction to "Novel Genetic Loci in Early-Onset Gout Derived From Whole-Genome Sequencing of an Adolescent Gout Cohort"

Ji A, Sui Y, Xuo X, et al. Novel Genetic Loci in Early-Onset Gout Derived From Whole-Genome Sequencing of an Adolescent Gout Cohort. Arthritis Rheumatol. 2025 Jan;77(1):107–115.

In the article noted above, the author name Xiapeng Ji is misspelled. The correct spelling is Xiaopeng Ji.

We apologize for this error.

American College of Rheumatology
Empowering Rheumatology Professionals

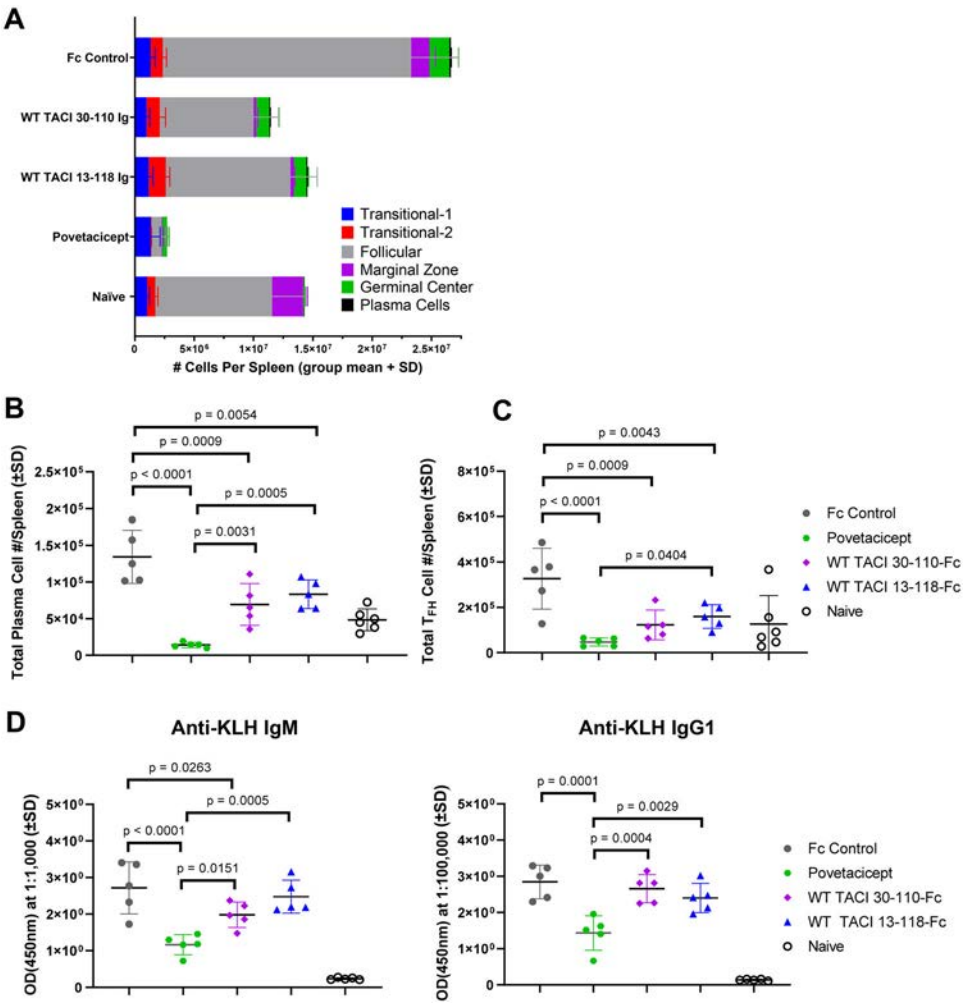
Arthritis & Rheumatology
Vol. 78, No. 3, March 2026, pp 782
DOI 10.1002/art.43181

CORRECTION

Correction to "Povetacicept, an Enhanced Dual APRIL/BAFF Antagonist That Modulates B Lymphocytes and Pathogenic Autoantibodies for the Treatment of Lupus and Other B Cell-Related Autoimmune Diseases"

Arthritis Rheumatol. 2023 Jul;75(7):1187-1202. doi: 10.1002/art.42462. Epub 2023 Apr 19.

In Figure 3A, the labels for the two WT TACI-Fc comparators on the y-axis were inverted. The 2nd bar from the top should have been labeled "WT TACI 30-110 Ig" and the 3rd bar from the top should have been labeled "WT TACI 13-118 Ig."



We apologize for this error.