Arthritis Care & Research

Aims and Scope

Arthritis Care & Research is an official journal of the American College of Rheumatology and the Association of Rheumatology Professionals, a division of the College. *Arthritis Care & Research* is a peer-reviewed journal that publishes both original research and review articles that promote excellence in the clinical practice of rheumatology. Relevant to the care of individuals with arthritis and related disorders, major topics are evidence-based practice studies, clinical problems, practice guidelines, health care economics, health care policy, educational, social, and public health issues, and future trends in rheumatology practice.

Volume 78 No. 6 June 2026

Arthritis Care & Research

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

Cover design: Sandra Pulmano

♾ This journal is printed on acid-free paper.

Arthritis Care & Research

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

VOLUME 78 • JUNE 2026 • NO. 6

Cover image: The image on the cover (Niessink et al; pages 788–797; Supplementary Figure 1) shows compensated polarized light microscopy images from synovial crystals.

Arthritis Care & Research
Vol. 78, No. 6, June 2026, pp 693–695
DOI 10.1002/acr.80003

American College of Rheumatology
Empowering Rheumatology Professionals

EDITORIAL

Using Genetic Risk Scores for Rheumatoid Arthritis–Associated Interstitial Lung Disease Risk Stratification

Gregory C. McDermott and Jeffrey A. Sparks

Rheumatoid arthritis–associated interstitial lung disease (RA-ILD) results in increased morbidity and mortality in patients with RA.[1–4] There is significant interest in screening strategies for RA-ILD to improve early diagnosis, especially for patients with fibrotic subtypes of RA-ILD, which have a particularly poor prognosis.[5–7] Although several genetic and clinical risk factors for fibrotic forms of RA-ILD have been identified,[8,9] there is uncertainty about the best approaches for identifying patients with RA who are at increased risk for fibrotic forms of RA-ILD.

In this issue of *Arthritis Care & Research*, Brink and colleagues[10] examine the performance of a previously published RA-ILD combined clinical and genetic risk score[11] in a Swedish multicenter cohort of patients with RA. The authors calculated a weighted genetic risk score including five loci previously identified in genetic studies of idiopathic pulmonary fibrosis and extended to RA-ILD.[8,12] Patients with RA were systematically screened with chest radiographs. For most patients, pulmonary fibrosis diagnosis was based on high-resolution computed tomography chest imaging. The authors report associations of clinical and genetic factors with pulmonary fibrosis in RA and display the area under the receiver operating characteristic curve (AUC) for the previously published genetic risk score and combined clinical/genetic risk score. The combined score included the weighted genetic risk score, age at RA diagnosis, male sex, ever smoking status, mean Disease Activity Score in 28 joints using the erythrocyte sedimentation rate over the first 24 months of cohort enrollment, and rheumatoid factor positivity.

Among 1,115 patients with early RA diagnosed in five Swedish rheumatology clinics, 60 (5.6%) had baseline pulmonary fibrosis or developed findings of pulmonary fibrosis over a mean follow-up of about nine years. Older age and positive rheumatoid factor were associated with pulmonary fibrosis, consistent with previously described clinical risk factors for RA-ILD.[1,13–16] The AUC for the genetic risk score was 0.62, and the AUC for the combined clinical/genetic risk score was 0.75. The authors determined that 36.1% of the cohort could potentially be spared screening if the combined clinical/genetic risk score were used with a screening cutoff at >90% sensitivity.

Despite the recognition of important genetic risk factors for RA-ILD,[8] combined clinical/genetic risk profiling has not been extensively investigated in RA-ILD. Several clinical risk factor–based scores have been developed.[17–22] However, only the Étude et Suivi des Polyarthrites Indifférenciées Récentes (ESPOIR) risk score additionally incorporated genetic factors.[19] The ESPOIR risk score included the common *MUC5B* promoter variant (rs35705950), which was also the strongest genetic risk factor identified in the current study, along with sex, age, and disease activity, and demonstrated an AUC of around 0.80 for identifying subclinical RA-ILD. The findings presented in this issue represent an important addition to the existing literature on RA-ILD risk profiling. This study provides an external validation of an existing clinical and genetic risk score. Because the original derivation was performed in the US Department of Veterans Affairs system with a predominantly male population, this replication in a predominantly female Swedish cohort strongly supports the validity of the findings and provides further support for incorporating genetic risk factors into RA-ILD risk profiling. Finally, the authors used a small number of candidate genes that were originally identified in patients with idiopathic pulmonary fibrosis versus healthy controls. Future studies should consider polygenic risk scores across the entire genome and pursue genome-wide association studies specifically devoted to patients with RA-ILD and subtypes versus controls with RA and no ILD.

The study has important limitations to consider. First, it does not report the performance of clinical risk factors alone for RA-ILD identification, and thus it is difficult to quantify the additive benefit of the genetic risk score. In the original derivation of the clinical and genetic risk scores, Wheeler and colleagues[11] reported that the addition of the genetic risk score to clinical factors resulted in an increase in the AUC from 0.635 to 0.675. Whether this modest

[1]Division of Rheumatology, Inflammation, and Immunity, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.80003.

Address correspondence via email to Gregory C. McDermott, MD, MPH, at gmcdermott@bwh.harvard.edu.

Submitted for publication December 18, 2025; accepted in revised form December 22, 2025.

improvement supports testing for RA-ILD risk alleles remains to be determined. Second, the focus on pulmonary fibrosis and the lack of universal high-resolution computed tomography chest screening likely resulted in underascertainment of nonfibrotic forms of RA-ILD and subclinical disease. Because the established RA-ILD genetic risk factors are specifically associated with the fibrotic usual interstitial pneumonia pattern of RA-ILD,[8] the performance of the genetic risk score may be overstated when the outcome is limited to pulmonary fibrosis and may miss ILD patterns with more inflammatory features. Although fibrotic RA-ILD has a particularly poor prognosis, the detection of other forms of RA-ILD remains clinically important.

In conclusion, Brink and colleagues[10] present an important external validation of a combined clinical and genetic risk score for the identification of pulmonary fibrosis in patients with RA. The addition of a genetic risk score to clinical factors resulted in an AUC of 0.75 for identifying patients with RA and pulmonary fibrosis. With decreasing costs of genetic sequencing and the increasing availability of consumer genetic sequencing services,[23,24] one can easily envision a future where genetic risk profiling will augment clinical risk factors for risk profiling for RA-ILD and other disease complications.

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 McDermott 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. Bongartz T, Nannini C, Medina-Velasquez YF, et al. Incidence and mortality of interstitial lung disease in rheumatoid arthritis: a population-based study. Arthritis Rheum 2010;62(6):1583–1591. doi:10.1002/art.27405
2. Hyldgaard C, Hilberg O, Pedersen AB, et al. A population-based cohort study of rheumatoid arthritis-associated interstitial lung disease: comorbidity and mortality. Ann Rheum Dis 2017;76(10):1700–1706. doi:10.1136/annrheumdis-2017-211138
3. Duarte AC, Porter JC, Leandro MJ. The lung in a cohort of rheumatoid arthritis patients-an overview of different types of involvement and treatment. Rheumatology (Oxford) 2019;58(11):2031–2038. doi:10.1093/rheumatology/kez177
4. Raimundo K, Solomon JJ, Olson AL, et al. Rheumatoid arthritis-interstitial lung disease in the United States: prevalence, incidence, and healthcare costs and mortality. J Rheumatol 2019;46(4):360–369. doi:10.3899/jrheum.171315
5. Solomon JJ, Chung JH, Cosgrove GP, et al. Predictors of mortality in rheumatoid arthritis-associated interstitial lung disease. Eur Respir J 2016;47(2):588–596. doi:10.1183/13993003.00357-2015
6. Singh N, Varghese J, England BR, et al. Impact of the pattern of interstitial lung disease on mortality in rheumatoid arthritis: a systematic literature review and meta-analysis. Semin Arthritis Rheum 2019;49(3):358–365. doi:10.1016/j.semarthrit.2019.04.005
7. Kim EJ, Collard HR, King TE Jr. Rheumatoid arthritis-associated interstitial lung disease: the relevance of histopathologic and radiographic pattern. Chest 2009;136(5):1397–1405. doi:10.1378/chest.09-0444
8. Juge P-A, Lee JS, Ebstein E, et al. MUC5B promoter variant and rheumatoid arthritis with interstitial lung disease. N Engl J Med 2018;379(23):2209–2219. doi:10.1056/NEJMoa1801562
9. McDermott GC, Hayashi K, Juge P-A, et al. Impact of sex, serostatus, and smoking on risk for rheumatoid arthritis-associated interstitial lung disease subtypes. Arthritis Care Res (Hoboken) 2025;77(2):185–194. doi:10.1002/acr.25432
10. Brink M, Wheeler A, England BR, et al. Validation of a genetic risk score combined with clinical variables for predicting pulmonary fibrosis in early rheumatoid arthritis. Arthritis Care Res (Hoboken) 2026;78(6):696–702. doi:10.1002/acr.25696
11. Wheeler AM, Baker JF, Riley T, et al. Development and internal validation of a clinical and genetic risk score for rheumatoid arthritis-associated interstitial lung disease. Rheumatology (Oxford) 2025;64(1):268–275. doi:10.1093/rheumatology/keae001
12. Allen RJ, Guillen-Guio B, Oldham JM, et al. Genome-wide association study of susceptibility to idiopathic pulmonary fibrosis. Am J Respir Crit Care Med 2020;201(5):564–574. doi:10.1164/rccm.201905-1017OC
13. Kelly CA, Saravanan V, Nisar M, et al; British Rheumatoid Interstitial Lung (BRILL) Network. Rheumatoid arthritis-related interstitial lung disease: associations, prognostic factors and physiological and radiological characteristics--a large multicentre UK study. Rheumatology (Oxford) 2014;53(9):1676–1682. doi:10.1093/rheumatology/keu165
14. Doyle TJ, Patel AS, Hatabu H, et al. Detection of rheumatoid arthritis-interstitial lung disease is enhanced by serum biomarkers. Am J Respir Crit Care Med 2015;191(12):1403–1412. doi:10.1164/rccm.201411-1950OC
15. Restrepo JF, del Rincón I, Battafarano DF, et al. Clinical and laboratory factors associated with interstitial lung disease in rheumatoid arthritis. Clin Rheumatol 2015;34(9):1529–1536. doi:10.1007/s10067-015-3025-8
16. Kiely P, Busby AD, Nikiphorou E, et al. Is incident rheumatoid arthritis interstitial lung disease associated with methotrexate treatment? Results from a multivariate analysis in the ERAS and ERAN inception cohorts. BMJ Open 2019;9(5):e028466. doi:10.1136/bmjopen-2018-028466
17. Sparks JA, Dieudé P, Hoffmann-Vold A-M, et al. Design of ANCHOR-RA: a multi-national cross-sectional study on screening for interstitial lung disease in patients with rheumatoid arthritis. BMC Rheumatol 2024;8(1):19. doi:10.1186/s41927-024-00389-4
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. doi:10.1002/art.42860
19. Juge PA, Granger B, Debray MP, et al. A risk score to detect subclinical rheumatoid arthritis-associated interstitial lung disease. Arthritis Rheumatol 2022;74(11):1755–1765. doi:10.1002/art.42162
20. Koduri GM, Podlasek A, Pattapola S, et al. Four-factor risk score for the prediction of interstitial lung disease in rheumatoid arthritis. Rheumatol Int 2023;43(8):1515–1523. doi:10.1007/s00296-023-05313-6
21. 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. doi:10.1016/j.reumae.2022.03.004

22. Paulin F, Doyle TJ, Mercado JF, et al. Development of a risk indicator score for the identification of interstitial lung disease in patients with rheumatoid arthritis. Reumatol Clin (Engl Ed) 2021;17(4):207–211. doi:10.1016/j.reumae.2019.05.006
23. Allyse MA, Robinson DH, Ferber MJ, et al. Direct-to-consumer testing 2.0: emerging models of direct-to-consumer genetic testing. Mayo Clin Proc 2018;93(1):113–120. doi:10.1016/j.mayocp.2017.11.001
24. Johansen Taber KA, Dickinson BD, Wilson M. The promise and challenges of next-generation genome sequencing for clinical care. JAMA Intern Med 2014;174(2):275–280. doi:10.1001/jamainternmed.2013.12048

Arthritis Care & Research
Vol. 78, No. 6, June 2026, pp 696–702
DOI 10.1002/acr.25696

American College of Rheumatology
Empowering Rheumatology Professionals

Validation of a Genetic Risk Score Combined With Clinical Variables for Predicting Pulmonary Fibrosis in Early Rheumatoid Arthritis

Mikael Brink,[1] Austin Wheeler,[2] Bryant R. England,[2] and Solbritt Rantapää-Dahlqvist[1]

Objective. Pulmonary fibrosis (PF) is a severe extra-articular manifestation of rheumatoid arthritis (RA). This study aimed to externally validate a genetic risk score (GRS) and a combined risk score (CRS) for predicting the risk of RA-associated PF in an independent cohort of patients with early RA.

Methods. This study used an inception cohort of 1,118 patients diagnosed with RA from northern Sweden between 1996 and 2016. Clinical data were systematically collected, and genotyping was performed for 12 single-nucleotide polymorphisms (SNPs) associated with idiopathic PF. Statistical analyses, including logistic regression and area under the curve (AUC) assessments, were conducted to evaluate the performance of the GRS and in combination with clinical data as the CRS in predicting RA-PF development.

Results. Of the 1,115 patients with complete data, 60 (5.6%) were diagnosed with PF. PF was significantly associated with age, rheumatoid factor positivity, disease activity, and *MUC5B* (rs35705950) and *FAM13A*(rs2609255) SNPs. The GRS demonstrated a significant association with RA-PF (odds ratio 2.6, 95% confidence interval 1.6–4.5), whereas the CRS exhibited superior performance (AUC 0.75, $P < 0.001$) compared to the GRS alone (AUC 0.62). The combined risk score outperformed the GRS in discriminating RA-PF, indicating its potential utility in clinical practice.

Conclusion. This study provides external validation of the Veterans Affairs Rheumatoid Arthritis Registry interstitial lung disease GRS (VARA-ILD-GRS) and the VARA-ILD-CRS in an RA cohort, demonstrating their generalizability and effectiveness in identifying individuals at high risk for RA-ILD. The findings support the integration of genetic and clinical data in risk stratification models, which could significantly improve screening strategies for patients with RA at risk of developing PF.

INTRODUCTION

Rheumatoid arthritis–associated interstitial lung disease (RA-ILD) is a severe extra-articular manifestation affecting approximately 10% of individuals with RA, leading to significant morbidity and mortality.[1,2] Despite the identification of several clinical risk factors, such as older age, male sex, smoking history, severe articular disease, and autoantibody seropositivity,[3,4] particularly in usual interstitial pneumonia,[5] there remains a lack of standardized tools for risk stratification in RA-ILD. The overlap of genetic risk factors between RA-ILD and idiopathic pulmonary fibrosis (IPF), particularly the *MUC5B* rs35705950 promoter variant, has demonstrated the potential for genetic risk variants to enhance the prediction of RA-ILD.

Wheeler et al developed and internally validated a clinical and genetic risk score (GRS) for RA-ILD using data from the Veterans Affairs Rheumatoid Arthritis Registry (VARA), a multicenter national US cohort.[6] Their study demonstrated that a combined clinical and genetic risk model outperformed clinical factors alone in discriminating RA-ILD, highlighting the utility of incorporating genetic information into risk stratification models. This approach aligns with previous findings that genetic variants associated with IPF also contribute to RA-ILD risk or pulmonary fibrosis (PF).[7–9]

Supported by the Swedish Research Council (2018-02551), King Gustaf V's 80-Year Fund, the Swedish Rheumatism Association, and Umeå University. Dr Wheeler's work was supported by the Rheumatology Research Foundation. Dr England's work was supported by the US Department of Veterans Affairs (CX002203), and the Rheumatology Research Foundation.

[1]Department of Public Health and Clinical Medicine, Rheumatology, Umeå University, Umeå, Sweden; [2]Division of Rheumatology & Immunology, Department of Internal Medicine, University of Nebraska Medical Center & VA Nebraska-Western Iowa Health Care System, Omaha, Nebraska.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25696).

Author disclosures and graphical abstract are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25696.

Address correspondence via email to Solbritt Rantapää-Dahlqvist, PhD, MD, at solbritt.rantapaa.dahlqvist@umu.se.

Submitted for publication June 10, 2025; accepted in revised form October 21, 2025.

SIGNIFICANCE & INNOVATIONS

- The study successfully validated a genetic risk score (GRS) and a combined risk score predicting rheumatoid arthritis–associated pulmonary fibrosis.
- Combining a GRS with clinical risk factors performed significantly better than the GRS alone.
- These risk scores can identify at-risk patients with RA-associated interstitial lung disease, improving the selection of patients for screening.

In this study, we aimed to externally validate the GRS developed by Wheeler et al in an independent RA cohort. By using the same genetic loci and clinical data collection methods, we sought to test the robustness and generalizability of the VARA RA-ILD GRS for predicting development of PF in an RA population. With validation of the clinical utility of the GRS in identifying individuals at high risk for RA-ILD, these tools could support targeted screening and early intervention strategies.

PATIENTS AND METHODS

Patients and collection of clinical data. The cohort of patients with early RA (eRA) has previously been used to investigate the associations between selected genotypes and PF.[8] Briefly, this inception cohort included patients diagnosed with eRA according to the American College of Rheumatology classification criteria.[10] Patients were consecutively included at the time point of RA diagnosis between January 1, 1996, and December 31, 2016, at five rheumatology clinics in northern Sweden. All patients were observed until December 31, 2016, or death, with a mean follow-up time of 9.4 (SD 4.9) years. Clinical data, including the Disease Activity Score in 28 joints (DAS28),[11] pharmacological treatments, and smoking history, were systematically recorded at baseline and at 6, 12, 18, and 24 months after diagnosis and subsequently at all clinical visits. The presence of rheumatoid factor (RF) was assessed at baseline using routine laboratory methods, and the presence of anti–cyclic citrullinated peptide (anti-CCP) antibodies was assessed using an anti-CCP2 test (Euro-Diagnostica AB), as previously described.[8]

Genotyping. DNA samples were collected at baseline and stored at −80°C. Genotyping was performed using the Global Screen Assay (Illumina) to analyze 571,151 genome-wide single-nucleotide polymorphisms (SNPs) at deCode Genetics, as previously reported.[12] Information for selected gene variants was extracted from the genotype dataset. The genome-wide association study (GWAS) data used were stored by the National Bioinformatics Infrastructure Sweden at SciLifeLab. The SNP for *MUC5B* (rs35705950) was imputed, and the sequence variants for imputation were identified through whole-genome sequencing of 67,645 individuals of Islandic, Danish, Norwegian, and Swedish origin.[12] Samples (n = 193) showing quality scores <90% were reanalyzed using quantitative polymerase chain reaction (qPCR) assays and TaqMan SNP Genotyping Assays (Applied Biosystems) specific for the rs35705950 SNP. For this study, genotype data at 12 loci were extracted: rs35705950 (*MUC5B*), rs2736100 (*TERT*), rs111521887 and rs5743890 (*TOLLIP*), rs2076295 (*DSP*), rs7887 (*EHMT2*), rs2034650 (*IVD*), rs2609255 (*FAM13A*), rs4727443 (*LOC100128334/LOC105375423*), rs11191865 (*OBFC1*), rs1278769 (*ATP11A*), rs6793295 (*LRRC34*), and rs12610495 (*DPP9*). Genetic ancestry analysis was performed for Danish, Swedish, and Norwegian sample sets separately included in the GWAS.[12]

Pulmonary examinations. Radiographs of the lungs were obtained at baseline and during follow-up if patients presented any symptoms of cough, dyspnea, or chest pain or before initiating biologic disease-modifying antirheumatic drugs. High-resolution computed tomography (HRCT) of the chest was performed in cases with signs of pathology on routine plain radiographs or clinically suspected pulmonary involvement. The diagnosis of PF was based on HRCT findings, with reticular patterns, honeycombing, or traction bronchiectasis and occasionally with ground-glass in a few cases being indicative of PF.[13] In five patients, the diagnosis of PF was already evident and diagnosed on the routine plain radiograph examinations. In this study, no further diagnostic evaluation was performed to separate usual interstitial pneumonia and nonspecific interstitial pneumonia disease. Clinical and pulmonary examinations have previously been presented in detail.[8]

Statistical analysis. Statistical analyses were conducted using R software version 4.4.1 (R Foundation for Statistical Computing).[14] Descriptive data were summarized as proportions, means, or medians. Frequencies were compared using Pearson's chi-square test with Yates' continuity correction, and for continuous data, Student's *t*-test was used. A Mann-Whitney U test was conducted to compare the different GRS between RA-PF and RA without PF. Associations between PF and potential predictors, including genetic markers, were analyzed using logistic regression analyses, presented as odds ratios (ORs) with 95% confidence intervals (CIs). Areas under the curve (AUCs) were compared using DeLong's test for two correlated receiver operating characteristic curves using the pROC library v.1.18.5.[15] For eight individuals, the mean DAS28 value for the first 24 months ($DAS28_{mean24m}$) was imputed using predictive mean matching, and ever smoking ($smoking_{ever}$) was imputed for 19 individuals using logistic regression by using the mice library v. 3.16.0.[16]

Calculation of genetic and clinical risk scores. All risk scores were calculated as previously reported by Wheeler et al.[6] In brief, the GRS was calculated with each of the five SNPs

represented, assuming autosomal dominant inheritance as 0 (wild type) or 1 (variant allele): VARA-ILD-GRS = *MUC5B* × ln(2.47) + *DSP* × ln(1.21) + *LRRC34* × ln(1.08) + *OBFC1* × ln(1.19) + *FAM13A* × ln(0.74).

Similarly, the VARA-ILD combined risk score (VARA-ILD-CRS) model was calculated as reported by Wheeler et al, but with small modifications. In the original score, the variable β coefficient (ie, ln OR) for categorical variables (smoking history, male sex, RF positivity) was multiplied by 1 if the factor was present or 0 if the factor was absent. For age, 69.5 from the individual's current age was subtracted, and the difference was multiplied by the β coefficient, and the individual's mean DAS28 using the C-reactive protein level (DAS28-CRP) was multiplied by the β coefficient. Modifications of this original score were performed for disease activity and age. In our validation cohort, we instead used the mean DAS28 using the erythrocyte sedimentation rate (DAS28-ESR) over the first 24 months after diagnosis to catch the burden of early inflammation, and instead of using age at the most recent encounter, we used age at the time point of diagnosis of RA (continuing to center age by subtracting 69.5). The modified VARA-ILD-CRS (mVARA-ILD-CRS) = 0.9857579 × GRS + 0.0249686 × ($age_{inclusion}$ − 69.5) + 0.2992501 × sex_{male} + 0.3963635 × $smoking_{ever}$ + 0.2376407 × $DAS28_{mean24m}$ + 0.610918 × $RF_{positive}$ + (−4.442266). The probability was then, as in the original publication, obtained by calculating exp(mVARA-ILD-CRS)/(1 + exp[mVARA-ILD-CRS]).

Ethics. The study complies with the Declaration of Helsinki, and the Regional Ethics Committees at Umeå University, Sweden, approved this study (no. Dnr 2017-432-32M, 2019-02039). The patients provided informed consent to participate.

RESULTS

Patient characteristics and variant allele frequencies associated with RA-PF. Of the 1,118 patients with eRA included in our inception cohort, a GRS could successfully be calculated for 1,115 individuals with complete data, of whom 60 (5.6%) were diagnosed with PF at the time of RA diagnosis or developed PF during follow-up. Variables related to development of PF were age at inclusion ($P < 0.001$), RF positivity ($P < 0.05$), and $DAS28_{mean24m}$ ($P < 0.05$) (Table 1). Of the 12 analyzed SNPs, two were significantly associated with RA-PF in unadjusted analyses: *MUC5B* (rs35705950) (OR 3.70 [95% CI 2.18–6.32], $P < 0.001$) and *FAM13A* (rs2609255) (OR 1.91 [95% CI 1.13–3.25], $P < 0.05$) (Supplementary Table 1).

Model performance of GRS in the validation cohort. The five included SNPs (*MUC5B* [rs35705950], *DSP* [rs2076295], *LRRC34* [rs6793295], *OBFC1* [rs11191865], and *FAM13A* [rs2609255]) in the GRS constructed by Wheeler et al[6] were analyzed using multivariable logistic regression analyses, with two SNPs continuing to show significant association with RA-PF: rs35705950 (*MUC5B*, OR 3.02 [95% CI 1.95–4.65], $P < 0.001$) and rs2609255 (*FAM13A*, OR 1.82 [95% CI 1.20–2.72], $P < 0.01$).

The GRS calculated according to Wheeler et al showed a significant association with RA-PF in our cohort, with higher values in RA-PF compared to RA without PF, as expected (mean 0.63, median 0.79, range −0.13 to 1.3 and mean 0.41, median 0.27, range −0.3 to 1.3, respectively; $P < 0.01$) (Supplementary Figure 1). The OR for RA-PF increased for each unit of the GRS at 2.6 (95% CI 1.6–4.5; $P < 0.0001$).

Model performance of CRS in the validation cohort. Univariable logistic regression for the clinical variables showed significant association with RA-PF for age, $DAS28_{mean24m}$, and RF positivity ($P < 0.05$ for all) (Table 2). In a multivariable logistic regression analysis including the same clinical variables, the significance remained for age and RF ($P < 0.01$) (data not shown).

Table 1. Demographic and clinical data of the included patients with eRA, stratified for development of PF*

	eRA with pulmonary fibrosis (n = 60)	eRA without fibrosis (n = 1,058)
Age at RA diagnosis, mean (SD), y	64.8 (10.3)[a]	57.4 (14.1)
Age at PF diagnosis, mean (SD), y	73.9 (9.6)	–
Male sex, n (%)	23 (38.3)	311 (29.4)
Ever smoker, n (%)	43/59 (72.9)	644/1,041 (60.9)
Current smoker, n (%)	11/58 (19.0)	203/1,043 (19.5)
ACPA+, n (%)	43 (80.0)	705 (69.2)
RF+, n (%)	53 (88.3)[b]	770 (72.8)
HLA-SE+, n (%)	23 (63.9)	371 (58.7)
Follow-up time from RA diagnosis until Dec 31, 2016, or death, mean (SD), y	8.7 (4.8)	9.5 (5.0)
BMI, mean (SD) kg/m^2	27.0 (3.7)[c]	26.4 (4.5)[c]
DAS28-ESR (index date), mean (SD)	4.8 (1.4)	4.8 (1.4)
DAS28-ESR AUC_{24},[d] mean (SD)	90.0 (18.8)[e]	82.7 (21.5)
Death during study period, n (%)	27 (45.0)[a]	130 (12.3)
Duration of RA at time point of PF diagnosis, mean (min to max), y	5.4 (−0.58 to 15.6)	–
HRCT, n (%)	55/60 (91.7)	–

* ACPA, anti–citrullinated protein/peptide antibody; AUC, area under the curve; BMI, body mass index; DAS28, Disease Activity Score in 28 joints; eRA, early RA; ESR, erythrocyte sedimentation rate; HLA-SE, HLA shared epitope, HRCT, high-resolution computed tomography; IQR, interquartile range; PF, pulmonary fibrosis; RA, rheumatoid arthritis; RF, rheumatoid factor.
[a] $P < 0.001$.
[b] $P < 0.01$.
[c] Data available in 50 and 849 patients, respectively.
[d] AUC calculated for the first 24 months after diagnosis.
[e] $P < 0.05$.

Table 2. Multivariable logistic regression models for variables used in the combined risk score for outcome in patients with rheumatoid arthritis with or without pulmonary fibrosis in the validation cohort and in the original cohort according to Wheeler et al[6]*

	Validation cohort model				Original cohort model	
	Univariable model		Multivariable model		Multivariable model	
Variable	OR (95% CI), *P* value	β coefficient	OR (95% CI), *P* value	β coefficient	OR (95% CI), *P* value	β coefficient
Genetic risk score	2.64 (1.55–4.50), <0.001	0.9719446	2.36 (1.24–4.48), <0.01	0.8576770	2.68 (1.95–3.69), <0.001	0.9857579
Age, per 1-y increase over age 69.5 y	1.05 (1.02–1.07), <0.001	0.0473284	1.04 (1.01–1.07), <0.01	0.0416284	1.03 (1.01–1.04), 0.001	0.0249686
Male sex	1.49 (0.87–2.55), 0.143	0.4008486	1.07 (0.9–3.19), 0.1	0.5298953	1.35 (0.74–2.47), 0.33	0.2992501
Smoking history (ever)	1.56 (0.88–2.78), 0.128	0.4462943	1.40 (0.7–2.82), 0.343	0.3375456	1.49 (1.00–2.20), 0.05	0.3963635
Mean DAS28-ESR during 24 mo/DAS28-CRP	1.42 (1.05–1.9), <0.05	0.3472252	1.42 (0.99–2.02), 0.055	0.3471921	1.27 (1.12–1.44), <0.001	0.2376407
RF positive	2.83 (1.27–6.3), <0.05	1.0409517	2.96 (1.14–7.7), <0.05	1.0859854	1.84 (1.23–2.76), 0.003	0.610918
Model constant	–	–	–	−5.5932484	–	−4.442266

* CI, confidence interval; CRP, C-reactive protein; DAS28, Disease Activity Score in 28 joints; OR, odds ratio; RF, rheumatoid factor.

When also including the GRS as a covariate in the model, the GRS, age, and RF positivity remained significantly predictive of PF ($P < 0.05$) (Table 2).

The GRS was then combined with clinical variables into the VARA-ILD-CRS and validated in our cohort. The CRS was found to be significantly associated with RA-PF development, with higher values in RA-PF compared to RA without PF, as expected (mean 0.12, median 0.12, range 0.03–0.24 and mean 0.076, median 0.064, range 0.0052–0.34, respectively; $P < 0.0001$) (Supplementary Figure 2). The OR for RA-PF using the CRS was 1.72 (95% CI 1.43–2.1, $P < 0.0001$) per 0.05-unit change in the CRS.

The CRS was categorized into quartiles, with the highest quartile (>0.10016) showing an OR for RA-PF of 2.7 (95% CI 1.8–5.0), and that was significantly higher compared to the lowest quartile (<0.04015) ($P < 0.0001$) (data not shown). The CRS discriminated RA-PF with an AUC of 0.75 (95% CI 0.69–0.81), which was significantly better than the GRS showing an AUC of 0.62 (95% CI 0.54–0.69) ($P < 0.001$) (Figure 1). To validate the performance of the CRS, a table was created for different cutoff levels and their sensitivity, specificity, positive and negative predictive values, and positive and negative likelihood ratios (Supplementary Table 2). The optimal cutoff by maximizing the Youden index was ≥0.07999, yielding a sensitivity of 0.789 and a specificity of 0.653. Setting the sensitivity to ≥0.9 led to a cutoff of 0.05, as in the derivation study.[6] Using this cutoff would eliminate 36.1% of the cohort from needing to undergo further testing. In contrast, prioritizing specificity (≥0.9) resulted in a sensitivity of 0.316 and a cutoff value of 0.1458237.

Because the components have different scales, we evaluated which variables contributed the most to the CRS. We present the distribution of the contribution for the RA-PF and RA without PF groups separately in Figure 2. The highest average contributions to the CRS were found for DAS28-ESR$_{mean24m}$, RF, and GRS. In relative terms, the GRS and age variables differed the most between RA-PF and RA without PF.

DISCUSSION

In this study of 1,115 patients with eRA, we externally validated a GRS and a CRS for predicting the risk of RA-PF. The original score was developed for RA-ILD in a male-predominant US veteran RA cohort.[6] This external validation study evaluated the application of this score in an inception RA cohort from northern Sweden, with a higher proportion of females. Despite the differences in these cohorts, the performances of the GRS and CRS were similar. Thus, this RA-ILD CRS appears to be a promising tool for use to risk stratify RA-ILD and inform RA-ILD screening approaches.

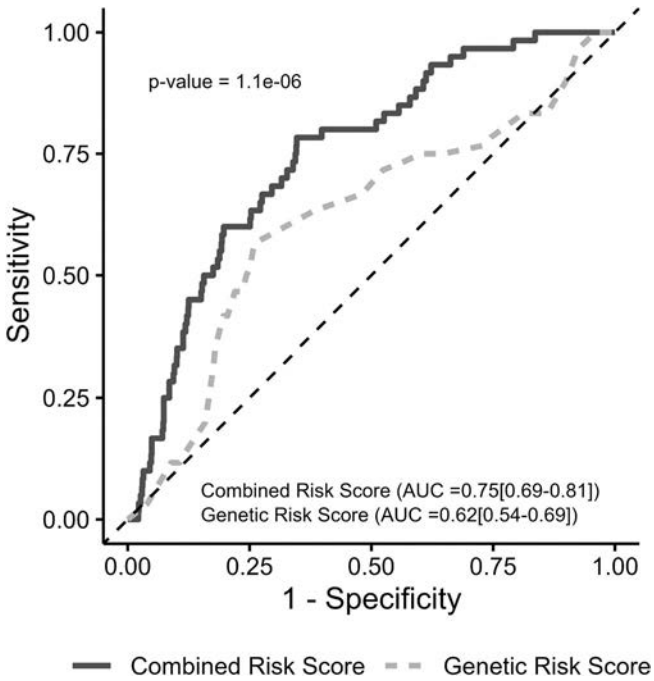


Figure 1. Comparison of receiver operating characteristic curves for combined and genetic risk scores for the presence of rheumatoid arthritis–associated pulmonary fibrosis. AUC, area under the curve.

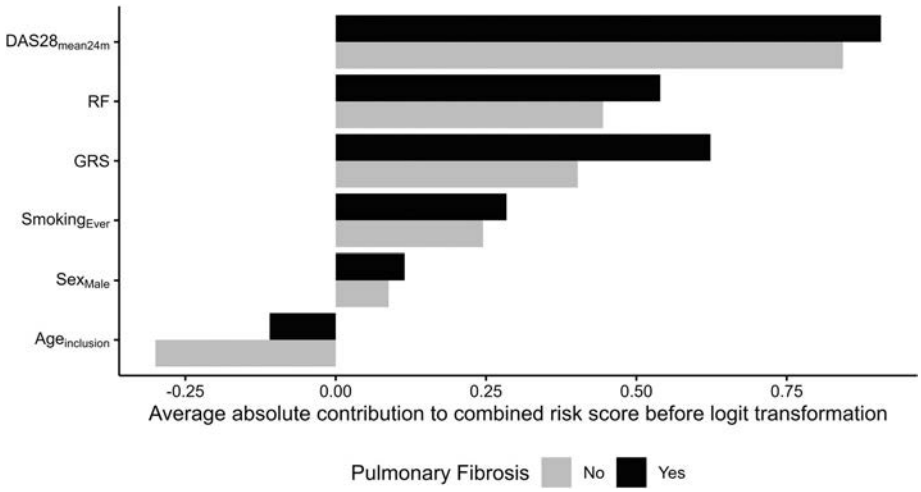


Figure 2. Average contribution of each variable included in the combined risk score, stratified by the presence of pulmonary fibrosis. DAS28, Disease Activity Score in 28 joints; GRS, genetic risk score; RF, rheumatoid factor.

Discrimination of PF in RA with the CRS was higher in our cohort (AUC 0.75 [95% CI 0.69–0.81]) than in the original publication (AUC 0.675 [95% CI 0.639–0.711]), demonstrating the validity of this tool in an alternative study population. As in the original study, we found the CRS to outperform the GRS alone. The similar results are quite remarkable given different frequencies of the genotypes, study populations, and variability of clinical data that were included. Thus, it would be expected for the VARA-ILD-CRS to generalize to other study populations and settings. Despite our patients being of Scandinavian ancestry, there was good agreement between the VARA-ILD-GRS and the VARA-ILD-CRS in RA based on an American population.

We confirmed that the *MUC5B* promotor variant was the strongest genetic risk factor, as was shown in our previous publication.[8] We also found a significant association with *FAM13A* (rs2609255), which was not significant in the study by Wheeler et al.[6] This supports the continued need for evaluation of genetic variants among diverse cohorts, as findings for specific SNPs outside of *MUC5B* have varied across populations. This also highlights the value of using a GRS, such as the VARA-ILD-GRS, which allows for a broader representation of our current understanding of genetic risk and moderates the nuanced differences in individual allele frequencies and associations across populations. The performance of the GRS in this study demonstrates the external validity of such approaches.

Polygenic risk scores for IPF, but not RA, were associated with RA-ILD in the COPDgene cohort.[17] In another study evaluation of two large cohorts (MESA and COPDgene), RA-related HLA-DRB1 was not associated with interstitial lung abnormalities.[18]

Our cohort was composed of consecutively included patients with eRA between 1996 and 2016, whereas Wheeler et al used a prospective RA cohort of US veterans with more established disease initiated in 2003. Both cohorts had a high frequency of cigarette smoking history (current cohort 60%–70% vs VARA cohort 77%–85%). Perhaps related to these high cigarette smoking frequencies, we were not able to detect a significant contribution of smoking to RA-PF risk in the multivariable model, despite smoking having previously been established as a risk factor for both RA and RA-PF.[3,19] Given high background smoking frequency in these cohorts, it is possible that the genetic risk component is more discriminative for RA-PF vs RA alone. Validation of the GRS and CRS in a cohort with less frequent cigarette smoking history would be valuable.

Because of differences in cohorts and data availability, some components of the CRS required modification. Wheeler et al calculated the mean DAS28-CRP preceding ILD diagnosis but also showed consistent performance when using the most recent DAS28-CRP. In adapting the CRS to this cohort, we calculated the DAS28-ESR as a mean value over the first 24 months. This measure of disease activity was similarly associated with RA-PF, suggesting that alternative measures of average disease activity using DAS28 variants are suitable for calculating the CRS. Whether this extends to alternative RA disease activity measures, such as the Clinical Disease Activity Index or Routine Assessment of Patient Index Data 3, and adjustment of these disease activity measure values to maintain consistent CRS values will need to be examined in future studies.

The sensitivity and specificity of the CRS were comparable to those reported previously by Wheeler et al.[6] At the cutoff of 0.05 proposed previously to ensure a 90% sensitivity for RA-ILD detection, the sensitivity was 93.3% and the specificity was 37.7% in the current cohort. Applying this cutoff would exclude approximately 36% of the patients with eRA in our study from the need for additional testing with HRCT and pulmonary function

tests (PFTs). If providers were comfortable with a test sensitivity of 80%, a cutoff of 0.07 could be used. At this value, the proportion of patients with eRA in our cohort who would avoid the need for additional ILD testing increases to 55%. This illustrates the potential value of an RA-ILD screening strategy that incorporates a highly sensitive VARA-ILD-CRS to identify patients with RA who warrant further ILD testing with HRCT and PFTs.

MUC5B in our cohort was determined from imputation, with a smaller number showing quality scores <90% that were reanalyzed using qPCR assays and TaqMan SNP Genotyping Assays. There was very good agreement between those cases tested by both methods. The sequence variants for the imputation were identified using cases with Scandinavian origin.[12] These findings suggest that imputation is a valid method for ascertaining *MUC5B* status.

A limitation of our study is that the HRCT examinations were not performed randomly or on all included patients. Rather, these were performed among patients with abnormalities on the plain radiographs or who had clinical indications for HRCT with the development of defined symptoms or signs suspicious for RA-PF. The HRCT examinations have been performed over a period of almost 20 years, and methodologic improvements during this time could affect the results. Consequently, we refrained from further diagnostic evaluations besides PF.

In summary, we externally validated the VARA-ILD-GRS and VARA-ILD-CRS among a cohort of patients with eRA based in northern Sweden. Despite differences in the study populations and implementing adaptations to the clinical risk score components, our findings showed a similar ability to discriminate RA-PF and performance metrics at various cut points. These results emphasize the potential role of combined genetic and clinical risk scores to inform RA-ILD screening strategies.

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 Rantapää-Dahlqvist 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. Marigliano B, Soriano A, Margiotta D, et al. Lung involvement in connective tissue diseases: a comprehensive review and a focus on rheumatoid arthritis. Autoimmun Rev 2013;12(11):1076–1084. doi:https://doi.org/10.1016/j.autrev.2013.05.001
2. Spagnolo P, Lee JS, Sverzellati N, et al. The lung in rheumatoid arthritis: focus on interstitial lung disease. Arthritis Rheumatol 2018;70(10):1544–1554. doi:https://doi.org/10.1002/art.40574
3. Kelly CA, Saravanan V, Nisar M, et al; British Rheumatoid Interstitial Lung (BRILL) Network. Rheumatoid arthritis-related interstitial lung disease: associations, prognostic factors and physiological and radiological characteristics--a large multicentre UK study. Rheumatology (Oxford) 2014;53(9):1676–1682. doi:https://doi.org/10.1093/rheumatology/keu165
4. Chartrand S, Lee JS, Swigris JJ, et al. Clinical characteristics and natural history of autoimmune forms of interstitial lung disease: a single-center experience. Lung 2019;197(6):709–713. doi:https://doi.org/10.1007/s00408-019-00276-7
5. McDermott GC, Hayashi K, Juge P-A, et al. Impact of sex, serostatus, and smoking on risk for rheumatoid arthritis-associated interstitial lung disease subtypes. Arthritis Care Res (Hoboken) 2025;77(2):185–194. doi:https://doi.org/10.1002/acr.25432
6. Wheeler AM, Baker JF, Riley T, et al. Development and internal validation of a clinical and genetic risk score for rheumatoid arthritis-associated interstitial lung disease. Rheumatology (Oxford) 2025;64(1):268–275. doi:https://doi.org/10.1093/rheumatology/keae001
7. Juge PA, Lee JS, Ebstein E, et al. MUC5B promoter variant and rheumatoid arthritis with interstitial lung disease. N Engl J Med 2018;379(23):2209–2219. doi:https://doi.org/10.1056/NEJMoa1801562
8. Jönsson E, Ljung L, Norrman E, et al. Pulmonary fibrosis in relation to genetic loci in an inception cohort of patients with early rheumatoid arthritis from northern Sweden. Rheumatology (Oxford) 2022;61(3):943–952. doi:https://doi.org/10.1093/rheumatology/keab441
9. Klein J, Wheeler AM, Baker JF, et al. MUC5B Promoter variant and survival in rheumatoid arthritis-associated interstitial lung disease. Rheumatology (Oxford) 2025;64(SI):SI47–SI54. doi:https://doi.org/10.1093/rheumatology/keae615
10. Arnett FC, Edworthy SM, Bloch DA, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31(3):315–324. doi:https://doi.org/10.1002/art.1780310302
11. Prevoo MLL, van 't Hof MA, Kuper HH, et al. Modified disease activity scores that include twenty-eight-joint counts. Development and validation in a prospective longitudinal study of patients with rheumatoid arthritis. Arthritis Rheum 1995;38(1):44–48. doi:https://doi.org/10.1002/art.1780380107
12. Saevarsdottir S, Stefansdottir L, Sulem P, et al; Members of the DBDS Genomic Consortium; Danish RA Genetics Working Group; Swedish Rheumatology Quality Register Biobank Study Group (SRQb). Multiomics analysis of rheumatoid arthritis yields sequence variants that have large effects on risk of the seropositive subset. Ann Rheum Dis 2022;81(8):1085–1095. doi:https://doi.org/10.1136/annrheumdis-2021-221754
13. Raghu G, Remy-Jardin M, Myers JL, et al; American Thoracic Society, European Respiratory Society, Japanese Respiratory Society, and Latin American Thoracic Society. Diagnosis of idiopathic pulmonary fibrosis. An official ATS/ERS/JRS/ALAT clinical practice guideline. Am J Respir Crit Care Med 2018;198(5):e44–e68. doi:https://doi.org/10.1164/rccm.201807-1255ST
14. R Development Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing; 2022.
15. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 2011;12(1):77. doi:https://doi.org/10.1186/1471-2105-12-77
16. van Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. J Stat Softw 2011;45(3):1–67. doi:https://doi.org/10.18637/jss.v045.i03
17. McDermott GC, Moll M, Cho MH, et al. Polygenic risk scores for rheumatoid arthritis and idiopathic pulmonary fibrosis and associations with RA, interstitial lung abnormalities, and quantitative interstitial

abnormalities among smokers. Semin Arthritis Rheum 2025;72:152708. doi:https://doi.org/10.1016/j.semarthrit.2025.152708

18. Kim JS, Flack KF, Malik V, et al. Genomic and serological rheumatoid arthritis biomarkers, *MUC5B* promoter variant, and interstitial lung abnormalities. Ann Am Thorac Soc 2025;22(1):64–71. doi:https://doi.org/10.1513/AnnalsATS.202403-238OC
19. Klareskog L, Stolt P, Lundberg K, et al. A new model for an etiology of rheumatoid arthritis: smoking may trigger HLA-DR (shared epitope)-restricted immune reactions to autoantigens modified by citrullination. Arthritis Rheum 2006;54(1):38–46. doi:https://doi.org/10.1002/art.21575

Arthritis Care & Research
Vol. 78, No. 6, June 2026, pp 703–712
DOI 10.1002/acr.25691

American College of Rheumatology
Empowering Rheumatology Professionals

Mortality in US Veterans With Rheumatoid Arthritis Treated With Immune Checkpoint Inhibitors

Madeline L. O'Sullivan,[1] Grant W. Cannon,[1] Brian C. Sauer,[1] Jorge Rojas,[2,3] Gary A. Kunkel,[1] Jessica A. Walsh,[1] Shardool Patel,[1] Punyasha Roul,[4] Joshua F. Baker,[5] Bryant R. England,[4] Ted R. Mikuls,[4] and Tawnie J. Braaten[1]

Objective. This investigation compared all-cause and cause-specific mortality in patients with and without rheumatoid arthritis (RA) in the Veterans Health Administration (VHA) following immune checkpoint inhibitor (ICI) cancer treatment.

Methods. Veterans with RA and a control set of Veterans without RA who were matched on age, sex, and year of Veterans Affairs enrollment and had received an ICI were identified. All-cause and cause-specific mortality were obtained. Survival from the time of ICI initiation was evaluated using Cox models, Kaplan-Meier curves, and log rank testing.

Results. There were 301 patients with RA and 2,114 controls without RA treated with an ICI. The majority of the participants were white, male, and current/former smokers. Lung cancer was the most common malignancy (51.2%), pembrolizumab was the most frequently used ICI (43.9%), and most patients received ICI monotherapy (97.1%). Cox proportional hazard ratio comparison of all-cause mortality in patients with RA to controls without RA was 1.08 (95% confidence interval [CI] 0.94–1.25) for the crude analysis and 1.09 (95% CI 0.94–1.25) for the adjusted analysis. Cause of death was similar in the two groups, most frequently neoplasm in 93.0% and 90.9% for RA and non-RA groups, respectively ($P = 0.737$). Deaths due to infection were rare in both groups (<1.0%).

Conclusion. Patients with RA who received ICIs for the treatment of malignancy did not experience increased mortality or differences in cause of death compared with patients without RA receiving ICIs. These preliminary data suggest ICI therapy may be considered as part of cancer treatment in RA patients based on individual patient circumstances.

INTRODUCTION

Immune checkpoint inhibitors (ICIs) have provided a breakthrough in the treatment of several cancers, including advanced and metastatic disease. ICIs block inhibitory signals of T-cell activation to allow tumor-reactive T-cells to mount an effective antitumor response. Despite their effectiveness, ICIs are known to cause a broad spectrum of unintended immune-related adverse events (irAEs) affecting multiple organ systems.[1,2] The severity of these irAEs varies widely, with more severe forms requiring cessation of ICIs and sometimes systemic immunosuppression. Because of concerns that patients with underlying autoimmune disease may experience more severe, even life-threatening irAEs, flares of their underlying autoimmune disease, or that baseline immunosuppression may decrease ICI effectiveness, these patients were excluded from clinical trials.[3,4] Polymorphisms in programmed cell death protein (PD)-1 and Cytotoxic T-lymphocyte associated protein-4, the targets of ICI therapy, have also been associated with rheumatoid arthritis (RA) risk, which raises further concern that patients with RA may be at higher risk of treatment-related AEs.[5–7]

It remains unknown whether these theoretical risks result in an actual increased overall mortality or differences in cause of death in patients with autoimmune disease. Patients with RA

Supported by the Center of Excellence for Suicide Prevention, Joint Department of Veterans Affairs (VA) and Department of Defense (DoD) Suicide Data Repository – National Death Index and the Western Institute for Veterans Research. Dr England's work was supported by the VA Clinical Science Research and Development (IK2 CX002203) and the Rheumatology Research Foundation. Dr Mikuls' work was supported by the VA (Merit Award BX003635), DoD (grant PR200793), and NIH (grant U54-GM-115458).

[1]Salt Lake City VA and University of Utah; [2]Puget Sound VA Medical Center, Seattle, Washington; [3]University of Utah, Salt Lake City; [4]VA Nebraska-Western Iowa Health Care System and University of Nebraska Medical Center, Omaha; [5]Corporal Michael J. Crescenz VA and University of Pennsylvania, Philadelphia.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25691).

Author disclosures and graphical abstract are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25691.

Address correspondence via email to Grant W. Cannon, MD, at grant.cannon@va.gov.

Submitted for publication March 3, 2025; accepted in revised form October 7, 2025.

SIGNIFICANCE & INNOVATIONS

- This is the largest study to date examining both all-cause and cause-specific mortality of patients with pre-existing rheumatoid arthritis (RA) treated with immune checkpoint inhibitors (ICIs).
- No differences in cause-specific or all-cause mortality were observed in patients with RA in comparison with controls without RA.
- These data suggest that treatment with ICIs should be considered in patients with RA if indicated as part of cancer management.

undergoing cancer treatment with ICIs are of particular interest given the relatively high prevalence and the increased risk of malignancy (ie, lung cancer) inherent to the disease, and therefore may more frequently require ICI therapy.[8] There is a critical need to more fully understand the safety of ICIs in patients with RA because most previous studies have focused primarily on rates of, and treatments for, RA flares complicating cancer treatment.[3,9] Much of the existing literature did not report on treatment efficacy, mortality rates, or specific cause of death in this population.[8]

More data about the use of ICIs in patients with RA will help guide providers and patients to make informed decisions when weighing potential risks and benefits of cancer treatment options. Therefore, the objective of this investigation was to determine and compare mortality rates and specific causes of death in US Veterans with and without pre-existing RA undergoing cancer treatment with ICIs.

METHODS

Study design and patient selection

This retrospective cohort study evaluated US Veterans enrolled in the Veterans Health Administration (VHA). The VHA Corporate Data Warehouse (CDW) provided clinical and administrative data for care delivered by VHA providers from 145 VHA medical centers. This study was approved by the University of Utah Institutional Review Board (IRB #12917).

We classified Veterans as having RA using a validated algorithm requiring that the patient had both (1) two or more *International Classification of Diseases* (ICD) codes for RA at least 30 days apart, with at least one ICD code from a rheumatologist, and (2) treatment with a disease modifying anti-rheumatic drug (DMARD) or laboratory evidence of a positive rheumatoid factor or anti-cyclic citrullinated peptide (anti-CCP) antibody.[10] Patients were included as an RA case if RA criteria were met any time before the first ICI infusion using all data available in the CDW from 1999 moving forward to the time of first ICI infusion. Patients meeting RA criteria after ICI infusion were excluded. Rheumatology encounters were identified in CDW as clinic visits using the Veterans Affairs (VA) rheumatology stop code 315. Control patients were selected from a selected cohort of patients without RA. This non-RA cohort consisted of up to ten patients not fulfilling RA criteria that were matched to each patient with RA with the same year of birth, sex, and year of initial VHA enrollment. Patients in the RA group and non-RA control group receiving one or more ICI infusions between June 6, 2011, and February 14, 2023, were identified using pharmacy databases and included in this study (Supplementary Figure 1). Inclusion in the study required only a single ICI infusion because our goal was to evaluate survival based on the intent to initiate ICI therapy.

Patient characteristics

The CDW data domains used included patient (eg, demographics), inpatient, and outpatient encounters; inpatient and outpatient pharmacy dispensing; vital signs; and laboratory and chemistry domains. We used the Compensation and Pension Records Interchange system, which provides national "read-only" access to Veteran electronic health records and medical notes to inform algorithm development and confirm the correct classification of clinical concepts.[11,12] Demographic data and baseline patient characteristics included age at first ICI infusion, sex, race, ethnicity, smoking status, and laboratory reports of rheumatoid factor and anti-CCP. Sex, race, and ethnicity were self-reported and recorded in CDW. The Care Assessment Needs (CAN) classification[13] is a VA comorbidity score that uses clinical and demographic characteristics to predict future hospitalization and mortality.[14] The CAN score includes predictors across six categories: sociodemographics, medical conditions, vital signs, previous use of health services, medications, and laboratory tests, and is expressed as a percentile of probabilities ranging from 0 percentile (lowest risk) to 99th percentile (highest risk). The CAN score also provides a mortality estimate over the following 90 days and 1 year. Health care use was measured as the number of hospitalizations and clinical encounters documented in CDW during the year before ICI therapy.

Cancer diagnosis and stage

The cancer diagnosis and cancer stage was ascertained from the VA Oncology Raw Domain, which contributes to the VA Central Cancer Registry.[15] For reporting and analysis, cancers were grouped as gastrointestinal, genitourinary, head and neck, lung (non-small cell lung cancer [NSCLC]), lung–other (not NSCLC), melanoma, and other.

ICI treatment

Pharmacy CDW data for atezolizumab, avelumab, cemiplimab, durvalumab, ipilimumab, nivolumab, and pembrolizumab were extracted. For each patient, the ICI infusion data collected

were the date of the first ICI infusion, the number of unique agents received, the number of infusions for each agent, and the duration between the first and last ICI infusion. Patients were classified as combination ICI therapy if two different ICIs were used within two weeks of the initial ICI infusion (initial ICI combination therapy) or if a new and different ICI was added within two weeks after at least two infusions of the same ICI ("base" ICI drug) and the base ICI drug was infused again after the new and different ICI (add-on combination therapy). For example, if a patient had been receiving multiple infusions of ipilimumab and then had an infusion of nivolumab that was followed by another infusion of ipilimumab after four weeks, this pattern would qualify as an "add-on" combination therapy with initial monotherapy. In contrast, if a patient was treated with ipilimumab for several infusions and then treated with nivolumab without subsequent ipilimumab, this pattern would be classified as a "switch" in therapy and not combination therapy.

Cancer treatments before ICI infusions

Both oral and intravenous cancer treatments, other than ICI, were identified in CDW for each patient. The specific agents and number of unique drugs by each patient before ICI treatment were identified.

DMARD treatment

CDW pharmacy data were searched for any exposure to the following DMARDs before the first ICI infusion for each patient: abatacept, adalimumab, anakinra, azathioprine, baricitinib, certolizumab, etanercept, golimumab, hydroxychloroquine, infliximab, leflunomide, methotrexate, rituximab, sarilumab, sulfasalazine, tocilizumab, tofacitinib, and upadacitinib. Case and control subjects with any DMARD exposure during the 90 days before ICI were reported. All exposures to methotrexate and rituximab were reported irrespective of indication for either RA or cancer treatment. Control patients receiving DMARDs other than methotrexate and rituximab were reviewed for indication.

Glucocorticoid treatment

Both intravenous and oral dispensing episodes for the following glucocorticoid drugs—betamethasone, budesonide, dexamethasone, hydrocortisone, methylprednisolone, prednisolone, and prednisone—were extracted from CDW for each patient receiving an ICI. For each dispensing episode, the date of medication dispensing, medication unit dose, prednisone equivalent unit dose, number of units dispensed, and days of treatment were recorded. Using these data, patients were identified with any glucocorticoid exposure at any time before and within 90 days before first ICI infusion; any prednisone exposure any time before and within 90 days before first ICI infusion; intravenous and oral dexamethasone exposure within 90 days before first ICI infusion; and total cumulative glucocorticoid, dexamethasone, and prednisone during the 90 days before first ICI infusion. Total glucocorticoid exposure during the 90 days before initial ICI treatment was used for the covariate analysis. All glucocorticoid doses were reported as prednisone equivalent.

Study outcome

The primary outcome was survival with cause-specific mortality being a secondary outcome. The VA-maintained Death Ascertainment File (DAF) provided all-cause mortality, and linkage with the National Death Index (NDI) allowed evaluation of the specific causes of death, including neoplasm and infection.[16,17] Although DAF all-cause mortality was available until August 1, 2023, NDI cause-specific mortality was only available on deaths occurring before December 31, 2020. We followed patients from first ICI treatment to date of death or end of study period for all-cause mortality and censored data at December 31, 2020, for patients being evaluated for cause-specific mortality.

Statistical analysis

Descriptive statistics were calculated for baseline demographic characteristics at the date of first ICI infusion. Differences between groups with categorical data were compared with Fisher exact test, and continuous data were compared with *t* test for normally distributed data and Wilcoxon rank-sum test for skewed data with the *P* value reported. If data were missing or not documented, it was classified as unknown. Unknown was treated as a categorical variable in the model. Crude mortality rates were calculated for both RA and non-RA groups. Cox proportional hazards regression models assessed mortality risk in patients with RA compared with patients without RA and censored at the end of the study period on August 1, 2023. Models were adjusted for age, sex, race, smoking history, cancer type, CAN score 90-day mortality probability, and total glucocorticoid exposure during the 90 days before first ICI treatment.

For the cause-specific mortality analysis, data was limited to the 192 patients with RA and 1353 controls without RA who had initiation of ICI therapy before February 31, 2020. All data in the cause-specific mortality data were censored at December 31, 2020. All analyses were conducted using Stata v18.1 (StataCorp, College Station, TX).

RESULTS

Patient characteristics

Among 73,677 Veterans with RA there were 301 (0.41%) treated with ICI. Among the 727,627 patients without RA there were 2,114 (0.29%) treated with an ICI (Supplementary Figure 1).

The mean age was 71.2 years, a majority reported white race (74.0%), non-Hispanic ethnicity (93.9%), were male (93%), and current or former smokers (53.8%) (Table 1). Smoking status was unknown for 31.8% and 33.8% of RA cases and controls, respectively. No statistically significant differences in demographics or smoking status were identified. The absolute comorbidity CAN scores for 1 year and 90 days before the first ICI infusion and 1 year and 90-day mortality prediction scores showed higher scores in patients with RA in comparison with controls without RA, which were statistically significant for the 90-day scores (Table 1).

Healthcare use before the first ICI infusion showed similar hospitalization during the year before the first ICI infusion, with 167 (55.5%) patients with RA and 1160 (54.9%) of controls without RA having at least one hospitalization with similar numbers of hospitalization in the two groups (Table 1). The overall outpatient

Table 1. Demographic, laboratory, comorbidity, and health care use data at immune checkpoint inhibitor initiation*

Patient characteristics	RA (N = 301)	RA-control (N = 2114)	*P* value
Age at ICI start, mean (SD)	71.4 (6.5)	71.1 (7.1)	0.53
Male, n (%)	286 (95.0)	1956 (92.5)	0.15
Race, n (%)			
White	239 (79.4)	1549 (73.2)	0.08
Black	47 (15.6)	434 (20.5)	
Other/unknown	15 (5.0)	131 (6.2)	
Ethnicity, n (%)			
Not Hispanic	279 (92.7)	1990 (94.1)	0.24
Hispanic	15 (5.0)	66 (3.1)	
Other/unknown	7 (2.3)	58 (2.7)	
Smoking status, n (%)			
Current/former	160 (53.1)	1140 (53.9)	0.38
Never	45 (15.0)	258 (12.2)	
Unknown	96 (31.9)	716 (33.9)	
Rheumatoid factor status, n (%)			
Positive	191 (63.4)	61 (2.8)	<0.0001
Negative	95 (31.6)	412 (19.4)	
Not tested	15 (5.0)	1641 (77.6)	
anti-CCP status, n (%)			
Positive	150 (49.8)	8 (0.3)	<0.0001
Negative	106 (35.2)	143 (6.7)	
Not tested	45 (15.0)	1963 (92.8)	
Comorbidity assessment, mean (SD)			
CAN comorbidity score			
One-year score	89.5 (10.8)	87.9 12.5)	0.534
90-d score	89.5 (11.1)	87.7 (13.0)	0.535
Mortality probability at 1 year	0.18 (0.14)	0.17 (0.14)	0.535
Mortality probability at 90 days	0.06 (0.06)	0.05 (0.06)	0.537
Healthcare utilization in year before initial ICI infusion			
Hospitalizations during the year before first ICI infusion			
Patients with ≥1 hospitalizations, n (%)	167 (55.5)	1160 (54.9)	0.85
Average number of admissions for all patients, mean (SD)	1.16 (1.64)	1.17 (1.59)	0.50
Average number of admissions for patients with ≥1 admissions, mean (SD)	2.08 (1.72)	2.13 (1.60)	0.22
Outpatient clinic visits during year before first ICI infusion			
All specialty visits			
Patients with ≥1 visit, n (%)	300 (99.7)	2102 (99.4)	1.00
Average number of visits, mean (SD)	18.0 (11.5)	14.3 (10.4)	0.0002
Average number of visits for patients with ≥1 visits, mean (SD)	18.0 (11.5)	14.3 (10.4)	0.0002
Hematology/oncology clinic visits			
Patients with ≥1 visit, n (%)	294 (97.7)	2058 (97.4)	1.00
Average number of visits, mean (SD)	8.2 (7.3)	7.4 (7.1)	0.07
Average number of visits for patients with ≥1 visits, mean (SD)	8.4 (7.9)	7.6 (7.1)	0.07
Rheumatology clinic visits			
Patients with ≥1 visit, n (%)	210 (69.8)	46 (2.2)	<0.0001
Average number of visits for patients, mean (SD)	2.5 (2.6)	0.1 (0.4)	<0.0001
Average number of visits for patients with ≥1 visits, mean (SD)	3.6 (2.4)	2.3 (1.7)	0.0007

* anti-CCP, anti-cyclic citrullinated peptide; CAN, Care Assessment Need; ICI, immune checkpoint inhibitor; RA, rheumatoid arthritis.

clinical services use appeared slightly higher in patients with RA, with an average of 18.0 ± 11.5 visits per year for patients with RA in comparison with 14.3 ± 10.4 visits per year for patients without RA, with this difference statistically significant (P = 0.002). Hematology/oncology clinic visits were very similar for patients with RA at 8.2 ± 7.3 per year compared with patients without RA, with 7.4 ± 11.5 visits per year (P = 0.07). Rheumatology clinic visits were more common for patients with RA at 2.5 ± 2.6 per year, in comparison with 0.1 ± 0.4 per year for patients without RA (<0.0001) (Table 1).

Cancer diagnosis and stage

Lung cancer was the most common malignancy type—NSCLC (40.7%) and other lung cancer (10.5%)—followed by genitourinary (13.1%), gastrointestinal (9.9%), melanoma (8.5%), and head and neck (5.7%). Stage IV (35.6%) cancer was most common, followed by Stage III (22.4%) and Stage II (13.1%). No stage could be determined in 698 (28.9%) of patients. There were no clinical or statistically significant differences in the underlying cancer diagnosis (P = 0.10) or stage (P = 0.60) between RA and non-RA groups (Table 2).

ICI treatment

Initial ICI therapy was monotherapy in 294 (97.7%) of patients with RA and 2051 (97.0%) of patients without RA (Table 2). PD-1 inhibitors were the most used agents, with pembrolizumab being the most common first ICI in both groups, followed by nivolumab. All patients on initial combination ICI therapy were treated with ipilimumab and nivolumab. There were 6 (1.9%) patients with RA and 38 (1.8%) patients without RA who had subsequent "add on" combination therapy, with all these cases of combination therapy being ipilimumab and nivolumab except one patient without RA who had "add on" therapy with ipilimumab and pembrolizumab. The total number of patients with combination at some point during their ICI therapy was 13 (4.3%) of patients with RA and 102 (4.8%) of patients without RA (Table 2).

Cancer treatment before ICI infusions

Complete information on the specific cancer treatments received before ICI treatment is included in Supplemental Table 1. For 74 (24.6%) patients with RA and 542 (25.6%) patients without RA, ICI treatment was the first cancer therapy received. Of the 227 (75.6%) patients with RA and 1,572 (74.4%) patients without RA who previously received chemotherapy, the average number of drugs received was 1.7 ± 0.9 (range 1–6) and 1.8 ± 1.1 (range 1–11), respectively, which was not statistically significantly different (P = 0.98) (Table 2). A complete list of all non-ICI cancer treatments received is included in Supplementary Table 1.

DMARD therapy

All 301 (100%) patients with RA and 76 (3.6%) patients without RA received at least one DMARD before ICI therapy. Looking specifically at the 90 days before ICI treatment, there were 138 (46.2%) of patients with RA with DMARD treatment within 90 days and 11 (0.5%) of patients without RA receiving one or more DMARDs during the same period (Table 3). In patients with RA, the average number of DMARDs was 0.7 ± 0.9 (range 0–4). Methotrexate and rituximab may have been used as chemotherapy. Individual indications for DMARD use were not confirmed except for individual chart reviews of the three patients without RA receiving DMARDs other than methotrexate or rituximab within 90 days of ICI initiation. This review identified one patient each taking etanercept for psoriasis, azathioprine for celiac disease, and sulfasalazine for a seronegative inflammatory arthritis, which did not meet RA criteria.

Glucocorticoid therapy

Any glucocorticoid therapy before ICI infusion was observed in 284 (94.4%) of patients with RA and 1614 (76.3%) of patients without RA, with the difference being statistically significant (P < 0.0001) (Table 3). Higher glucocorticoid use was also seen in patients with RA in comparison with patients without RA for any prednisone any time before ICI and use of any glucocorticoid and prednisone during the 90-day periods before ICI, which was statistically significant (P < 0.0001) (Table 3). Dexamethasone was used in 65 (21.6%) of patients with RA and 528 (25.0%) of patients without RA in the 90-day period before ICI, which was not statistically significant (P = 0.22) (Table 3). Total cumulative glucocorticoid use during the 90-day period before ICI was 282 ± 627 mg for patients with RA in comparison with 229 ± 644 mg for non-RA controls, which was not statistically significant (P = 0.22) (Table 3). Total cumulative dose of all prednisone and dexamethasone showed a trend for high use of prednisone in patients with RA and higher use of dexamethasone in patients without RA, although these differences were not statistically significant different (Table 3).

Survival

Survival was poor in both groups, with a one-year survival of 45.8% (95% confidence interval [CI] 40.0–51.6) in patients with RA vs 51.1% (95% CI 48.9–53.3) in patients without RA (P = 0.09), and two-year survival 28.7% (95% CI 23.4–34.3) in patients with RA vs 31.3% (95% CI 29.2–33.4) in patients without RA (P = 0.41). There was no significant survival difference between the groups on the Kaplan-Meier survival (log-rank

Table 2. Cancer diagnosis, cancer stage, and cancer treatments*

	RA (N = 301)	Non-RA (N = 2114)	*P* value
Cancer diagnosis, n (%)			
Gastrointestinal	22 (7.3)	218 (10.3)	0.10
Genitourinary	41 (13.6)	275 (13)	
Head and neck	14 (4.7)	124 (5.9)	
Lung—NSCLC	129 (42.9)	854 (40.4)	
Lung—other	28 (9.3)	226 (10.7)	
Melanoma	37 (12.3)	168 (7.9)	
Other	30 (10)	249 (11.8)	
Cancer stage, n (%)			
II	43 (14.3)	274 (13.0)	0.60
III	71 (23.6)	469 (22.2)	
IV	97 (32.2)	763 (36.1)	
Unknown	90 (29.9)	608 (28.8)	
ICI treatment			
Initial ICI in patient initiated on ICI monotherapy, n (%)			
PD-1 inhibitors			
Cemiplimab	5 (1.7)	15 (0.7)	0.33
Nivolumab	80 (26.6)	610 (28.9)	
Pembrolizumab	133 (44.2)	928 (43.9)	
PD-L1 inhibitors			
Atezolizumab	38 (12.6)	231 (10.9)	
Avelumab	2 (0.7)	5 (0.2)	
Durvalumab	32 (10.6)	225 (10.6)	
CTLA-inhibitor			
Ipilimumab	4 (1.3)	37 (1.8)	
Total patients with initial monotherapy	294 (97.7)	2051 (97.0)	
Combination ICI therapy			
Initial ICI combination therapy			
Ipilimumab and Nivolumab	7 (2.3)	63 (2.9)	
Add on ICI combinations therapy			
Ipilimumab and Nivolumab	6 (1.9)	38 (1.7)	
Ipilimumab and Pembrolizumab	0	1 (0)	
Total patients with combination therapy	13 (4.3)	102 (4.8)	
Number of unique ICIs per patient, n (%)			
1	274 (91.0)	1930 (91.3)	0.22
2	24 (8.0)	177 (8.4)	
3	3 (1.0)	7 (0.4)	
Additional information on ICIs, mean (SD)			
Unique drugs per patient	1.1 (0.3)	1.1 (0.3)	0.85
Number of ICI infusion per patient	8.7 (12.5)	9.1 (11.7)	0.57
Mean ICI Rx duration, days[a]	256 (394)	263 (393)	0.49
Cancer medical therapies before first ICI infusion			
Number of cancer drugs per patient before ICI, n (%)			
0	74 (24.6)	542 (25.6)	0.98
1	121 (40.2)	848 (40.1)	
2	67 (22.3)	439 (20.8)	
3	26 (8.6)	181 (8.6)	
4	10 (3.3)	54 (2.6)	
5 or more	3 (1.0)	50 (2.4)	
Number of non-ICI cancer drugs per patient before ICI, mean (SD)			
All patients	1.3 (1.1)	1.3 (1.2)	0.77
Patients with one or more non-ICI cancer drugs[b]	1.7 (0.9)	1.8 (1.1)	0.98

* ICI, immune checkpoint inhibitor; NSCLC, non-small cell lung cancer; PD, programmed cell death protein; RA, rheumatoid arthritis; Rx, prescription.
[a] Days between first infusion date and last infusion date.
[b] Includes only patients receiving ≥1 non-ICI cancer drugs before first ICI infusions (RA = 227, non-RA = 1,572).

P = 0.267) (Figure 1). Cox proportional hazards regression modeling showed no difference in mortality risk between the groups in both the unadjusted analysis and when adjusted for age, sex, race, smoking history, cancer type, CAN score for 90-day mortality rate, or cumulative glucocorticoid exposure within 90 days before ICI treatment. For the full cohort, the crude and adjusted hazard ratios were 1.08 (95% CI 0.94–1.25), *P* = 0.267 and 1.09 (95% CI 0.94–1.25), and *P* = 0.256, respectively (Table 4).

The subgroup analysis of patients with NSCLC, the largest cancer subgroup, was conducted using the same methodology

Table 3. DMARD and glucocorticoid therapy before immune checkpoint inhibitor treatment*

	RA (N = 301)	Non-RA (N = 2114)	*P* value
DMARD therapy, n (%)			
DMARDs within 90 days of first ICI infusion			
Abatacept	3 (1)	0 (0)	<0.0001
Adalimumab	6 (2)	0 (0)	
Azathioprine	4 (1.3)	1 (0)	
Etanercept	10 (3.3)	1 (0)	
Hydroxychloroquine	75 (24.9)	0 (0)	
Infliximab	1 (0.3)	0 (0)	
Leflunomide	13 (4.3)	0 (0)	
Methotrexate IV	1 (0.3)	2 (0.1)	
Methotrexate oral	54 (17.9)	2 (0.1)	
Rituximab	4 (1.3)	4 (0.2)	
Sulfasalazine	38 (12.6)	1 (0)	
Tocilizumab	1 (0.3)	0 (0)	
Tofacitinib	3 (1)	0 (0)	
Number of DMARDs within 90 days of first ICI infusion, n (%)			
0	162 (53.8)	2103 (99.5)	<0.0001
1	75 (24.9)	11 (0.5)	
2	55 (18.3)	0 (0)	
3	8 (2.7)	0 (0)	
4	1 (0.3)	0 (0)	
Mean number of DMARDs within 90 days, mean ± SD	0.7 (0.9)	0.01 (0.07)	<0.0001
Glucocorticoid therapy before first ICI infusion			
Any glucocorticoids before ICI, n (%)	284 (94.4)	1613 (76.3)	<0.0001
Any prednisone before ICI, n (%)	261 (86.7)	992 (46.89)	<0.0001
Any glucocorticoids during 90 days before ICI, n (%)	143 (47.5)	715 (33.8)	<0.0001
Any prednisone during 90 days before ICI, n (%)	85 (28.2)	225 (10.6)	<0.0001
Either IV or oral dexamethasone during 90 days before ICI, n (%)	65 (21.6)	528 (25.0)	0.22
Total glucocorticoid during 90 days before ICI, mean ± SD[a]	288 (627)	229 (644)	0.14
Total dexamethasone during 90 days before ICI, mean ± SD	122 (528)	176 (576)	0.13
Total prednisone during 90 days before ICI, mean ± SD	158 (342)	46 (252)	<0.0001

* DMARD, disease modifying antirheumatic drug; ICI, immune checkpoint inhibitor; IV, intravenous; RA, rheumatoid arthritis.
[a] All doses reported a prednisone mg equivalence.

that was used for the full population (Table 4, Figure 1). The NSCLC RA and NSCLC non-RA groups were very similar in patient characteristics, comorbidities, and other parameters (Supplementary Table 2). There was a one-year survival of 48.4% (95% CI 39.6–57.3) compared with 55.2% (95% CI 51.7–58.6) and two-year survival of 31.0% (95% CI 22.5–39.6) compared with 34.9% (95% CI 31.4–38.3) in the RA and non-RA groups, respectively. There was no significant survival difference between the groups (log-rank $P = 0.514$) (Figure 1). For the NSCLC subgroup, the crude and adjusted hazard ratios were 1.07 (95% CI 0.87–1.33) ($P = 0.515$) and 1.02 (95% CI 0.82–1.28), ($P = 0.830$), respectively (Table 4).

Cause-specific mortality

For the cause-specific analysis, only patients with first ICI infusion before December 31, 2020, were included, and data were censored after December 31, 2020. Of the 192 patients with RA in this subgroup, 114 (59.4%) patients died before December 31, 2020, and of 1352 of the patients without RA in this subgroup, 837 (61.9%) died before December 31, 2020. The cause of death was similar in both groups (χ^2 0.80, $P = 0.77$), with the most common cause of death being neoplasm in 93.0% of patients with RA and 90.9% of patients without RA (Table 5). Deaths due to infection were rare in both groups (<1.0%). For this subgroup with NDI data, the one-year survival of 47.4% (95% CI 40.2–54.5) vs 52.2% (95% CI 49.6–54.9) ($P = 0.21$) and two-year survival of 33.8% (95% CI 27.1–40.6) vs 36.4% (95% CI 33.8–39.0) ($P = 0.40$) was like that seen in the full cohort. For this subgroup, the crude and adjusted hazard ratios were 1.03 (95% CI 0.85–1.26) ($P = 0.754$) and 1.00 (95% CI 0.81–1.22) ($P = 0.963$), respectively (Table 4).

DISCUSSION

In this retrospective cohort study of Veterans, patients with pre-existing RA who received ICIs for cancer did not experience increased mortality or differences in cause of death compared with patients without RA receiving ICI treatment. These findings significantly expand on the limited existing literature reporting similar risk of mortality in RA and matched non-RA controls.[18] Our subgroup analysis of NSCLC was also consistent with a recent

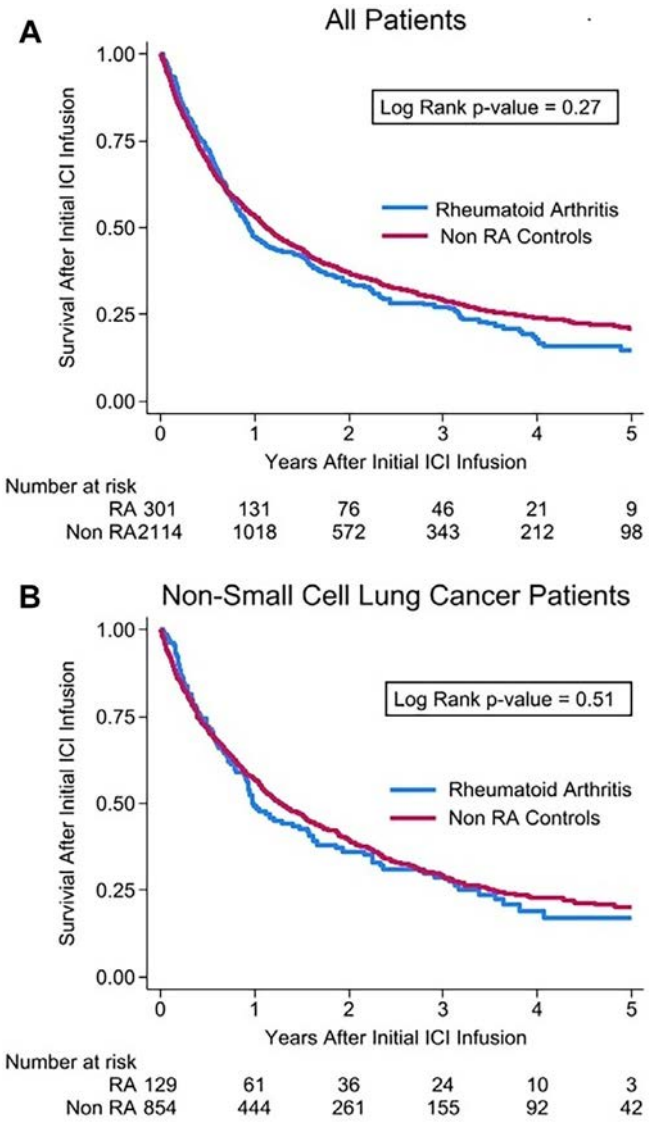


Figure 1. Kaplan-Meier survival analysis after first ICI infusion. There was no significant difference in mortality rates between groups over time up to five years after the first ICI infusion for (A) the full patient group and (B) a subgroup of patients with NSCLC. ICI, immune checkpoint inhibitor; NSCLC, non-small cell lung cancer.

publication showing no difference of survival after ICI therapy in patients with RA and patients without RA in the Medicare population.[19] There are no studies comparing the survival of patients with RA receiving ICI therapy to survival with traditional chemotherapy programs.

Previous studies of patients with various autoimmune diseases assessed ICI-related outcomes, including the risk of flare of the pre-existing autoimmune disease after ICI exposure, the risk of developing a new autoimmune disease after ICI exposure, and the effectiveness of ICIs for cancer treatment. In particular, several studies reported that patients with a pre-existing autoimmune disease treated with an ICI experience flares of their disease, but most of these flares were mild and often effectively managed with standard doses of steroids.[3,20–24] The literature to date has been conflicting regarding the risk for developing a new autoimmune disease after ICI exposure. For example, Cortellini et al found a significantly increased risk of irAEs in patients receiving ICI with an underlying autoimmune disease, but a large meta-analysis found no association between the presence of baseline autoimmune disease and development of irAEs, severity of irAEs, or overall response rate.[25,26] Regarding effectiveness of ICIs for cancer treatment, a few smaller studies have shown similar effectiveness of ICIs for cancer treatment in terms of overall survival and response rates in patients with and without pre-existing autoimmune disease. These studies included small samples of patients with RA among other autoimmune diseases.[19,27]

Less is known about patients with RA, specifically those who are receiving ICIs because of their exclusion from early clinical trials. A major limitation of existing literature is the heterogeneity of autoimmune disease included in these retrospective studies. Most of the literature on use of ICIs in patients with RA is from observational studies examining differences in treatment efficacy and rates of adverse effects in patients with a variety of autoimmune diseases, with 16% of patients represented in a large meta-analysis carrying a pre-existing diagnosis of RA.[4,28] Efuni et al reported incidence of disease flares and other irAEs in a subgroup of 22 patients with RA treated with ICIs but did not examine mortality and did not include a comparator group.[3] A larger study of 100 patients with RA receiving ICIs found that almost half experienced flares within three months, but disease flares were not associated with differences in overall mortality.[9] A smaller study that included 87 patients with RA undergoing ICI therapy, found that patients with RA had more frequent irAEs overall, but no difference in the risk of mortality or severe irAEs compared with matched controls without RA.[18]

Strengths of this work include examining all-cause and cause-specific mortality in the largest cohort of patients with RA on ICI therapy to date. The diagnosis of RA was determined using a validated algorithm with a positive predictive value of over 90%. Clinical and pharmacy information were taken from a standard electronic medical record system across the VHA to provide nationwide uniform data collection. For mortality data, we were able to confirm mortality rates from national administrative data and specific cause of death that is available through linkages with the NDI.

There are also limitations to our work. We recognize that the number of covariates that can be included in our model may be limited by our sample size. Separate analyses were undertaken to determine if there were significant changes in the model by including cancer stage, previous cancer therapy, number and duration of ICI infusions, or changing the glucocorticoid covariate to a cumulative dose for either all glucocorticoids or prednisone alone. These modifications in the model did not produce substantial differences from those reported with the reported covariate analysis.

Table 4. Cox proportional hazard ratio comparison of all-cause mortality in patients with RA with controls without RA*

	Hazard ratio	95% CI	*P* value
Unadjusted			
All patients[a]	1.08	0.94–1.25	0.267
Patients with non-small cell lung cancer[b]	1.07	0.87–1.33	0.515
Patients with NDI data[c]	1.03	0.85–1.26	0.754
Adjusted[d]			
All patients[a]	1.09	0.94–1.25	0.256
Patients with non-small cell lung cancer[b]	1.02	0.82–1.28	0.830
Patients with NDI data[c]	1.00	0.81–1.22	0.963

* CAN, Care Assessment Needs; CI, confidence interval; ICI, immune checkpoint inhibitor; NDI, National Death Index; NSCLC, non-small cell lung cancer; RA rheumatoid arthritis.
[a] 301 patients with RA; 2,114 control patients without RA.
[b] 129 patients with RA; 854 control patients without RA with NSCLC.
[c] 192 patients with RA; 1,353 control patients without RA, data censored at December 31, 2020.
[d] Adjusted for age, sex, race, smoking history, cancer type, CAN score 90-day mortality probability, and total glucocorticoid exposure during 90 days before ICI, except not cancer type adjustment on the NSCLC sub-analysis.

The potential for channeling bias, with patients with RA being selectively channeled away from ICI therapy because of concerns for irAEs or other issues, is a question that cannot fully be addressed by this work. However, the observation that patient characteristics, including cancer type, cancer stage, comorbidity, and health care use, were similar in the RA and non-RA groups, as well as the similarity in ICI drug selection and treatment patterns, is reassuring that channeling bias was not a major issue. Also, the similar ICI treatment patterns in the RA and non-RA groups in the subgroup analysis of NSCLC further suggests that channeling basis is not a major issue in the management of NSCLC in patients with RA.

Additionally, most patients in our study received ICI monotherapy. Given the increased risk of irAEs with ICI combination therapy, our data may not be as generalizable to the combination therapy population.[29] The lack of specific indication for the glucocorticoid prescriptions is a limitation of our study. Our sample included predominately elderly men, consistent with the VHA population, which may also limit the generalizability of our results, although as noted before, our results were very similar to those seen in the Medicare population, which suggests some generalizability. We did not assess the rates or severity of RA flares, the development of other de novo irAEs, or cancer response outcomes as part of this study. Finally, our data are retrospective, making this analysis subject to residual confounding, which could affect our results.

In summary, this investigation supports the hypothesis that patients with RA have similar clinical outcomes with ICI therapy to patients without RA. These results are consistent with similar reported observations for other autoimmune diseases and build on previous work in this area. These results provide critical information for rheumatology and oncology providers to consider when discussing potential risks and benefits of cancer treatment options in patients with underlying RA. Although these data suggest that a diagnosis of RA alone should not prevent consideration for ICI therapy when appropriately indicated, there remains a critical need for individualized risk-benefit assessment in decision making for cancer treatment recommendations in patients with RA. This individualized approach should include consideration of potential irAEs, cancer diagnosis and stage, the treatment options being considered, and the patient's individual characteristics.

Table 5. Specific cause of death from NDI following immune checkpoint inhibitor treatment*

	RA (N = 192)	Non-RA (N = 1,353)	*P* value
Living patients at the time of NDI report, n (%)	78 (40.6)	516 (38.1)	0.527
Deceased patients, n (%)	114 (59.4)	837 (61.9)	
Neoplasms, n (%)	106 (93.0)[a]	751 (90.9)	0.737
Respiratory system diseases, n (%)	3 (2.6)	16 (1.9)	
Circulatory system diseases, n (%)	2 (1.7)	34 (4.1)	
Infectious diseases, n (%)	1 (0.8)	8 (0.9)	
Other/unknown in NDI data, n (%)[b]	2 (1.7)	17 (2.0)	
Unknown—not included in NDI data, n (%)[c]	0 (0)	10 (0.7)	

* Analysis limited to patients with initial immune check point inhibitor infusion before the final date of National Death Index data was available December 31, 2020, with all data censored at this date. NDI, National Death Index; RA, rheumatoid arthritis; VA, Veterans Affairs.
[a] Percent reported as percentage of deaths, not total population at risk.
[b] NDI data report the cause of death as unknown or single event (eg, suicide, homicide, etc.).
[c] VA data documented patient death before December 31, 2020, but death was not captured by the NDI.

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 Cannon 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. Martins F, Sofiya L, Sykiotis GP, et al. Adverse effects of immune-checkpoint inhibitors: epidemiology, management and surveillance. Nat Rev Clin Oncol 2019;16(9):563–580.
2. Cappelli LC, Shah AA, Bingham CO III. Immune-related adverse effects of cancer immunotherapy- implications for rheumatology. Rheum Dis Clin North Am 2017;43(1):65–78.
3. Efuni E, Cytryn S, Boland P, et al. Risk of toxicity after initiating immune checkpoint inhibitor treatment in patients with rheumatoid arthritis. J Clin Rheumatol 2021;27(7):267–271.
4. Boland P, Pavlick AC, Weber J, et al. Immunotherapy to treat malignancy in patients with pre-existing autoimmunity. J Immunother Cancer 2020;8(1):8.
5. Li G, Shi F, Liu J, et al. The effect of CTLA-4 A49G polymorphism on rheumatoid arthritis risk: a meta-analysis. Diagn Pathol 2014;9(1):157.
6. Zhou C, Gao S, Yuan X, et al. Association between *CTLA-4* gene polymorphism and risk of rheumatoid arthritis: a meta-analysis. Aging (Albany NY) 2021;13(15):19397–19414.
7. Lee YH, Bae SC, Kim JH, et al. Meta-analysis of genetic polymorphisms in programmed cell death 1. Associations with rheumatoid arthritis, ankylosing spondylitis, and type 1 diabetes susceptibility. Z Rheumatol 2015;74(3):230–239.
8. Beydon M, Pinto S, De Rycke Y, et al. Risk of cancer for patients with rheumatoid arthritis versus general population: a national claims database cohort study. Lancet Reg Health Eur 2023;35:100768.
9. McCarter KR, Arabelovic S, Wang X, et al. Immunomodulator use, risk factors and management of flares, and mortality for patients with pre-existing rheumatoid arthritis after immune checkpoint inhibitors for cancer. Semin Arthritis Rheum 2024;64:152335.
10. Chung CP, Rohan P, Krishnaswami S, et al. A systematic review of validated methods for identifying patients with rheumatoid arthritis using administrative or claims data. Vaccine 2013;31(suppl 10): K41–K61.
11. United States Department of Veterans Affairs. Health systems research: corporate data warehouse. https://www.hsrd.research.va.gov/for_researchers/cdw.cfm
12. Fihn SD, Francis J, Clancy C, et al. Insights from advanced analytics at the veterans health administration. Health Aff (Millwood) 2014;33(7): 1203–1211.
13. Wang L, Porter B, Maynard C, et al. Predicting risk of hospitalization or death among patients receiving primary care in the Veterans Health Administration. Med Care 2013;51(4):368–373. https://doi.org/10.1097/MLR.0b013e31827da95a
14. Nelson KM, Chang ET, Zulman DM, et al. Using predictive analytics to guide patient care and research in a national health system. J Gen Intern Med 2019;34(8):1379–1380.
15. United States Department of Veterans Affairs. Veterans affairs central cancer registry. https://www.data.va.gov/dataset/Veterans-Affairs-Central-Cancer-Registry-VACCR-/jvmd-8fgj/about_data
16. US Centers for Disease Control and Prevention. National death index. https://www.cdc.gov/nchs/ndi/?CDC_AAref_Val=https://www.cdc.gov/nchs/ndi/index.htm
17. United States Department of Veterans Affairs death ascertainment file. https://www.hsrd.research.va.gov/for_researchers/cyber_seminars/archives/video_archive.cfm?SessionID=6303
18. McCarter KR, Wolfgang T, Arabelovic S, et al. Mortality and immune-related adverse events after immune checkpoint inhibitor initiation for cancer among patients with pre-existing rheumatoid arthritis: a retrospective, comparative, cohort study. Lancet Rheumatol 2023;5(5): e274–e283.
19. Jannat-Khah DP, Xie F, Saxena A, et al. Survival in immune checkpoint inhibitor–treated patients with rheumatoid arthritis and non–small cell lung cancer: an observational cohort study. Arthritis Care Res (Hoboken) 2025;77(10):1214–1221.
20. Gutzmer R, Koop A, Meier F, et al; German Dermatooncology Group (DeCOG). Programmed cell death protein-1 (PD-1) inhibitor therapy in patients with advanced melanoma and preexisting autoimmunity or ipilimumab-triggered autoimmunity. Eur J Cancer 2017;75:24–32.
21. Johnson DB, Sullivan RJ, Ott PA, et al. Ipilimumab therapy in patients with advanced melanoma and preexisting autoimmune disorders. JAMA Oncol 2016;2(2):234–240.
22. Mitchell EL, Lau PKH, Khoo C, et al. Rheumatic immune-related adverse events secondary to anti-programmed death-1 antibodies and preliminary analysis on the impact of corticosteroids on anti-tumour response: a case series. Eur J Cancer 2018;105:88–102.
23. Alexander S, Swami U, Kaur A, et al. Safety of immune checkpoint inhibitors in patients with cancer and pre-existing autoimmune disease. Ann Transl Med 2021;9(12):1033.
24. Tison A, Garaud S, Chiche L, et al. Immune-checkpoint inhibitor use in patients with cancer and pre-existing autoimmune diseases. Nat Rev Rheumatol 2022;18(11):641–656.
25. Yeung C, Kartolo A, Holstead R, et al. Safety and clinical outcomes of immune checkpoint inhibitors in patients with cancer and preexisting autoimmune diseases. J Immunother 2021;44(9):362–370.
26. Cortellini A, Buti S, Santini D, et al. Clinical outcomes of patients with advanced cancer and pre-existing autoimmune diseases treated with anti-programmed death-1 immunotherapy: a real-world transverse study. Oncologist 2019;24(6):e327–e337.
27. Tang H, Zhou J, Bai C. The efficacy and safety of immune checkpoint inhibitors in patients with cancer and preexisting autoimmune disease. Front Oncol 2021;11:625872.
28. Abdel-Wahab N, Shah M, Lopez-Olivo MA, et al. Use of immune checkpoint inhibitors in the treatment of patients with cancer and pre-existing autoimmune disease: a systematic review. Ann Intern Med 2018;168(2):121–130.
29. Wolchok JD, Chiarion-Sileni V, Gonzalez R, et al. Overall survival with combined nivolumab and ipilimumab in advanced melanoma. N Engl J Med 2017;377(14):1345–1356.

Arthritis Care & Research
Vol. 78, No. 6, June 2026, pp 713–722
DOI 10.1002/acr.25685

American College of Rheumatology
Empowering Rheumatology Professionals

Objectively Measured Sleep Disorders in Rheumatoid Arthritis and Their Association With Disease Activity

Patricia Katz,[1] Sarah L. Patterson,[1] Mary C. Nakamura,[2] Aric A. Prather,[1] Joonsuk Park,[1] and Katie L. Stone[1,3]

Objective. Poor self-reported sleep is common in rheumatoid arthritis (RA), but studies using objective sleep measures are rare. We report the prevalence of objectively measured sleep characteristics in individuals with RA and their association with clinical measures of RA disease activity.

Methods. Data were from a longitudinal study with four measurement periods at 6-month intervals. Sleep data were collected by actigraphy over 7 days at each period. Primary actigraphy sleep variables were sleep efficiency (SE; time asleep/time in bed), time awake after sleep onset (WASO), and total sleep time. At baseline, WatchPAT devices were also used for two nights to assess presence of sleep-disordered breathing (SDB; primarily obstructive sleep apnea (OSA) indicated by the apnea-hypopnea index [AHI]). Disease activity measures (Clinical Disease Activity Index [CDAI], Disease Activity Score [DAS] 28 joints, DAS C-reactive protein [CRP], and DAS erythrocyte sedimentation rate) were completed by rheumatologists before each sleep-monitoring period. Prevalence of sleep problems was estimated, and associations of sleep characteristics with RA disease activity were analyzed with regression models.

Results. Of 133 individuals enrolled, 116 had sufficient actigraph wear time for scoring at baseline, and 63 completed WatchPAT monitoring. More than 40% of the cohort had poor SE (<85%) at all measurement periods completed. More than half with WatchPAT assessments had moderate to severe OSA (AHI ≥ 15). Lower SE and higher WASO were associated with greater disease activity across all measures. AHI was associated with CDAI and DAS-CRP.

Conclusion. Objectively measured sleep problems were frequent and associated with RA disease activity. Given high rates of sleep disorders and their significant negative health effects, greater attention to sleep disorders among individuals with RA is warranted.

INTRODUCTION

Several studies have shown the high frequency of self-reported poor sleep among people with rheumatoid arthritis (RA), with the prevalence as high as 67%.[1–3] The gold standard for identifying sleep disorders is polysomnography (PSG). The few sleep studies using PSG among people with RA have been relatively small but have shown high rates of poor sleep, primarily sleep fragmentation.[4,5] Sleep-disordered breathing (SDB), specifically obstructive sleep apnea (OSA), also appears to be disproportionately present among individuals with RA based on data from clinical sleep medicine studies as well as data reported from screening measures.[6–9] It is worth noting that estimates of the prevalence of sleep disorders based on these PSG studies may be biased upward because they are based on referrals for clinical evaluations.

More recently, data have begun to accrue from objective home-based measurements of sleep such as actigraphy.[10–12] These studies have shown evidence of poor sleep quality based on sleep efficiency (time sleeping / time in bed) and associations with symptoms such as pain and fatigue. Prior studies using both objective and self-reported measures of sleep or sleep disruption have also shown links between sleep disturbance and lowered pain thresholds, higher levels of pain and fatigue, and depressive symptoms among people with RA.[4,8,11,13,14]

Supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases, NIH (grant R01-AR-072040).

[1]University of California San Francisco; [2]San Francisco Veterans Affairs Health Care System, San Francisco, California; [3]California Pacific Medical Research Institute, San Francisco.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25685).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25685.

Address correspondence via email to Patricia Katz, PhD, at patti.katz@ucsf.edu.

Submitted for publication April 25, 2025; accepted in revised form September 23, 2025.

SIGNIFICANCE & INNOVATIONS

- Most of this sample had at least one sleep disorder. More than half had poor sleep efficiency and short sleep times, and more than half had sleep-disordered breathing (obstructive sleep apnea [OSA]) of at least moderate severity.
- Worse sleep efficiency, greater time awake after sleep onset, and presence of OSA were each significantly associated with clinical rheumatoid arthritis (RA) disease activity measures.
- Given the high rates of sleep disorders and their significant negative health effects, greater attention to sleep disorders among individuals with RA is warranted. Screening for sleep-disordered breathing disorders such as OSA may be particularly critical.

Although there have been some reports of the association between self-reported sleep disturbance and RA disease activity,[15,16] prior studies have not examined the association of objective measurement of sleep with clinical assessments of RA disease activity. There are biologic reasons to propose that such a link exists. In addition to the impact of poor sleep on pain and pain thresholds, which may affect components of RA disease activity measures, cytokines important to RA disease activity, such as interleukin-6 and tumor necrosis factor α, also play a role in regulating sleep[17,18] Here, we report the prevalence of objectively measured sleep characteristics including the presence of OSA among a cohort of individuals with RA, the overlap of sleep characteristics, and the association with clinical measures of RA disease activity.

METHODS

Participants and study design

Study data were from a longitudinal study of sleep disorders in RA. Participants were recruited from previous studies as well as from rheumatology clinics at an academic medical center and a safety-net hospital. Inclusion criteria were age ≥18 years, physician diagnosis of RA, fluency in English or Spanish, not currently pregnant, and not currently being treated for OSA.

The original study design consisted of four data-collection periods at 6-month intervals. At the beginning of each study period, participants were to complete an in-person visit that included a physician examination, blood draw, and orientation to sleep-measurement devices and protocols, followed by up to 9 nights of at-home sleep monitoring using two devices. In March 2020, all research activities were stopped because of the COVID-19 pandemic. In June 2020, we were able to resume the study remotely; however, in-person visits were not reinstated. The study was approved by the University of California San Francisco institutional review board (approval number 17-21790), and all participants provided informed consent.

Sleep measurement

At the initial study period, participants underwent 9 nights of home sleep monitoring, including 2 consecutive nights with a WatchPAT device (Itamar Medical, Caesarea, Israel) to assess SDB, followed by 7 consecutive nights with an actigraph to assess additional 24-hour sleep-wake data. Participants also completed sleep diaries that included information about time going to bed, time arising, and actigraph nonwear time to aid in scoring. Follow-up actigraph assessments were made at 6, 12, and 18 months after baseline, following the same protocol.

WatchPAT. WatchPAT is a validated diagnostic wrist-worn device that measures multiple channels of data (peripheral arterial tone, oximetry, actigraphy, heart rate, and chest motion) with three points of contact (wrist, finger, and chest sensor). Data from the WatchPAT were transferred to the proprietary zzzPAT software for analysis. The WatchPAT yields estimates on a variety of measures, including the apnea-hypopnea index (AHI), reflecting the number of apnea or hypopnea episodes per hour of sleep. AHI was averaged over the 2 nights and was categorized as normal (<5), mild OSA (5 to <15), moderate OSA (15 to <30), or severe OSA (≥30). WatchPAT also provided the oxygen desaturation index (ODI), reflecting the number of average times per hour that blood oxygen level decreases. WatchPAT has been validated against PSG for identification of SDB and correspondence of AHI determinations.[19,20]

Actigraph. Participants were provided with an actigraph (GT9X, Actigraph Corp., Pensacola, FL, USA), a watch-like device that has shown good correspondence with PSG in detecting sleep/wake patterns,[21] to wear for the remaining 7 nights of the baseline period and for the follow-up periods. Participants were instructed to wear the actigraph on their nondominant wrist for 7 consecutive days and remove only when showering. The GT9X captures objectively measured sleep through wrist movement activity. The device has a validated three-axis accelerometer with a sample rate of 80 Hz, an integrated gyroscope, and a magnetometer. Data were processed using 1-minute epochs (activity counts/minute) and scored using ActiGraph, LLC ActiLife software (version 6.13.4). Participants' bedtimes were cross-checked using daily diaries. Nonwear time was determined by a combination of an off-wrist detector in the device, a nonwear algorithm, and a review by an actigraphy data scorer.[22,23] Nonwear times were set to missing. Based on the recommended duration of 72 hours to 14 consecutive days for assessment of sleep disorders with actigraphy,[24] a minimum of 4 valid days of wear time was required to be included in the analysis. The Cole-Kripke sleep scoring algorithm was used to determine sleep from wake.[25]

Measures of interest from the actigraph were mean sleep efficiency (defined as total sleep time / time in bed; includes time

before falling asleep, ie, sleep onset latency), time awake after sleep onset (WASO), and mean total sleep time (during night-time while in bed). All values were averaged over all nights during the wear period.

As noted above, in March 2020 all research activities were stopped because of the COVID-19 pandemic. Beginning in June 2020, we were able to resume data collection with the actigraphs by sending the devices out by mail. However, because in-person training on use of the WatchPAT was required, we did not resume use of these devices.

Clinical measures of RA disease activity

During the in-person visits immediately before each sleep-measurement period, one of two board-certified rheumatologists conducted an examination to assess the tender joint count and swollen joint count and provide a global assessment of disease activity. Blood was drawn for the measurement of erythrocyte sedimentation rate (ESR) and C-reactive protein (CRP) by the clinical laboratory. The Clinical Disease Activity Index (CDAI)[26] and Disease Activity Score (DAS) with 28 joints with both ESR and CRP[27,28] were calculated. The CDAI is calculated from 28-joint swollen and tender joint counts and patient and physician global assessments. The DAS-ESR and DAS-CRP are calculated from 28-joint swollen and tender joint counts, a patient global assessment, and either ESR or CRP, respectively.

Other data collection

Participants completed structured telephone interviews immediately before each sleep-measurement period. The interview covered sociodemographic information, general health factors, medication use, and screening for depression and anxiety. Sociodemographic characteristics included age, sex, race, ethnicity, education, and income. General health factors were body mass index (BMI), smoking, comorbid conditions (hypertension, heart disease, chronic obstructive pulmonary disease, and diabetes), and RA duration. Medications included glucocorticoid (GC) use and current dose, conventional disease-modifying antirheumatic drugs (eg, methotrexate), tumor necrosis factor inhibitors (eg, adalimumab), abatacept, tocilizumab, and rituximab. Use of prescription pain medications, specifically medications containing codeine, hydrocodone, oxycodone, or fentanyl was ascertained; indications for pain medication were not obtained.

Analysis

Sleep characteristics from both the WatchPAT and actigraph at each time period were tabulated. Poor sleep efficiency was defined as <85%[29]; sleep efficiency <80% was also examined in a sensitivity analysis. Because of the known nonlinear association between sleep time and health outcomes,[30] we examined two three-level sleep categorizations. The first defined short sleep as an average of <7 hours,[31] and the second defined short sleep as <6 hours.[32,33] An average of >8 hours was defined as long sleep, with the 7- to 8-hour or 6- to 8-hour groups, respectively, used as the reference group in analyses. Individual results for total sleep time, sleep efficiency, and WASO were also compared with published sex and age-group reference values taken from second night PSG of healthy adults without known sleep disorders.[34] These reference values are shown in Supplemental Table S1.

Overlap of sleep disorders at the baseline assessment was examined. The cross-sectional association of actigraph-assessed sleep characteristics with RA disease activity measures at each time period was assessed with mixed-effects linear regression to permit inclusion of data from all observation periods. Both unadjusted analyses and analyses adjusting for age, sex, BMI, number of comorbid conditions, and RA duration were conducted. Because a relatively small number of participants had both WatchPAT and clinical disease activity data (n = 33–37), the association of WatchPAT-assessed sleep characteristics with RA disease activity were adjusted for BMI only. Sensitivity analyses in the WatchPAT sample examined whether the association between SDB and disease activity was independent of sleep efficiency by adding sleep efficiency to the models. The data underlying this article will be shared upon reasonable request to the corresponding author.

RESULTS

Of the 133 individuals who enrolled in the study, 116 had sufficient actigraph wear time for scoring at baseline, 63 completed WatchPAT monitoring, and 58 completed both WatchPAT and actigraph assessments at baseline. Actigraph data were available for 83, 64, and 66 participants at 6, 12, and 18 months of follow-up, respectively. Forty-two participants completed four actigraph assessments, 36 completed two, 15 completed three, and 23 completed only baseline. Because of the relatively large loss (26%) to follow-up at the 6-month follow-up, we examined differences in age, sex, sleep parameters, disease duration, and all measures of RA disease activity between those who were lost to follow-up and those who were not; no significant differences were found (data not shown). Missing actigraph data through the study were primarily because of nonuse (25 of the 56 total missing cases over all four measurement periods) and malfunction in uploading data for scoring (10 of 56) (see Supplemental Table S2). Considering all data-collection points, 63% of contacts were during the COVID pandemic (between July 17, 2020, and February 8, 2022; 43%, 54%, 79%, and 91% of baseline, 6-month, 12-month, and 18-month follow-ups, respectively) and were conducted remotely.

Table 1. Participant baseline characteristics*

Characteristics	Entire cohort (N = 116)	WatchPat subset (n = 63)
Sociodemographics		
Age, mean ± SD, y	57.9 ± 13.3	57.5 ± 12.6
Female sex, % (n)	88.8 (103)	81.0 (51)
Race, % (n)		
Asian	6.0 (7)	6.4 (4)
Black	7.8 (9)	11.1 (7)
White	54.3 (63)	39.7 (25)
Other	31.9 (37)	42.9 (27)
Latino/a	33.6 (39)	39.7 (25)
Education, HS or less, % (n)	28.5 (33)	33.3 (21)
General health		
BMI, mean ± SD, kg/m^2	27.6 ± 5.9	28.6 ± 5.4
Smoking, % (n)		
Ever	40.5 (47)	46.0 (29)
Current	5.2 (6)	9.5 (6)
Other chronic health conditions, % (n)		
Hypertension	38.8 (45)	48.6 (17)
Asthma	15.5 (18)	20.6 (13)
COPD	7.0 (8)	8.1 (5)
Heart disease	6.0 (7)	9.5 (6)
Diabetes	12.2 (14)	11.3 (7)
RA characteristics, mean ± SD		
Disease duration, y	16.4 ± 13.0	16.5 ± 12.8
RADAI (range 0–10)	3.3 ± 1.9	3.5 ± 1.9
Pain rating (range 0–10)	3.3 ± 2.5	3.6 ±2.4
RA affect (range 0–100)	29.4 ± 22.8	32.0 ± 22.9
RA medications, % (n)		
Glucocorticoids		
Any	31.0 (36)	31.8 (20)
High dose	6.9 (8)	9.5 (6)
Conventional DMARD	63.8 (74)	66.7 (42)
TNFi	40.5 (47)	37.1 (13)
Abatacept	3.5 (4)	0
Tocilizumab	4.3 (5)	4.8 (3)
Rituxan	4.3 (5)	6.4 (4)
Any pain medications in the past month	28.5 (33)	41.3 (26)
RA disease activity, mean ± SD		
CDAI (range 0–76)	14.6 ± 11.6 (n = 42)	13.6 ± 9.8 (n = 37)
DAS-ESR (range 0–9.4)	3.1 ± 1.1 (n = 38)	3.0 ± 1.0 (n = 34)
DAS-CRP (range 0–9.4)	3.1 ± 1.3 (n = 38)	3.3 ± 1.2 (n = 33)

* Score ranges are possible minimum and maximum scores. BMI, body mass index; CDAI, Clinical Disease Activity Index; COPD, chronic obstructive pulmonary disease; CRP, C-reactive protein; DAS, Disease Activity Score; DMARD, disease-modifying antirheumatic drug; ESR, erythrocyte sedimentation rate; HS, high school; RA, rheumatoid arthritis; RADAI, RA Disease Activity Index; TNFi, tumor necrosis factor inhibitor.

Characteristics of the analytic sample are shown in Table 1. The mean ± SD age was 58 ± 13 years, almost 90% were female, 34% were Latino/a, 29% had high school education or less, and 5.2% were current smokers. The most common comorbid conditions were hypertension (38.8%), asthma (15.5%), and diabetes (12.2%). The mean ± SD disease duration was approximately 16 ± 13 years. A total of 31% were currently using GCs, and 6.9% were using moderate- or high-dose GCs (≥7.5 mg/day). Additional details on RA medications are shown in Table 1. Approximately one-quarter reported the use of prescription pain medications in the past month.

CDAI scores were available for 42 participants at baseline, 29 at the 6-month period, and 11 at the 12-month period, for a total of 82 observations. DAS scores were available for 38 participants at baseline, 23 at the 6-month period, and 8 at the 12-month period, for a total of 69 observations. The mean ± SD CDAI score at baseline were 14.6 ± 11.6. The mean ± SD DAS-ESR and DAS-CRP scores were 3.1 ± 1.1 and 3.4 ± 1.3, respectively

Characteristics of the 63 participants in the WatchPAT subset are also shown in Table 1. Eighty-one percent were female, 40% were Latino/a, 33% had high school or less education, and 9.5% were current smokers. Thirty-two percent were currently using GCs, and 9.5% were using high-dose GCs.

Actigraph assessments of sleep

The mean ± SD sleep efficiency at baseline assessment was 83.8% ± 5.7% (Table 2) and was relatively consistent across the four measurement periods. More than half of the sample had a mean sleep efficiency <85% at baseline; 60.2%, 48.4%, and 57.6% had sleep efficiency <85% at the 6-, 12-, and 18-month periods, respectively; and >40% had sleep efficiency <85% at every measurement period in which they participated. Sleep efficiency <80% was present in approximately one-quarter to one-third of the participants over the four measurement periods. At each period, approximately two-thirds of the participants had worse sleep efficiency than the published age and sex reference values.[35]

The mean ± SD WASO time was 71.2 ± 32.0 minutes at baseline, with slightly longer wake times at the follow-up periods. More than two-thirds of the cohort had WASO >60 minutes at each period and more than half had WASO >60 minutes at every period. Between one-quarter and one-third had WASO >90 minutes at each period. High WASO according to the age and sex reference values was present for between 85.9% and 87.1% across the four measurement periods, and 81% had high WASO according to age and sex reference values at every measurement period.

The mean ± SD total sleep time was 6.9 ± 1.0 hours at baseline, with similar times at the other measurement periods. At least half of the sample had an average of <7 hours sleep per night at each measurement period, and roughly one in five had an average of <6 hours per night. Approximately half had short sleep according to the age and sex reference values at each time period.

WatchPAT assessments

More than half of the participants with WatchPAT assessments (n = 63) had moderate to severe OSA (AHI > 15); only

Table 2. Actigraphy results*

	By measurement period				Cumulative		
	Baseline	6 months	12 months	18 months	At all measurement periods	Ever (≥1 period)	Never
N	116	83	64	66	–	–	–
Sleep characteristic							
Sleep efficiency, mean ± SD	83.8 ± 5.7	83.5 ± 5.2	83.5 ± 5.9	82.8 ± 7.2	–	–	–
Sleep efficiency <85%, % (n)	51.7 (60)	60.2 (50)	48.4 (31)	57.6 (38)	41.4 (48)	69.8 (81)	30.2 (35)
Sleep efficiency <80%, % (n)	23.3 (27)	26.5 (22)	25.0 (16)	33.3 (22)	15.5 (18)	38.8 (45)	61.2 (71)
Poor sleep efficiency for age and sex, % (n)	69.3 (81)	71.1 (59)	67.2 (43)	68.2 (45)	60.3 (70)	76.7 (89)	23.3 (27)
WASO, mean ± SD, min	71.2 ± 32.0	78.1 ± 29.8	78.1 ± 30.9	82.4 ± 40.0	–	–	–
WASO >60 min, % (n)	69.8 (81)	72.3 (60)	67.2 (43)	68.2 (45)	56.9 (66)	80.2 (93)	19.8 (23)
WASO >90 min, % (n)	25.9 (30)	28.9 (24)	32.8 (21)	34.9 (23)	18.1 (21)	42.2 (49)	57.8 (67)
High WASO for age and sex, % (n)	87.1 (101)	85.5 (71)	85.9 (55)	84.9 (56)	81.0 (94)	93.1 (108)	6.9 (8)
Total sleep time, mean ± SD, h	6.9 ± 1.0	6.7 ± 1.0	6.9 ± 1.0	6.8 ± 1.1	–	–	–
Sleep time <7 h, % (n)	53.5 (62)	57.8 (48)	50.0 (32)	59.1 (39)	40.5 (47)	64.7 (75)	35.3 (41)
Sleep time <6 h, % (n)	18.1 (21)	21.7 (18)	20.3 (13)	18.2 (12)	11.2 (13)	28.5 (33)	71.6 (83)
Short sleep for age and sex, % (n)	51.7 (60)	55.4 (46)	46.9 (30)	51.6 (38)	40.5 (47)	62.9 (73)	37.1 (43)

* Sleep efficiency is time sleeping divided by the time in bed for sleep, expressed as a percentage. Age and sex reference values are from Hertenstein et al[35] (see Supplemental Table S1). WASO, time awake after sleep onset.

9.5% had a normal AHI (AHI < 5) (Table 3). Of those who had moderate to severe OSA, only 54.3% (19 of 35) reported that they had ever received a diagnosis of OSA. A substantial portion of the sample had periods of low oxygen (O_2) saturation. Five percent had more than 30 minutes with <85% O_2 saturation. More than three-quarters had some time with <90% O_2 saturation, with 14% recording >120 minutes at <90%.

Overlap of sleep conditions

In the actigraph cohort, approximately one-third had both low sleep efficiency and short sleep, 18% had short sleep only, and 16% had low sleep efficiency only (Figure 1). Among those with WatchPAT and actigraph assessments, 21% (12 of 58) had OSA only, 7% had low sleep efficiency only, 14% had short sleep only, and 22% (13 of 58) had combined OSA, poor sleep efficiency, and short sleep. Only 5% (3 of 58) had none of the sleep disorders.

Table 3. WatchPAT results (n = 63)*

Sleep characteristics	Data
AHI, mean ± SD (range)	18.8 ± 11.5 (1.4–50.5)
ODI, mean ± SD (range)	9.6 ± 8.2 (0.2–36.4)
AHI category, % (n)	
No OSA (0 to <5)	9.5 (6)
Mild (5 to <15)	34.9 (22)
Moderate (15 to <30)	38.1 (24)
Severe (≥30)	17.5 (11)
Ever diagnosed with OSA, % (n)	
None/mild AHI (0 to <15), n = 28	3.6 (1)
Moderate/severe (≥15), n = 35	54.3 (19)
Oxygen desaturation time, % (n)	
Any time <80% O_2 saturation	10.3 (6)
Any time <85% O_2 saturation	43.8 (25)
>30 min <85%	5.3 (3)
Any time <90% O_2 saturation	76.9 (40)
>30 min	39.6 (23)
>60 min	27.6 (16)
>120 min	13.8 (8)

* AHI, apnea-hypopnea index; O_2, oxygen; ODI, oxygen desaturation index; OSA, obstructive sleep apnea.

Association of sleep conditions with RA disease activity

Among the subset of participants with in-person visits and measures of disease activity, poorer sleep efficiency was significantly associated with greater disease activity by all three disease activity measures in analyses adjusted for age, sex, BMI, comorbidities, and RA duration (Table 4). Greater WASO time was also significantly associated with greater disease activity. There were no statistically significant associations between sleep time and disease activity measures.

Examination of the components of the disease activity measures showed that the one most consistently associated with sleep efficiency was the tender joint count (Supplemental Tables S3 and S4). Sleep efficiency <80% was associated with both tender and swollen joint counts, physician global assessment, and CRP. WASO was associated with tender and swollen joint counts and physician global assessment.

In analyses adjusted for BMI, greater (worse) AHI obtained from the WatchPAT measurements was significantly associated with worse disease activity by CDAI and DAS-CRP (Table 4). A higher ODI was also associated with higher CDAI scores in adjusted analyses. Further adjustment for sleep efficiency slightly attenuated the parameter estimates, but statistically significant associations remained between CDAI scores and both AHI and

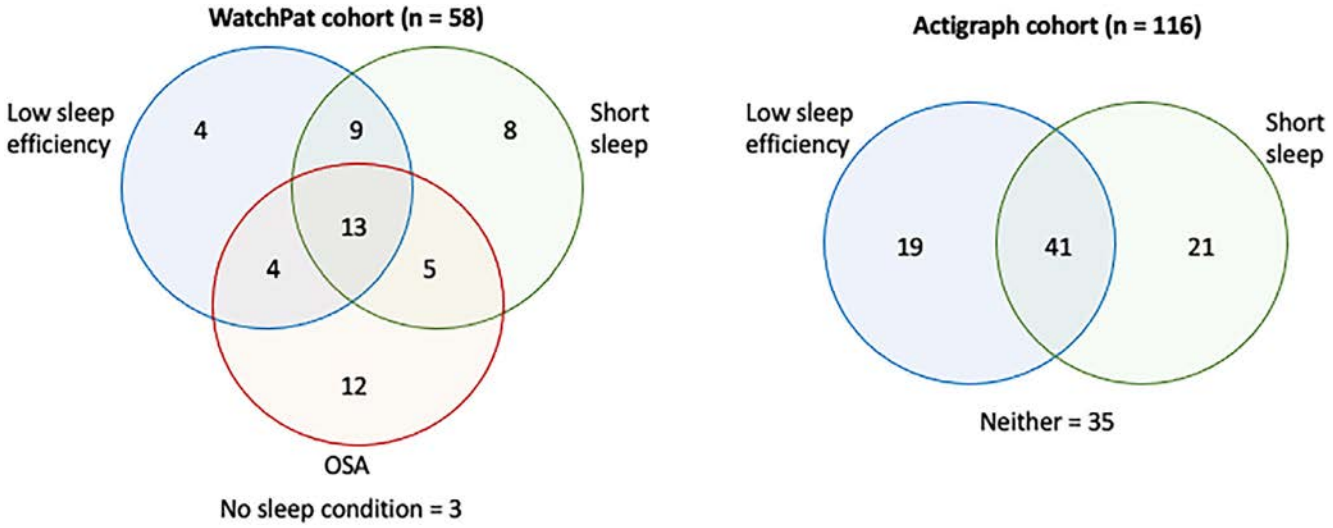


Figure 1. Overlap among sleep conditions at baseline (sleep efficiency <85% and sleep <7 hours). The WatchPAT cohort includes those with one condition only (OSA only, 12 [21%]; low efficiency only, 4 [7%]; and short sleep only, 8 [14%]), two conditions (OSA and low efficiency, 5 [9%]; OSA and short sleep, 5 [9%]; and low efficiency and short sleep, 9 [16%]), all three conditions (13 [22%]), and none of the conditions (3 [5%]). The Actigraph cohort includes those with low efficiency only (19 [16%]), short sleep only (21 [18%]), both (41 [35%]), and neither (35 [30%]). OSA, obstructive sleep apnea.

ODI and between DAS-CRP and AHI ≥ 15 (Supplemental Tables S5 and S6).

DISCUSSION

In this study, we sought to identify the prevalence of objectively measured sleep characteristics in individuals with RA and their association with clinical measures of RA disease activity. Most of this sample had at least one sleep disturbance or disorder. More than half had poor sleep efficiency defined as <85% at baseline, with similar proportions at each measurement period. With a lowered criterion of <80%, almost one-quarter were still classified with poor sleep efficiency. At each measurement period, the mean time awake after sleep onset was >1 hour. At all measurement periods, more than half of the sample had <7 hours of measured sleep, the Centers for Disease Control and Prevention criterion for adequate sleep; roughly one in five had <6 hours. When comparing results with age and sex reference values, more than two-thirds had poor sleep efficiency at each measurement period, approximately three-quarters had high WASO, and approximately half had inadequate sleep time. Worse sleep efficiency and greater WASO were each significantly associated with clinical RA disease activity measures.

Among the smaller cohort with both WatchPAT and actigraph assessments, more than half had moderate to severe OSA, yet less than one-quarter had ever had a diagnosis of OSA. Only 5% (n = 3) had no sleep problem, whereas 22% had combined high AHI, poor sleep efficiency, and short sleep.

Few studies have reported measured sleep efficiency in RA. Two recent studies reported mean sleep efficiency in RA cohorts of 83% and 86%,[11,12] quite similar to what we found. In general population studies of adults of similar age to our cohort, sleep efficiency estimates derived from actigraphy have been similar, ranging from 82% to 85%.[36,37] Estimates derived from PSG have been slightly higher, 90.5% for women and 86.4% for men on a second night of measurement.[38] However, sleep parameters generally vary by age and sex, and when we compared sleep efficiency with PSG-derived age and sex reference values from individuals without RA or a known sleep disorder,[35] approximately 70% had poor sleep efficiency at each measurement period. WASO time also appeared to be higher than in the general population. We found mean WASO times of well over an hour at each measurement period, whereas general population studies of adults in the same age range as our cohort reported average WASO times ranging from 25 to 64 minutes.[36–38] Compared with age and sex reference values,[35] approximately 85% of our sample had high WASO. Combined, our sleep efficiency and WASO data suggest that objectively measured sleep quality is substantially worse among people with RA compared with the general adult population. The mean ± SD sleep time in our cohort (6.9 ± 1.0 hours) was slightly lower than findings from a general population study of objectively measured sleep among older adults,[36] but when compared with age and sex reference values from healthy adults, approximately half of the participants in this study had short sleep times.

The prevalence of severe OSA in this sample—more than half had AHI ≥ 15—was considerably higher than estimates from general population studies. For example, a study from the 2005 to 2006 National Health and Nutrition Examination Survey reported physician-diagnosed OSA in 7.1%,[39] and data from the Nurses Health Study, Nurses Health Study II, and Health Professionals Follow-up Study estimated a prevalence of clinically

Table 4. Association of sleep parameters with disease activity*

Sleep parameters	Unadjusted			Adjusted[a]		
	CDAI	DAS-CRP	DAS-ESR	CDAI	DAS-CRP	DAS-ESR
Actigraphy						
Observations, n	82	69	68	–	–	–
Participants, n	54	48	47	–	–	–
Sleep efficiency (continuous)						
Coefficient	**−0.60**	**−0.051**	**−0.055**	**−0.73**	**−0.07**	**−0.07**
95% CI	**−1.1 to −0.12**	**−0.10 to −0.003**	**−0.11 to 0.001**	**−1.17 to −0.29**	**−0.11 to −0.02**	**−0.13 to −0.02**
P	**0.013**	**0.038**	**0.056**	**0.001**	**0.005**	**0.005**
Sleep efficiency <85%						
Coefficient	4.22	0.39	0.39	**5.34**	**0.56**	**0.65**
95% CI	−0.88 to 9.33	−0.11 to 0.90	−0.17 to 0.96	**0.31 to 10.39**	**0.05 to 1.06**	**0.10 to 1.20**
P	0.11	0.13	0.18	**0.038**	**0.030**	**0.021**
Sleep efficiency <80%						
Coefficient	**9.89**	**0.98**	**0.88**	**10.43**	**0.97**	**0.94**
95% CI	**4.22 to 15.54**	**0.42 to 1.54**	**0.24 to 1.52**	**5.12 to 15.75**	**0.44 to 1.51**	**0.35 to 1.54**
P	**0.001**	**0.001**	**0.007**	**0.000**	**0.000**	**0.002**
WASO						
Coefficient	**0.14**	**0.013**	**0.015**	**0.14**	**0.012**	**0.014**
95% CI	**0.06 to 0.22**	**0.005 to 0.02**	**0.006 to 0.024**	**0.07 to 0.21**	**0.005 to 0.02**	**0.007 to 0.02**
P	**0.001**	**0.002**	**0.001**	**0.000**	**0.001**	**0.000**
Total sleep time[a]						
<7 h						
Coefficient	−3.34	−0.41	−0.51	−2.50	−0.29	−0.38
95% CI	−8.60 to 1.91	−0.94 to 0.12	−1.09 to 0.07	−7.86 to 2.87	−0.82 to 0.25	−0.95 to 0.20
P	0.21	0.13	0.08	0.36	0.29	0.20
>8 h						
Coefficient	2.59	−0.02	0.25	2.15	−0.07	0.07
95% CI	−7.50 to 12.69	−0.98 to 0.94	−0.80 to 1.30	−7.68 to 11.98	−1.00 to 0.86	−0.95 to 1.10
P	0.62	0.96	0.65	0.67	0.88	0.89
Total sleep time[b]						
<6 h						
Coefficient	0.92	−0.09	−0.29	1.61	0.018	−0.09
95% CI	−4.94 to 6.77	−0.66 to 0.48	−0.92 to 0.34	−4.1 to 7.32	−0.55 to 0.59	−0.71 to 0.53
P	0.76	0.76	0.37	0.58	0.95	0.77
>8 h						
Coefficient	4.58	0.19	0.48	3.51	0.08	0.26
95% CI	−5.25 to 14.40	−0.74 to 1.12	−0.54 to 1.50	−6.09 to 13.12	−0.83 to 0.99	−0.73 to 1.25
P	0.36	0.69	0.35	0.47	0.87	0.61
WatchPAT, n	37	34	33	–	–	–
AHI						
Coefficient	**0.26**	0.02	0.02	**0.40**	**0.034**	0.031
95% CI	**0.01 to 0.51**	−0.007 to 0.049	−0.01 to 0.05	**014 to 0.67**	**0.003 to 0.067**	−0.01 to 0.069
P	**0.043**	0.13	0.22	**0.004**	**0.033**	0.10
AHI ≥ 15						
Coefficient	5.1	**0.77**	0.68	**7.09**	**0.98**	0.84
95% CI	−1.6 to 11.8	**0.05 to 1.49**	−0.18 to 1.53	**0.15 to 14.02**	**0.23 to 1.73**	−0.06 to 1.75
P	0.13	**0.039**	0.12	**0.045**	**0.01**	0.067
ODI						
Coefficient	0.32	0.02	0.03	**0.53**	0.043	0.044
95% CI	−0.05 to 0.68	−0.02 to 0.06	−0.02 to 0.07	**0.13 to 0.92**	−0.004 to 0.09	−0.01 to 0.10
P	0.09	0.24	0.26	**0.011**	0.074	0.12

* Bold font denotes statistically significant results, $P < 0.05$. Tabled values are coefficient, 95% CI, and *P* value from mixed-effects linear regression analyses. Actigraph analysis–adjusted analyses control for age, sex, body mass index, number of comorbid conditions, and rheumatoid arthritis duration. WatchPAT analysis was adjusted only for body mass index. AHI, apnea-hypopnea index; CDAI, Clinical Disease Activity Index; CI, confidence interval; CRP, C-reactive protein; DAS, Disease Activity Score; ESR, erythrocyte sedimentation rate; ODI, oxygen desaturation index; WASO, time awake after sleep onset.
[a] Reference group: 7 to 8 hours.
[b] Reference group: 6 to 8 hours.

relevant OSA (ie, AHI ≥ 15) among individuals aged 60 to 65 years of 8.2% in women and 14.6% in men.[40]

The rates of OSA were also higher than what has been seen in PSG studies of sleep among individuals with RA. For example, Mutoh et al and Wali et al found OSA, defined as AHI ≥ 15, in 20.9% and 22.9%, respectively, of patients with RA tested by PSG.[6,41] The prevalence of any OSA (AHI ≥ 5) in our cohort (90.5%) was much higher that reported in other studies, in which

rates were generally 58% to 68%.[41–43] One study based on medical records found rates of OSA similar to those we found,[9] but as noted for previous PSG studies, these were derived from clinical sleep studies in which a higher rate might be expected. Claims-based studies have also found a significantly higher risk of OSA for individuals with RA.[44]

Sleep efficiency and WASO were each significantly associated with all three clinical measures of RA disease activity, and measures of SDB in the smaller subset were significantly associated with two of the three. Previous research has shown an association between self-reported sleep and RA disease activity and pain thresholds[15] and between objective measures of sleep and specific measures of inflammation,[17] but this study is the first to report the associations with clinical RA disease activity measures. In addition, we found that both sleep efficiency and WASO were consistently associated with both tender and swollen joint counts and with physician global ratings. Although we did not assess the directionality of these findings, they suggest that either sleep disturbances may lead to greater RA-related pain and disease activity and/or that RA pain and disease activity may impair sleep, the latter being important because of known associations of sleep disturbances with other health conditions such as cardiovascular disease. In either case, these associations support the clinical relevance of the high prevalence of sleep disorders that we found.

Although some general population studies show the association of both short and long sleep times with negative health outcomes, in this study, sleep time was not associated with disease activity as expected. This finding suggests that the quantity of sleep may be less important than the quality of sleep. Similar results have been found for other conditions such as cognitive function and dementia.

In addition to the associations we found between sleep disturbances and RA disease activity, there are other health effects from poor sleep that should also be considered. Of particular relevance to RA, sleep disturbances are associated with lower pain thresholds generally, a finding that seems magnified in individuals with RA.[4] Poor sleep is also associated with higher levels of fatigue, higher levels of depressive symptoms, cognitive dysfunction, poor physical function, and greater systemic inflammation, all of which are already more prevalent in people living with RA.[11,45–47] OSA is associated with increased risk of cardiovascular disease, stroke, and cognitive dysfunction, again, all of which are already more prevalent in people with RA.[47,48]

The limitations of this study include the discontinuation of WatchPAT measurement and disease activity assessments with the onset of the COVID pandemic, resulting in a relatively small sample for these measures. However, the measurements were available on a sample larger than most studies in RA. We report associations between sleep and disease activity only. The relatively small number of observations with disease activity measures at follow-up prohibited us from examining longitudinal or time-ordered associations between sleep disorders and disease activity. In-home measures of sleep may be different from what PSG studies would find, although both the actigraphy and WatchPAT devices have been validated against PSG studies. In addition, some argue that home-based measures of sleep may be more accurate reflections of usual sleep than laboratory-based PSG. Our sample of individuals with RA had relatively low disease activity and pain; studies of individuals with greater disease activity may yield different results. Although our study sample was larger than most studies of objective sleep measurement in RA, it may still have been too small to yield significant differences in associated variables. We did not find differences in baseline sleep or disease activity parameters between those who remained in the study and those who dropped out at 6 months; however, it is possible that unmeasured variables created a systematic bias. A relatively large portion of the data for the study was collected during the COVID-19 pandemic, and sleep may have universally been more disturbed during this time.[49]

The limitations of the study are balanced by its strengths. Sleep was measured objectively using well-tested and validated devices, protocols, and scoring methods. The study cohort was diverse and larger than any previously published studies of objective sleep measurement in RA. Sleep was assessed over up to four measurement periods over 18 months, and clinical RA disease measures were available for a substantial portion of the measurement periods.

In summary, the frequency of objectively measured sleep disorders in this cohort was higher than would be expected compared with general population studies and age and sex reference values. There appears to be a significant association between sleep disorders—poor sleep efficiency, in particular—and RA disease activity. Given the high rates of sleep disorders, the significant negative general and RA-specific health effects of sleep disorders, and the existence of treatments that can improve sleep quality, our findings point to a critical need for greater attention to sleep disorders among individuals living with RA.

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 Katz 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. Westhovens R, Van der Elst K, Matthys A, et al. Sleep problems in patients with rheumatoid arthritis. J Rheumatol 2014;41(1):31–40.

2. McKenna S, Donnelly A, Fraser A, et al. Sleep and physical activity: a survey of people with inflammatory arthritis and their engagement by health professionals in rheumatology in Ireland. Disabil Rehabil 2018; 40(19):2260–2266.
3. Zhang L, Shen B, Liu S. Rheumatoid arthritis is associated with negatively variable impacts on domains of sleep disturbances: evidence from a systematic review and meta-analysis. Psychol Health Med 2021;26(3):267–277.
4. Irwin MR, Olmstead R, Carrillo C, et al. Sleep loss exacerbates fatigue, depression, and pain in rheumatoid arthritis. Sleep 2012;35(4): 537–543.
5. Bjurström MF, Olmstead R, Irwin MR. Reciprocal relationship between sleep macrostructure and evening and morning cellular inflammation in rheumatoid arthritis. Psychosom Med 2017;79(1): 24–33.
6. Mutoh T, Okuda Y, Mokuda S, et al. Study on the frequency and risk factors of moderate-to-severe sleep apnea syndrome in rheumatoid arthritis. Mod Rheumatol 2016;26(5):681–684.
7. Wilton KM, Matteson EL, Crowson CS. Risk of obstructive sleep apnea and its association with cardiovascular and noncardiac vascular risk in patients with rheumatoid arthritis: a population-based study. J Rheumatol 2018;45(1):45–52.
8. Katz P, Pedro S, Michaud K. Sleep disorders among individuals with rheumatoid arthritis. Arthritis Care Res (Hoboken) 2023;75(6):1250–1260.
9. George RJ, Kumar R, Achenbach SJ, et al. Sleep disorders in rheumatoid arthritis: incidence, risk factors and association with dementia. Semin Arthritis Rheum 2025;73:152722.
10. McKenna S, Tierney M, O'Neill A, et al. Sleep and physical activity: a cross-sectional objective profile of people with rheumatoid arthritis. Rheumatol Int 2018;38(5):845–853.
11. McBeth J, Dixon WG, Moore SM, et al. Sleep disturbance and quality of life in rheumatoid arthritis: prospective mHealth study. J Med Internet Res 2022;24(4):e32825.
12. Aydemir B, Muhammad LN, Song J, et al. Characterization of sleep disturbance in established rheumatoid arthritis patients: exploring the relationship with central nervous system pain regulation. BMC Rheumatol 2024;8(1):33.
13. Song J, Muhammad LN, Neogi T, et al. Pain sensitization as a potential mediator of the relationship between sleep disturbance and subsequent pain in rheumatoid arthritis. Arthritis Care Res (Hoboken) 2023;75(4):778–784.
14. Aydemir B, Schieir O, Valois MF, et al. Association between sleep disturbance and subsequent pain interference in patients with early rheumatoid arthritis. Arthritis Care Res (Hoboken) 2025;77(9):1078–1084.
15. Son CN, Choi G, Lee SY, et al. Sleep quality in rheumatoid arthritis, and its association with disease activity in a Korean population. Korean J Intern Med 2015;30(3):384–390.
16. Azzam A. The impact of sleep problems on rheumatoid arthritis disease activity. Sleep Med Res 2024;15(2):106–112.
17. Irwin MR, Straub RH, Smith MT. Heat of the night: sleep disturbance activates inflammatory mechanisms and induces pain in rheumatoid arthritis. Nat Rev Rheumatol 2023;19(9):545–559.
18. Vgontzas AN, Chrousos GP. Sleep, the hypothalamic-pituitary-adrenal axis, and cytokines: multiple interactions and disturbances in sleep disorders. Endocrinol Metab Clin North Am 2002; 31(1):15–36.
19. Pillar G, Berall M, Berry R, et al. Detecting central sleep apnea in adult patients using WatchPAT — a multicenter validation study. Sleep Breath 2020;24(1):387–398.
20. Jen R, Orr JE, Li Y, et al. Accuracy of WatchPAT for the diagnosis of obstructive sleep apnea in patients with chronic obstructive pulmonary disease. COPD 2020;17(1):34–39.
21. Budig M, Stoohs R, Keiner M. Validity of two consumer multisport activity tracker and one accelerometer against polysomnography for measuring sleep parameters and vital data in a laboratory setting in sleep patients. Sensors (Basel) 2022;22(23):9540.
22. Arguello D, Andersen K, Morton A, et al. Validity of proximity sensor-based wear-time detection using the ActiGraph GT9X. J Sports Sci 2018;36(13):1502–1507.
23. Troiano RP, Berrigan D, Dodd KW, et al. Physical activity in the United States measured by accelerometer. Med Sci Sports Exerc 2008;40(1):181–188.
24. Smith MT, McCrae CS, Cheung J, et al. Use of actigraphy for the evaluation of sleep disorders and circadian rhythm sleep-wake disorders: an American Academy of Sleep Medicine clinical practice guideline. J Clin Sleep Med 2018;14(7):1231–1237.
25. Cole RJ, Kripke DF, Gruen W, et al. Automatic sleep/wake identification from wrist activity. Sleep 1992;15(5):461–469.
26. Aletaha D, Nell VP, Stamm T, et al. Acute phase reactants add little to composite disease activity indices for rheumatoid arthritis: validation of a clinical activity score. Arthritis Res Ther 2005;7(4):R796–R806.
27. van der Heijde DM, van 't Hof MA, van Riel PL, et al. Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score. Ann Rheum Dis 1990;49(11): 916–920.
28. Johnson TM, Michaud K, England BR. Measures of rheumatoid arthritis disease activity. Arthritis Care Res (Hoboken) 2020;72 Suppl 10 (suppl 10):4–26.
29. Reed DL, Sacco WP. Measuring sleep efficiency: what should the denominator be? J Clin Sleep Med 2016;12(2):263–266.
30. Wang F, Sun Z, Lin F, et al. Nonlinear relationships between sleep duration, mental health, and quality of life: the dangers of less sleep versus more sleep. Sleep Med 2024;119:565–573.
31. Centers for Disease Control and Prevention. Sleep facts and stats. Published May 15, 2024. Accessed January 12, 2006. https://www.cdc.gov/sleep/data-research/facts-stats/index.html
32. Ford ES, Cunningham TJ, Croft JB. Trends in self-reported sleep duration among US adults from 1985 to 2012. Sleep 2015;38(5): 829–832.
33. Domínguez F, Fuster V, Fernández-Alvira JM, et al. Association of sleep duration and quality with subclinical atherosclerosis. J Am Coll Cardiol 2019;73(2):134–144.
34. Hertenstein E, Gabryelska A, Spiegelhalder K, et al. Reference data for polysomnograpy-measured and subjective sleep in healthy adults. J Clin Sleep Med 2018;14(4):523–532.
35. Hertenstein E, Gabryelska A, Spiegelhalder K, et al. Reference data for polysomnography-measured and subjective sleep in healthy adults. J Clin Sleep Med 2018;14(4):523–532.
36. Kurina LM, Thisted RA, Chen JH, et al. Actigraphic sleep characteristics among older Americans. Sleep Health 2015;1(4):285–292.
37. Evans MA, Buysse DJ, Marsland AL, et al. Meta-analysis of age and actigraphy-assessed sleep characteristics across the lifespan. Sleep 2021;44(9):zsab088.
38. Boulos MI, Jairam T, Kendzerska T, et al. Normal polysomnography parameters in healthy adults: a systematic review and meta-analysis. Lancet Respir Med 2019;7(6):533–543.
39. Ram S, Seirawan H, Kumar SK, et al. Prevalence and impact of sleep disorders and sleep habits in the United States. Sleep Breath 2010; 14(1):63–70.
40. Huang T, Lin BM, Markt SC, et al. Sex differences in the associations of obstructive sleep apnoea with epidemiological factors. Eur Respir J 2018;51(3):1702421.
41. Wali S, Mustafa M, Manzar D, et al. Prevalence of obstructive sleep apnea in patients with rheumatoid arthritis. J Clin Sleep Med 2020; 16(2):259–265.

42. Taylor-Gjevre RM, Gjevre JA, Nair B, et al. Components of sleep quality and sleep fragmentation in rheumatoid arthritis and osteoarthritis. Musculoskeletal Care 2011;9(3):152–159.
43. Gjevre JA, Taylor-Gjevre RM, Nair BV, et al. Do sleepy rheumatoid arthritis patients have a sleep disorder? Musculoskeletal Care 2012; 10(4):187–195.
44. Shen TC, Hang LW, Liang SJ, et al. Risk of obstructive sleep apnoea in patients with rheumatoid arthritis: a nationwide population-based retrospective cohort study. BMJ Open 2016; 6(11):e013151.
45. Spira AP, Covinsky K, Rebok GW, et al. Poor sleep quality and functional decline in older women. J Am Geriatr Soc 2012;60(6):1092–1098.
46. Irwin MR, Opp MR. Sleep health: reciprocal regulation of sleep and innate immunity. Neuropsychopharmacology 2017;42(1):129–155.
47. Leng Y, McEvoy CT, Allen IE, et al. Association of sleep-disordered breathing with cognitive function and risk of cognitive impairment. A systematic review meta-analysis. JAMA Neurol 2017;74(10):1237–1245.
48. Stone KL, Blackwell TL, Ancoli-Israel S, et al; Osteoporotic Fractures in Men (MrOS) Study Research Group. Sleep disordered breathing and risk of stroke in older community-dwelling men. Sleep 2016; 39(3):531–540.
49. Jahrami H, BaHammam A, Bragazzi N, et al. Sleep problems during the COVID-19 pandemic by population: a systemic review and meta-analysis. J Clin Sleep Med 2021;17:299–313.

Arthritis Care & Research
Vol. 78, No. 6, June 2026, pp 723–732
DOI 10.1002/acr.25684

The Renal Activity Index for Lupus Identifies Active Renal Disease and Treatment Response in Adult Patients With Systemic Lupus Erythematosus and Lupus Nephritis

Hermine I. Brunner,[1] Ellen M. Cody,[2] Prasad Devarajan,[1] Bin Huang,[1] Chen Chen,[1] Dominic Sinibaldi,[3] Madhu Ramaswamy,[3] Jacob Knagenhjelm,[4] Frederick Jones,[5] Philip Z. Brohawn,[3] Raj Tummala,[3] Catharina Lindholm,[4] and Wendy I. White[3]

Objective. We evaluated the ability of the Renal Activity Index for Lupus (RAIL) to discriminate active lupus nephritis (LN) in adult patients with active systemic lupus erythematosus (SLE) and differentiate LN treatment response.

Methods. Urine samples from adults with biopsy-proven active class III and IV LN from TULIP-LN (active LN group; ClinicalTrials.gov identifier: NCT02547922) and adults with active, nonrenal SLE from TULIP-1 (active SLE group; ClinicalTrials.gov identifier: NCT02446912) were used, and RAIL biomarkers (neutrophil gelatinase–associated lipocalin, kidney injury molecule 1, monocyte chemotactic protein 1, adiponectin, hemopexin, and ceruloplasmin) were measured in the urine at baseline for both studies and at weeks 12 and 24 for TULIP-LN only. The groups were compared at baseline, and changes in RAIL scores from baseline in the active LN group were compared between nonresponders and responders over time, ie, those with complete renal response (CRR), partial renal response (PRR), and urine protein–creatine ratio decrease ≥50% (UPCR50).

Results. At baseline, median (interquartile range [IQR]) concentrations of RAIL biomarkers were significantly higher ($P \leq 0.02$) in the active LN group (n = 128) versus the SLE control group (n = 48), as were RAIL scores (5.59 [IQR 4.31–6.47] vs 3.57 [IQR 2.78–4.47]; $P < 0.001$). At weeks 12 and 24, there were 25 and 31 patients with CRR, 39 and 54 patients with PRR, and 41 and 63 patients with UPCR50, respectively. Changes in RAIL scores from baseline to weeks 12 and 24 significantly differed between nonresponders and responders (PRR, CRR, UPCR50: all $P < 0.0006$), with lower scores in responders. For CRR versus non-response, median [IQR] RAIL scores decreased by –1.3 (–3.64 to –0.21) versus –0.39 at week 12, and –2.30 (–3.63 to –1.03) versus –0.88 (–2.20 to 0.33) at week 24, respectively.

Conclusions. The RAIL identifies active LN and longitudinally differentiates treatment response in adults with LN.

INTRODUCTION

In patients with systemic lupus erythematosus (SLE), lupus nephritis (LN) is considered one of the most common, severe organ manifestations and a leading cause of morbidity and mortality.[1,2] LN affects up to 40% of patients with SLE, with a higher prevalence observed in women, individuals of lower socioeconomic status, and individuals of African American, Asian, and Hispanic or Latino race and ethnicity.[1,3–6]

Previously presented at SLEuro 2022, The 13th European Lupus Meeting, Stockholm, Sweden, October 5–8, 2022; and the American College of Rheumatology Convergence, Philadelphia, PA, November 10–14, 2022.

Supported by AstraZeneca. This research was also supported by a Transformational Research Award of the Marianne and Ralph Falk Medical Trust as well as the National Institutes of Arthritis and Musculoskeletal and Skin Diseases Core Center, NIH (grants P30-AR-076316 to Dr Brunner, principal investigator [PI], and P30-AR-070549).

[1]Cincinnati Children's Hospital Medical Center, University of Cincinnati, Cincinnati, Ohio; [2]Medical College of Wisconsin, Milwaukee; [3]AstraZeneca, Gaithersburg, Maryland; [4]AstraZeneca, Gothenburg, Sweden; [5]AstraZeneca, Cambridge, United Kingdom. Dr Tummala's current address is AstraZeneca, Gaithersburg, Maryland.

Data are available on reasonable request. 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. Deidentified participant data can be made available on reasonable request to Catharina Lindholm or through the Vivli web-based data request platform. Reuse is permitted only with permission from AstraZeneca. The Clinical Study Protocol, Statistical Analysis Plan, and Clinical Study Report Synopsis are available at https://www.astrazenecaclinicaltrials.com/study/D3461C00007/ for TULIP-LN and at https://www.astrazenecaclinicaltrials.com/study/D3461C00005/ for TULIP-1.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25684).

Author disclosures and graphical abstract are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25684.

Address correspondence via email to Hermine I. Brunner, MD, at Hermine.brunner@cchmc.org.

Submitted for publication February 6, 2025; accepted in revised form October 7, 2025.

SIGNIFICANCE & INNOVATIONS

- Lupus nephritis (LN) confers significant morbidity and mortality to patients with systemic lupus erythematosus (SLE) in part because there is a lack of noninvasive, sensitive, and specific biomarkers to monitor renal activity in patients with LN. The Renal Activity Index for Lupus (RAIL) was developed as a noninvasive measure of LN activity based on the concentration of six urinary biomarkers (neutrophil gelatinase–associated lipocalin, kidney injury molecule 1, monocyte chemotactic protein 1, adiponectin, hemopexin, and ceruloplasmin) and was shown to accurately detect histologically active pediatric and adult patients with SLE and LN.
- We confirmed the high accuracy of the RAIL score to detect active LN in urine samples from adult patients with SLE and newly demonstrated the ability of the RAIL score to differentiate treatment response (agnostic of treatment type) over time in adult patients with LN.
- Thus, the RAIL allows for the noninvasive assessment of LN activity that could be readily used in the real-world clinical setting to optimize management of patients with LN.

There is a lack of noninvasive, accurate biomarkers to assess disease activity and evaluate response to treatment in patients with LN.[7,8] Baseline renal biopsies are the gold standard for LN diagnosis and are used to guide initial treatment decisions, based on the International Society of Nephrology/Renal Pathology Society (ISN/RPS) classification and the extent of kidney damage.[1,9,10] Repeat kidney biopsies are less established because they are invasive and costly, hence not feasible in routine surveillance as part of clinical practice or even in clinical trials.[1,11,12] Further, although repeated kidney biopsies are needed to fully understand the nature and extent of treatment impact on the kidneys, there is no consensus on the timing for postbaseline biopsies.[1,13–15]

Measurement of biomarkers in the urine may provide a noninvasive tool for estimating the degree of LN activity and the associated renal pathology in patients with LN, given that immune effectors mediating LN pathogenesis, including autoantibodies, complement, inflammatory cytokines, immune-mediators, and leukocytes, have been detected in urine samples from patients with LN.[12] Currently, the presence and degree of proteinuria is among the principal parameters used to clinically diagnose LN, gauge the degree of LN activity (or inflammation), and monitor treatment response.[1,16] However, proteinuria can come from pre-existing renal disease, hypertension, or renal damage, and some patients with marked kidney inflammation only have low-level proteinuria.[1,16] Likewise, other conventional SLE biomarkers, for example, complement proteins, and anti–double-stranded DNA (anti-dsDNA) antibodies, also cannot differentiate active renal disease from damage.[11,17]

Given the diversity of histologic findings on kidney biopsy, it is unlikely that a single biomarker measured in the urine suffices to estimate LN activity. Thus, the Renal Activity Index for Lupus (RAIL) was developed, whose scores are informed by the urine concentrations of six urine biomarkers.[18,19] RAIL biomarkers are neutrophil gelatinase–associated lipocalin (NGAL), monocyte chemotactic protein 1 (MCP-1), ceruloplasmin, adiponectin, hemopexin, and kidney injury molecule 1 (KIM-1).[18] Higher RAIL scores reflect higher active inflammation in LN, and changes in pediatric patients' RAIL scores reproduce the course of LN, that is, response to treatment and LN flare.[18,19] Two different RAIL algorithms have been developed, one for adults and one for pediatric patients with LN.[20] When used in pediatric patients, the pediatric RAIL scores were also able to predict National Institutes of Health Activity Index (NIH-AI) scores with >92% accuracy and Tubulointerstitial Activity Index scores with >80% accuracy.[18,20] When prospectively validated in an adult LN cohort, the RAIL has been found to accurately discriminate the level of LN activity as assessed by NIH-AI scores.[20] Furthermore, improvement in RAIL scores in pediatric patients with LN captured and even preceded clinical response of LN by at least three months with >90% accuracy.[21] However, the ability of the RAIL biomarkers to monitor therapeutic response over time has not yet been investigated in adults.

In the phase 3 Treatment of Uncontrolled Lupus via the Interferon Pathway (TULIP-1) multicenter trial (ClinicalTrials.gov identifier: NCT02446912), anifrolumab, a type I interferon receptor–targeting therapy, was well tolerated and provided therapeutic benefit in patients with moderate-to-severe SLE, excluding patients with severe, active LN.[22] This has led to anifrolumab subsequently receiving regulatory approval by the US Food and Drug Administration and the European Medicines Evaluation Agency, as well as agencies in other regions for moderate-to-severe SLE treatment.[23–27] In the phase 2 multicenter TULIP-LN study (ClinicalTrials.gov identifier: NCT02547922), anifrolumab was shown to have a consistent safety profile, and the intensified dosing regimen led to clinically meaningful improvements in LN end points in patients with SLE and active LN.[28]

In this study, we used data and samples from TULIP-1 and TULIP-LN to investigate the ability of the RAIL to discriminate patients requiring therapy for active LN from other patients with SLE and to differentiate LN treatment responders from nonresponders in adult patients, irrespective of LN treatment rendered. Deidentified participant data can be made available through the Vivli web-based data request platform.

MATERIALS AND METHODS

Patients. Full study design details, including patient inclusion and exclusion criteria for TULIP-1 and TULIP-LN, have been previously published.[22,28] In this analysis, the active LN group comprised 128 adult patients (18–70 years of age) from TULIP-

LN with biopsy-proven (biopsy within three months) active class III and IV LN, with or without coexistent class V LN overlap, according to the World Health Organization or ISN/RPS 2003 criteria.[10,28] Eligible patients had a urine protein–creatinine ratio (UPCR) from a 24-hour urine collection >1 mg/mg (113.17 mg/mmol), had an estimated glomerular filtration rate (eGFR) ≥35 mL/min/1.73 m^2, and fulfilled ≥4 of the 11 American College of Rheumatology SLE 1997 classification criteria.[28,29]

The comparator for this analysis (active SLE group) comprised 48 adult patients (18–68 years of age) from TULIP-1 with a diagnosis of moderate-to-severe active SLE despite standard therapy but without active kidney disease that required additional treatment, that is, a renal British Isles Lupus Assessment Group (BILAG) 2004 domain score of D or E.[22,29] We included this comparator group to study the ability of the RAIL to discriminate patients with active SLE from those with active LN.

As this was a secondary analysis of anonymized data, no ethics committee or institutional review board approvals were required. All such approvals were obtained in the original trials, which were conducted in accordance with the Declaration of Helsinki and the International Conference on Harmonisation Good Clinical Practice Guidelines[22,28]; independent ethics committee or independent institutional review board approvals were obtained, and all patients provided written informed consent.[22,28] Patients and/or the public were not involved in the design, conduct, reporting, or dissemination plans of this research.

Study design. TULIP-1 and TULIP-LN were both randomized, placebo-controlled, double-blind clinical trials of anifrolumab.[22,28] At baseline, in TULIP-1, patients received standard-of-care treatment, including glucocorticoids, antimalarials, azathioprine, mizoribine, mycophenolate mofetil [MMF] or mycophenolic acid, or methotrexate, for their active SLE. Race and ethnicity of the patients was based on patient self-report.

During TULIP-LN, patients received either placebo or anifrolumab every four weeks through week 48, in addition to standard-of-care treatment, that is, MMF and glucocorticoids.[22,28] Patients in TULIP-LN were required to discontinue treatment if they experienced predefined worsening of LN, defined as an LN-related confirmed eGFR decrease >30% from baseline to <60 mL/min/1.73 m^2 at any time, an eGFR decrease <75% from baseline to <60 mL/min/1.73 m^2 at week 12 or 24, or nephrotic range UPCR at week 12 or 24 (>3.5 mg/mg or <60% improvement in patients with UPCR >3 mg/mg at baseline).[28]

For the study, using a random sampling schedule developed by the study biostatistician, a subset of patients with active extrarenal SLE enrolled in TULIP-1 with available urine samples were included. Notably, this subset of patients from TULIP-1 included in this study are similar to the overall study population in TULIP-1 in terms of race, ethnicity, disease activity, and oral prednisone and MMF use.[22] Further, we included all patients participating in TULIP-LN for at least 24 weeks, irrespective of treatment allocation. Indeed, given the ongoing clinical trial program, this information could not be made available for this analysis. For all patients, we used urine samples obtained at baseline, and for the TULIP-LN participants, we also used urine samples collected at weeks 12 and 24.[22,28]

Course of renal disease during the study period. In the TULIP-LN study, complete renal response (CRR) was defined as a UPCR ≤0.7 mg/mg from a 24-hour urine collection, an eGFR ≥60 mL/min/1.73 m^2 or no decrease ≥20%, no treatment discontinuation, and no restricted medication use, including a mandatory oral glucocorticoid (prednisone) taper to a dosage of ≤15 mg/day by week 12 or <15 mg/day by week 24. Accordingly, for partial renal response (PRR), improvement in UPCR from a 24-hour urine collection to <1.0 mg/mg (if baseline UPCR ≤ 3.0 mg/mg) or to ≤3.0 mg/mg with >50% improvement (if baseline UPCR > 3 mg/mg) was required. The UPCR50 was another renal outcome and defined as a 50% reduction of proteinuria from baseline (all determined using a 24-hour urine collection). Although a 24-hour urine collection constitutes the gold standard for quantifying proteinuria, UPCR measurement from a random urine sample is considered an acceptable alternative and was available to measure proteinuria at all time points.[30]

Biomarker measurement. Spun urine samples were frozen within 24 hours from collection before testing in individual RAIL protein assays. RAIL urine biomarkers were quantified using single-plex assays per the manufacturers' instructions.

A Roche Cobas c 311 clinical chemistry analyzer, using a commercially available assay (BioPorto, catalog KIT ST001RA for NGAL; Roche Diagnostics, reference 03263991 190 for creatinine), was used to measure both human NGAL and urine creatinine concentrations. NGAL had a lower limit of detection (LLOD) of 9.8 ng/mL, and creatinine had an LLOD of 1.1 mg/dL.

Enzyme-linked immunosorbent assays were used to measure the remaining biomarkers; all sample testing was done in singlicate. Human urinary KIM-1 (R&D Systems, DKM100) had an LLOD of 0.009 ng/mL. Human MCP-1 (R&D Systems, DCP00) had an LLOD of 1.7 pg/mL. Human adiponectin (R&D Systems, DRP300) had an LLOD of 0.246 ng/mL. Human ceruloplasmin (Assaypro LLC, EC4201-1) had an LLOD of 0.085 ng/mL. Human hemopexin (Assaypro LLC, EH2001-1) had an LLOD of 4.2 ng/mL.

KIM-1 and MCP-1 used a four-parameter logistic curve to fit the standard curve, and adiponectin, ceruloplasmin, and hemopexin instead used a log/log curve. Analyte concentrations that were above or below the limits of detection were imputed by 50% of the LLOD and 50% over the upper limit of detection, respectively. Natural log-transformed quantities of NGAL, adiponectin, hemopexin, and ceruloplasmin are reported in ng/mL, whereas those of KIM-1 and MCP-1 are reported in pg/mL.

RAIL score derivation for use in adult populations. We have shown in the past that the RAIL biomarkers are produced locally in the kidney with active LN[31] and that standardization of biomarker quantities by urine creatinine concentration but not unspecific proteinuria results in the best reflection of histologic activity of LN as per renal biopsies (NIH-AI score) in both children and adults with LN.[18,20,31] When using RAIL scores, the age of the patient must be considered. RAIL scores were developed in pediatric patients with SLE and then validated in adult patients previously and in our study.[18,20] NGAL, KIM-1, and MCP-1 urine concentrations have been shown to increase with age, whereas adiponectin concentrations decrease with age.[11,32] Significantly higher concentrations of NGAL, hemopexin, and ceruloplasmin have been found in urine samples from female patients compared with male patients.[11,32] Further, eGFR declines with age. Given these differences, we have developed two RAIL algorithms for use in pediatric and adult populations.[33]

As previously published, the RAIL score for adults was calculated from the urine quantities of the RAIL biomarkers, after natural logarithmic transformation of values and standardization by urine creatinine concentrations, as follows: RAIL = –0.21 + 0.67 × NGAL + 0.28 × MCP1 – 0.12 × ceruloplasmin + 0.88 × adiponectin – 0.25 × hemopexin – 0.05 × KIM-1.[20] Herein, the RAIL had 80% accuracy to detect active LN as defined by an NIH-AI score of >10, and higher scores reflected higher LN activity.[20,33]

Statistical analyses and predictive modeling. Data were pooled agnostic of treatment received in TULIP-LN or TULIP-1. Using the RAIL algorithm for adults, RAIL scores presented are calculated from the urinary biomarker concentrations, standardized by urine creatinine concentration, as previously published.[20] Clinical characteristic and urinary biomarker comparisons at baseline between the active LN group and the active SLE group were performed using the Wilcoxon rank sum test. To assess the accuracy of the RAIL score, receiver operating characteristic (ROC) curve analyses were conducted to assess the ability of RAIL scores to distinguish the active LN group from the active SLE group, and we calculated the area under the curve (AUC; range 0–1). In general, values of the AUC are considered as outstanding, excellent, good, fair, or poor for AUCs of 0.90 to 1.00, 0.81 to 0.90, 0.71 to 0.80, 0.61 to 0.70, or 0.51 to 0.60, respectively. For patients from the TULIP-LN study, we compared RAIL scores from responders versus nonresponders at baseline and then again at week 12 and week 24. Three definitions of LN response were considered, ie, CRR, PRR, and UPCR50. Box and whisker plots were used to present the RAIL values at baseline, 12 weeks, and 24 weeks, and groups were compared using the Wilcoxon rank sum test. Further, generalized estimating equation (GEE) analyses were performed to evaluate the role of RAIL in predicting response status, adjusted for baseline UPCR, GFR, sex, race, disease duration, and oral prednisone and MMF usage.

RESULTS

Patients. This analysis included 128 patients with active LN (active LN group) and 48 patients with active SLE without clinically meaningful renal activity, that is, this SLE group consisted of 48 patients with a BILAG renal domain score of D/E (n = 11; D = previous involvement with no current activity, n = 25; or E = no previous involvement, n = 23)[22,29] and a (clinical) Systemic Lupus Erythematosus Disease Activity Index (SLEDAI) score of at least 4. No information about the class of LN or time since kidney biopsy was available for patients in the active SLE group with a BILAG renal domain score of D. Patient and baseline disease characteristics of the active LN group and the active SLE group are presented in Table 1.

Table 1. Patient demographics and disease characteristics*

Characteristics	Active LN group (n = 128)	Active SLE group (n = 48)
Age, y		
Median (IQR), IU	34 (25–42)	37 (29–48)
Mean ± SD	34.2 ± 10.7	38.7 ± 12.9
Female, n (%)	109 (83.2)	45 (95.7)
Race, n (%)		
White	56 (42.7)	36 (76.6)
Black/African American	7 (5.3)	5 (10.6)
Asian	27 (20.6)	4 (8.5)
Other	41 (31.3)	2 (4.3)
Hispanic or Latino ethnicity, n (%)	61 (46.6)	8 (17)
SLEDAI-2K score, median (IQR), IU		
Nonrenal	4 (4–6)	12 (8–14)
Renal	4 (4–8)	0 (0–0)
UPCR, mg/mg		
Median (IQR), IU	2.1 (1.2–4.0)	0.08 (0.05–0.15)
Mean ± SD	3.1 ± 2.7	0.12 ± 0.12
eGFR, median (IQR), mL/min/1.73 m^2	91.8 (63.1–125)	98.9 (85.3–120.3)
C3, median (IQR), g/L	0.8 (0.6–1)	0.94 (0.76–1.17)
C4, median (IQR), g/L	0.15 (0.1–0.2)	0.15 (0.1–0.23)
Increased anti-dsDNA antibodies, median (IQR), IU	74.2 (21.9–255.4)	13.6 (3.2–50.8)
Baseline SLE treatments		
Oral glucocorticoids, n (%)	127 (97.7)	43 (89.6)
Prednisone, median (IQR), mg/day	20 (15–30)	10 (5–15)
MMF (before randomization), n (%)[a]	95 (72.5)	11 (22.9)
MMF, median (IQR), g/day	2.0 (1.5–2.0)	2.0 (1.0–2.0)

* Data presented as n (% of N), unless otherwise specified. anti-dsDNA, anti–double-stranded DNA; eGFR, estimated glomerular filtration rate; IQR, interquartile range; LN, lupus nephritis; MMF, mycophenolate mofetil; SLE, systemic lupus erythematosus; SLEDAI-2K, Systemic Lupus Erythematosus Disease Activity Index 2000; UPCR, urine protein–creatinine ratio.
[a] All patients started MMF upon randomization as part of the study protocol.

At baseline, out of those with available LN class information, almost two-thirds of the patients in the active LN group had class IV (76/118 = 64.41%), approximately one-quarter had class III (26/118 = 22.03%), and the remainder of the active LN group showed an overlap of either class III/V or class IV/V (16/118 = 13.56%). Further, in the active LN group, the median age and the proportion of female patients were lower, whereas the proportion of Hispanic or Latino patients (46.56% vs 15.25%) was greater compared with the active SLE group. As expected from the study eligibility criteria, the median renal SLEDAI-2K score and the median proteinuria level were higher and the median eGFR was lower in the active LN group compared with the active SLE group. In the active SLE group, proteinuria was similar in patients with a baseline BILAG renal domain score of D versus those with a BILAG renal domain score of E (mean ± SD of BILAG renal domain score D 0.14 ± 0.13 vs E 0.11 ± 0.11; $P = 0.34$). Also expected was the more frequent use of MMF in the active LN group compared with the active SLE group, as all patients in TULIP-LN were required to take MMF as background standard of care from randomization onward.

RAIL biomarker levels and RAIL score at baseline. At baseline, individual RAIL biomarker levels were all significantly higher (all $P \leq 0.02$) in the urine samples from the active LN group compared with the active SLE group (Table 2). Urine creatinine concentration did not significantly differ between groups ($P = 0.05$). Median (interquartile range [IQR]) RAIL scores were significantly higher in the active LN group compared with the active SLE group (5.59 [IQR 4.31–6.47] vs 3.57 [IQR 2.78–4.47]; $P < 0.001$).

ROC analyses of RAIL performance to identify active LN requiring treatment. ROC analyses were conducted to evaluate the ability of the RAIL to identify active LN requiring treatment. When discriminating the active LN group from the active SLE group based on RAIL scores in ROC analysis, the AUC was 0.8012 (95% confidence interval [CI] 0.733–0.869), with an odds ratio (OR) of 2.0 (95% CI 1.5–2.6) (Figure 1). A RAIL score of 4.25 (statistically optimal threshold value) had a sensitivity of 76% and a specificity of 74% to discriminate between the active LN group and the active SLE group. After adjusting for baseline differences in MMF use, UPCR, anti-dsDNA antibodies, and race and ethnicity between groups, the RAIL score was outstanding in differentiating between the active LN group and the active SLE group (AUC 0.997, 95% CI 0.993–1). We also explored the ability of the anti-dsDNA antibodies and UPCR to differentiate between groups and determined that the AUCs were 0.707 (95% CI 0.622–0.792)and 0.989 (95% CI 0.975–1),

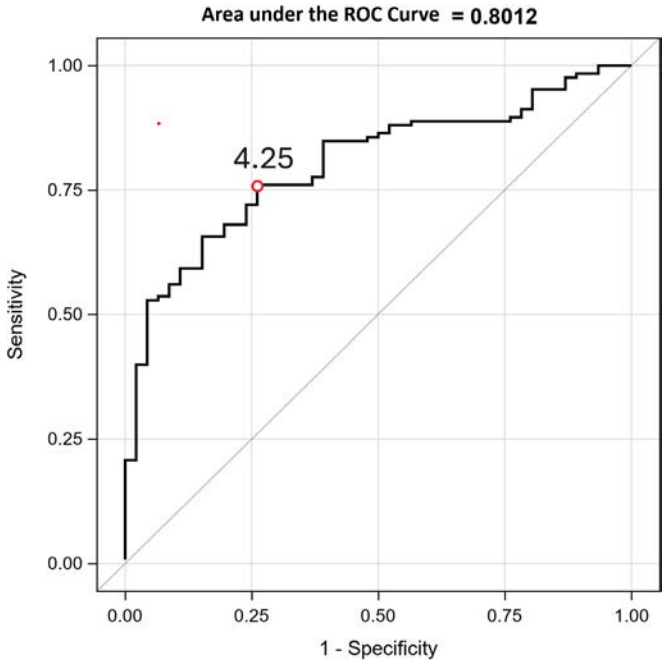


Figure 1. ROC curve of creatinine-adjusted RAIL scores in both the active LN group and the active SLE group. The statistically optimal threshold value to differentiate groups is represented by a diamond and indicates a sensitivity of 76% and a specificity of 74% based on a RAIL score of 4.25. LN, lupus nephritis; RAIL, Renal Activity Index for Lupus, based on the algorithm for adults and standardized by urine creatinine concentration; ROC, receiver operator characteristic; SLE, systemic lupus erythematosus.

Table 2. Baseline urine biomarkers*

Characteristics	Active LN group (n = 128)	Active SLE group (n = 48)	*P* value[a]
NGAL, ng/mL	33.3 (17.6–56.7)	19.8 (11.4–44.0)	0.02
MCP-1, pg/mL	658.2 (271.6–1,050.0)	234.3 (98.0–420.1)	<0.001
Ceruloplasmin, ng/mL	93.6 (44.5–311.3)	31.3 (9.35–71.2)	<0.001
Adiponectin, ng/mL	42.5 (16.7–139.6)	8.6 (1.9–22.3)	<0.001
Hemopexin, ng/mL	1,876.8 (745.1–4,743.4)	428.9 (189.0–941.5)	<0.001
KIM-1, pg/mL	1,673.5 (772.5–2,767)	864 (312–1,392)	<0.001
Urine creatinine, mg/mL	0.7 (0.5–1.3)	0.96 (0.5–1.7)	0.05
RAIL score[b]	5.59 (4.31–6.47)	3.57 (2.78–4.47)	<0.001

* Data are presented as median (IQR). IQR, interquartile range; KIM, kidney injury molecule; LN, lupus nephritis; MCP, monocyte chemotactic protein; NGAL, neutrophil gelatinase–associated lipocalin; RAIL, Renal Activity Index for Lupus; SLE, systemic lupus erythematosus.
[a] Nominal *P* values, Wilcoxon rank sum test.
[b] Scores are based on the algorithm for adults and standardized by urine creatinine concentration.

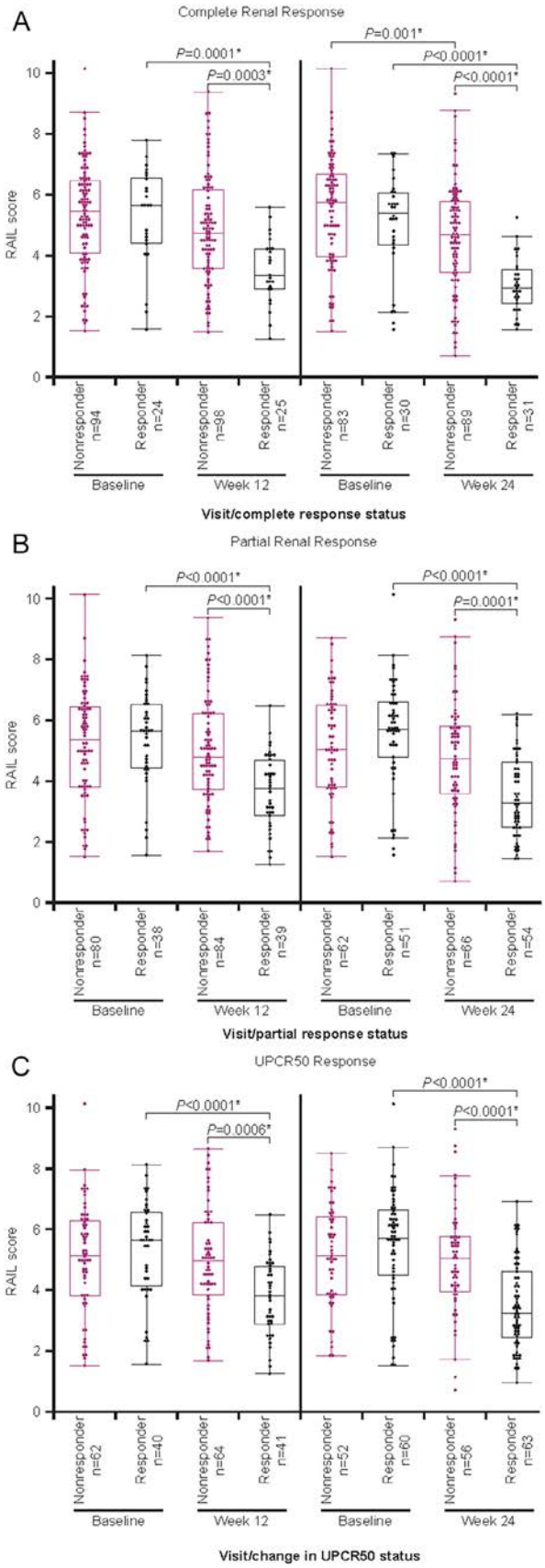


Figure 2. Legend on next column.

respectively, suggesting a fair ability of anti-dsDNA antibody levels but, as expected, an outstanding ability of proteinuria to discriminate groups (see Supplemental Figure S1).

Association of RAIL scores with baseline activity and longitudinal renal response. Renal response was studied using urine samples collected at baseline, week 12, and week 24 of the treatment period. There were 25, 39, and 41 patients who achieved CRR, PRR, and UPCR50 at week 12 and 31, 54, and 63 patients with CRR, PRR, and UPCR50 at week 24, respectively. There were no significant differences in baseline RAIL scores between subsequent responders (CRR, PRR, UPCR50) and nonresponders at weeks 12 and 24. Figure 2A compares side-by-side baseline RAIL scores and the RAIL scores of patients achieving CRR at week 12 (or week 24). Using a Wilcoxon rank sum test, RAIL scores were similar between groups at baseline, but there were significant differences in the RAIL scores between patients achieving CRR versus nonresponders at week 12 ($P < 0.0003$), and the same held true for week 24 ($P < 0.0001$). Indeed, compared with baseline, patients with CRR had a median decrease in RAIL score of −1.3 (IQR −3.64 to −0.21) at week 12 and −2.30 (IQR −3.63 to −1.03) at week 24. Figure 2B shows significant differences in RAIL scores between those who achieved PRR (or even CRR) and nonresponders at week 12 ($P < 0.0001$) and week 24 ($P < 0.0001$). Again, there were no differences in RAIL scores at baseline in patients achieving PRR versus those categorized as nonresponders at week 12 or week 24. As shown in Figure 2C, when considering the response measure UPCR50, RAIL scores of those achieving UPCR50 versus nonresponders significantly differed at week 12 ($P < 0.0006$) and at week 24 ($P < 0.0001$); again, RAIL scores of groups (PRR or nonresponders) were similar at baseline. There was a significant change in RAIL scores for those achieving PRR and UPCR50 between baseline and weeks 12 and 24 that was not observed in the nonresponder groups.

Figure 2. RAIL score changes over time based on renal response. Cross-sectional comparison of the RAIL score at baseline and at week 12 or week 24 is depicted for patients considered at these time points to have one of the evaluated renal outcomes. (A) CRR. (B) At least PRR. (C) UPCR50, ie, decrease of proteinuria of >50% from baseline. *Nominal *P* values from the Wilcoxon rank sum test. CRR was defined as 24-hour UPCR ≤0.7 mg/mg, eGFR ≥60 mL/min/1.73 m^2 or no decrease ≥20%, no treatment discontinuation, and no restricted medication use. PRR was defined as improvement in 24-hour UPCR to <1.0 mg/mg (if baseline UPCR ≤ 3 mg/mg) or to ≤3.0 mg/mg with >50% improvement (if baseline UPCR > 3 mg/mg). CRR, complete renal response; eGFR, estimated glomerular filtration rate; PRR, partial renal response; RAIL, Renal Activity Index for Lupus, based on the algorithm for adults and standardized by urine creatinine concentration; UPCR, urine protein–creatinine ratio; UPCR50, UPCR decrease >50% from baseline.

Table 3 summarizes the absolute change in UPCR, eGFR, and SLEDAI-2K over time in the active LN group for those achieving CRR at weeks 12 and 24 versus nonresponders. Notably, although nonresponders (those not achieving CRR) and those achieving CRR at week 24 had similar changes in UPCR and eGFR (both $P \leq 0.20$), changes in RAIL scores were significantly larger in those achieving CRR compared to nonresponders ($P = 0.004$). Similar observations were made for CRR at week 12.

Supplemental Table 1 provides the information for PRR, again suggesting a significant decrease in the RAIL scores at weeks 12 and 24 in those with PRR compared to nonresponders. For the UPCR50 outcome, comparable results were observed (data not shown). The decrease in RAIL scores was less substantial for those with PRR or UPCR50 compared to CRR at week 24, but differences in RAIL scores between those achieving PRR as compared to CRR did not reach statistical significance. In sensitivity analyses, we removed the restricted medication usage criterion when determining the course of LN over time, as is summarized for achieving CRR in Supplemental Table 2, no important differences compared with the primary analysis were observed.

Using GEE modeling, we further explored the ability of the RAIL score to predict response status at weeks 12 and 24 when also considering patient demographics, medication usage (prednisone, MMF), and LN parameters. A low baseline RAIL score ($P = 0.0007$) and a large decrease in the RAIL score from baseline over time ($P < 0.0001$) as well as baseline low-level proteinuria ($P = 0.0014$) and female sex ($P = 0.031$) were all statistically significant predictors of achieving CRR. For reaching PRR, a low baseline RAIL score ($P = 0.0473$) and a large decrease in the RAIL score from baseline ($P < 0.0001$) remained statistically significant predictors, whereas sex and baseline low-level proteinuria were not. Predictors of UPCR50 status were a low baseline RAIL score ($P = 0.0033$), a large decrease in the RAIL score from baseline ($P < 0.0001$), and baseline low-level proteinuria ($P = 0.0247$). Conversely, GFR at baseline, race, disease duration, and oral prednisone or MMF dosage at baseline were all not important predictors of response status. Additional details are provided in Supplemental Table 3.

Table 3. Change from baseline of proteinuria, eGFR, disease activity, and RAIL scores in the active LN group during the study for patients achieving CRR or nonresponders*,[b]

Outcome	Week 12	Week 24
Spot UPCR, mg/mg		
CRR	−1.0 (−1.67 to −0.46)	−1.18 (−1.94 to −0.62)
Nonresponder	−0.61 (−1.7 to 0.96)	−0.81 (−2.86 to 0.54)
P-value	0.03	0.20
eGFR, mL/min/1.73 m^2		
CRR	0.0 (−15.1 to 9.1)	0.0 (−1.2 to 10)
Nonresponder	0.0 (−9 to 17.7)	0.0 (−15.1 to 21.9)
P-value	0.29	0.47
SLEDAI-2K score		
Nonrenal		
CRR	−2.0 (−2.0 to 0.0)	0.0 (−2.0 to 0.0)
Nonresponder	0.0 (−2 to 0.0)	−2.0 (−3.3 to 0.0)
P-value	0.08	0.60
Renal		
CRR	−4.0 (−4.0 to 0)	−4.0 (−8 to 0.0)
Nonresponder	0.0 (−4.0 to 0.0)	0.0 (−4 to 0.0)
P-value	0.02	≤0.001
RAIL score[a]		
CRR	−1.3 (−3.64 to −0.21)	−2.30 (−3.63 to −1.03)
Nonresponder	−0.39 (−1.6 to 0.83)	−0.88 (−2.20 to 0.33)
P-value	0.01	0.004

* Medians and interquartile ranges are shown, and patients with partial renal response were excluded from the comparison. CRR, complete renal remission; eGFR, estimated glomerular filtration rate; LN, lupus nephritis; RAIL, Renal Activity Index for Lupus; SLEDAI-2K, Systemic Lupus Erythematosus Disease Activity Index 2000; UPCR, urine protein–creatinine ratio.
[a] Scores are based on the algorithm for adults and standardized by urine creatinine concentration.
[b] Nominal *P* values, Wilcoxon rank sum test.

DISCUSSION

Despite therapeutic advances in SLE management, LN remains a common severe disease manifestation, conferring significant morbidity and mortality.[1,3–5] The RAIL, based on a set of urine biomarkers, has been developed as a noninvasive tool to help diagnose LN and monitor the course of pediatric LN.[18,20] Here, we newly demonstrate the ability of individual RAIL biomarkers and the RAIL score, derived from the urine levels of all six RAIL biomarkers corrected for urine creatinine concentration, to differentiate adult patients requiring treatment for active LN from other patients with active extrarenal SLE. Further, we newly show that absolute change in RAIL scores differentiates LN responders from nonresponders in the TULIP-LN study, irrespective of the clinical definition of LN response considered (CRR, PRR, UPCR50). Collectively, these results suggest the RAIL is an effective, noninvasive tool for assessing LN activity and treatment response over time in adult patients with LN.

In this analysis, a RAIL score of 4.25 or higher differentiated patients with active LN from those with SLE with absent or inactive LN (active SLE group) with 80% accuracy, consistent with our earlier findings.[18,20] This observation could be exploited in clinical care for the surveillance of LN, especially in patients with known proteinuria from kidney damage. This is because we have shown in the past that the RAIL biomarkers do not reflect kidney damage.[19,20]

We conducted an exploratory analysis in which 12 patients with BILAG C renal domain scores were added to the active SLE group. Notably, this exploratory analysis showed that a RAIL score of 4.56 or higher discriminated the active LN group from the adapted active SLE group (with baseline BILAG renal domain scores of C/D/E) also with an accuracy of 80% (AUC 0.8004 [95% CI 0.740–0.867]), confirming the robustness of the RAIL score to discriminate active LN requiring treatment (BILAG renal domain

score of A or B) from absent or only minimally active LN (BILAG renal domain score of C/D/E).

Change in RAIL scores over time, rather than RAIL scores at baseline, differentiated LN responders from nonresponders at weeks 12 and 24 when other patient and disease characteristics were not considered in the analyses. This observation is congruent with a study in pediatric patients with LN in which a significant decrease in pediatric RAIL scores, also standardized by urine creatinine concentration, by at least 1.1 was associated with a complete response to LN induction therapy, irrespective of the type of immunosuppressive therapy rendered.[34]

We were unable to show that, in isolation, the RAIL score at baseline can forecast LN response at weeks 12 and week 24 (data not shown). However, RAIL score changes from baseline, either alone or together with the baseline RAIL score, were a strong predictor of LN responses. This corroborates our prior observations in pediatric patients with LN in which we reported that changes in RAIL scores anticipated the clinically observed course at least three months earlier.[21] After adjustment for patient and baseline disease characteristics (GEE models), both the baseline RAIL score and the change in RAIL score remain important predictors of LN response. These findings support that the RAIL score might be an important predictor of future LN response.

RAIL scores decreased from baseline during the LN study even in treatment nonresponders; however, decreases in scores only reached significance at week 24 for the CRR analysis. Some decreases in RAIL scores in nonresponders might be expected, as disease activity and the UPCR decreased in the patients whose treatment included MMF and often steroids during TULIP-LN[28]; despite this, improvement in RAIL scores was much smaller in nonresponders than in responders.

We did not evaluate the effects of treatment types on biomarker levels. However, prior observations in children that suggested accurate predictive outcomes were possible with RAIL scores regardless of treatment type, including the use of antihypertensive drugs that target the angiotensinogen system, steroids, cyclophosphamide, or MMF.[11,18,19,21,34]

Proteinuria and eGFR are included in current LN composite response measures, including those of TULIP-LN (CRR, PRR, UPCR50). UPCR levels only significantly differed in patients with CRR at week 12, but not week 24, and in patients with PRR at both time points. This could be due to the small sample size or reflect the dual role of proteinuria as a measure of both LN activity and damage. Given the lack of available follow-up kidney biopsies, we were unable to further evaluate this observation.

The short coming of proteinuria to gauge LN activity has long been recognized.[35] For example, proteinuria can reflect kidney damage or even comorbid disease, such as hypertension and diabetes mellitus.[36,37] Indeed, in this study, the change in UPCR in the active LN group at week 12 was similar in those categorized as having PRR or even CRR, further supporting the need for a more robust biomarker of LN activity besides proteinuria.

Although we consider the availability of kidney biopsy samples to confirm active LN at baseline in the active LN group as a strength of this study,[22,28] we were unable to confirm response to therapy by serial kidney biopsies because the study protocol did not mandate repeat biopsies. However, serial kidney biopsies for the sole purpose of research or a clinical trial may be ethically unacceptable. Further, additional validation studies are needed to expand on the interpretation of RAIL scores over time in adults, best with available serial kidney biopsies to estimate response to therapy.

In summary, these findings support the clinical utility of the RAIL biomarkers and the RAIL score to noninvasively assess kidney inflammation and response to therapy in adult LN, thus allowing a personalized monitoring approach. Future research is needed to evaluate the potential benefits of the RAIL compared to proteinuria histologic response of LN.

ACKNOWLEDGMENTS

The authors would like to thank the investigators, research staff, health care providers, patients, and caregivers who contributed to this study and the clinical trials on which its analyses are based. The authors would also like to thank Angela Merritt, Marc Sudman, Tingting Qiu, and James Rose for their assistance in research regulatory requirements, samples storage, initial statistics programming support, and biomarker assays, respectively.

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

Editorial support was provided by Jack Barton, PhD, of JK Associates Inc, part of Avalere Health. This support was funded by AstraZeneca. AstraZeneca also contributed to the statistical analysis and was a donor of the urine samples. AstraZeneca contributed to the study design but did not influence the decision to submit the manuscript for publication. Publication of this article was not contingent upon approval by AstraZeneca.

REFERENCES

1. Anders HJ, Saxena R, Zhao MH, et al. Lupus nephritis. Nat Rev Dis Primers 2020;6(1):7. doi:https://doi.org/10.1038/s41572-019-0141-9
2. Yurkovich M, Vostretsova K, Chen W, et al. Overall and cause-specific mortality in patients with systemic lupus erythematosus: a meta-analysis of observational studies. Arthritis Care Res (Hoboken) 2014; 66(4):608–616. doi:https://doi.org/10.1002/acr.22173

3. Hanly JG, O'Keeffe AG, Su L, et al. The frequency and outcome of lupus nephritis: results from an international inception cohort study. Rheumatology (Oxford) 2016;55(2):252–262. doi:https://doi.org/10.1093/rheumatology/kev311
4. Hoover PJ, Costenbader KH. Insights into the epidemiology and management of lupus nephritis from the US rheumatologist's perspective. Kidney Int 2016;90(3):487–492. doi:https://doi.org/10.1016/j.kint.2016.03.042
5. Pryor KP, Barbhaiya M, Costenbader KH, et al. Disparities in lupus and lupus nephritis care and outcomes among US Medicaid beneficiaries. Rheum Dis Clin North Am 2021;47(1):41–53. doi:https://doi.org/10.1016/j.rdc.2020.09.004
6. Suzon B, Louis-Sidney F, Aglaé C, et al. Good long-term prognosis of lupus nephritis in the high-income Afro-Caribbean population of Martinique with free access to healthcare. J Clin Med 2022;11(16):4860. doi:https://doi.org/10.3390/jcm11164860
7. Bertsias GK, Tektonidou M, Amoura Z, et al; European League Against Rheumatism and European Renal Association-European Dialysis and Transplant Association. Joint European League Against Rheumatism and European Renal Association-European Dialysis and Transplant Association (EULAR/ERA-EDTA) recommendations for the management of adult and paediatric lupus nephritis. Ann Rheum Dis 2012;71(11):1771–1782. doi:https://doi.org/10.1136/annrheumdis-2012-201940
8. Davidson A, Aranow C, Mackay M. Lupus nephritis: challenges and progress. Curr Opin Rheumatol 2019;31(6):682–688. doi:https://doi.org/10.1097/BOR.0000000000000642
9. Bajema IM, Wilhelmus S, Alpers CE, et al. Revision of the International Society of Nephrology/Renal Pathology Society classification for lupus nephritis: clarification of definitions, and modified National Institutes of Health activity and chronicity indices. Kidney Int 2018;93(4):789–796. doi:https://doi.org/10.1016/j.kint.2017.11.023
10. Weening JJ, D'Agati VD, Schwartz MM, et al; International Society of Nephrology Working Group on the Classification of Lupus Nephritis; Renal Pathology Society Working Group on the Classification of Lupus Nephritis. The classification of glomerulonephritis in systemic lupus erythematosus revisited. Kidney Int 2004;65(2):521–530. doi:https://doi.org/10.1111/j.1523-1755.2004.00443.x
11. Bennett MR, Ma Q, Ying J, et al. Effects of age and gender on reference levels of biomarkers comprising the pediatric Renal Activity Index for Lupus Nephritis (p-RAIL). Pediatr Rheumatol Online J 2017;15(1):74. doi:https://doi.org/10.1186/s12969-017-0202-0
12. Morell M, Pérez-Cózar F, Marañón C. Immune-related urine biomarkers for the diagnosis of lupus nephritis. Int J Mol Sci 2021;22(13):7143. doi:https://doi.org/10.3390/ijms22137143
13. De Rosa M, Azzato F, Toblli JE, et al. A prospective observational cohort study highlights kidney biopsy findings of lupus nephritis patients in remission who flare following withdrawal of maintenance therapy. Kidney Int 2018;94(4):788–794. doi:https://doi.org/10.1016/j.kint.2018.05.021
14. Malvar A, Pirruccio P, Alberton V, et al. Histologic versus clinical remission in proliferative lupus nephritis. Nephrol Dial Transplant 2017;32(8):1338–1344. doi:https://doi.org/10.1093/ndt/gfv296
15. Zickert A, Sundelin B, Svenungsson E, et al. Role of early repeated renal biopsies in lupus nephritis. Lupus Sci Med 2014;1(1):e000018. doi:https://doi.org/10.1136/lupus-2014-000018
16. Chedid A, Rossi GM, Peyronel F, et al. Low-level proteinuria in systemic lupus erythematosus. Kidney Int Rep 2020;5(12):2333–2340. doi:https://doi.org/10.1016/j.ekir.2020.09.007
17. Soliman S, Mohan C. Lupus nephritis biomarkers. Clin Immunol 2017;185:10–20. doi:https://doi.org/10.1016/j.clim.2016.08.001
18. Brunner HI, Bennett MR, Abulaban K, et al. Development of a novel renal activity index of lupus nephritis in children and young adults. Arthritis Care Res (Hoboken) 2016;68(7):1003–1011. doi:https://doi.org/10.1002/acr.22762
19. Brunner HI, Bennett MR, Mina R, et al. Association of noninvasively measured renal protein biomarkers with histologic features of lupus nephritis. Arthritis Rheum 2012;64(8):2687–2697. doi:https://doi.org/10.1002/art.34426
20. 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. doi:https://doi.org/10.1177/0961203316684212
21. Brunner HI, Bennett MR, Gulati G, et al. Urine biomarkers to predict response to lupus nephritis therapy in children and young adults. J Rheumatol 2017;44(8):1239–1248. doi:https://doi.org/10.3899/jrheum.161128
22. Furie RA, Morand EF, Bruce IN, et al; TULIP-1 study investigators. Type I interferon inhibitor anifrolumab in active systemic lupus erythematosus (TULIP-1): a randomised, controlled, phase 3 trial. Lancet Rheumatol 2019;1(4):e208–e219. doi:https://doi.org/10.1016/S2665-9913(19)30076-1
23. Burki TK. FDA approval for anifrolumab in patients with lupus. Lancet Rheumatol 2021;3(10):e689. doi:https://doi.org/10.1016/S2665-9913(21)00291-5
24. Rogers H, Adeniyi O, Ramamoorthy A, et al. Clinical pharmacology studies supporting oligonucleotide therapy development: an assessment of therapies approved and in development between 2012 and 2018. Clin Transl Sci 2021;14(2):468–475. doi:https://doi.org/10.1111/cts.12945
25. Shadid M, Badawi M, Abulrob A. Antisense oligonucleotides: absorption, distribution, metabolism, and excretion. Expert Opin Drug Metab Toxicol 2021;17(11):1281–1292. doi:https://doi.org/10.1080/17425255.2021.1992382
26. European Medicines Agency (EMA). Guideline on the investigation of drug interactions. Accessed March 14, 2024. https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-investigation-drug-interactions-revision-1_en.pdf
27. Food and Drug Administration (FDA). In vitro drug interaction studies - cytochrome P450 enzyme- and transporter-mediated drug interactions. guidance for industry. 2020. Accessed January 22, 2026. https://www.fda.gov/media/135587/download
28. Jayne D, Rovin B, Mysler EF, et al. Phase II randomised trial of type I interferon inhibitor anifrolumab in patients with active lupus nephritis. Ann Rheum Dis 2022;81(4):496–506. doi:https://doi.org/10.1136/annrheumdis-2021-221478
29. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1997;40(9):1725. doi:https://doi.org/10.1002/art.1780400928
30. Sammaritano LR, Askanase A, Bermas BL, et al. 2024 American College of Rheumatology (ACR) guideline for the screening, treatment, and management of lupus nephritis. Arthritis Rheumatol 2025;77(9):1115–1135. doi:https://doi.org/10.1002/art.43212
31. Suzuki M, Wiers K, Brooks EB, et al. Initial validation of a novel protein biomarker panel for active pediatric lupus nephritis. Pediatr Res 2009;65(5):530–536. doi:https://doi.org/10.1203/PDR.0b013e31819e4305
32. Aljaberi N, Bennett M, Brunner HI, et al. Proteomic profiling of urine: implications for lupus nephritis. Expert Rev Proteomics 2019;16(4):303–313. doi:https://doi.org/10.1080/14789450.2019.1592681
33. Austin HA III, Muenz LR, Joyce KM, et al. Diffuse proliferative lupus nephritis: identification of specific pathologic features affecting renal outcome. Kidney Int 1984;25(4):689–695. doi:https://doi.org/10.1038/ki.1984.75
34. Cody EM, Wenderfer SE, Sullivan KE, et al. Urine biomarker score captures response to induction therapy with lupus nephritis. Pediatr

Nephrol 2023;38(8):2679–2688. doi:https://doi.org/10.1007/s00467-023-05888-z.

35. Weeding E, Fava A, Magder L, et al. One-third of patients with lupus nephritis classified as complete responders continue to accrue progressive renal damage despite resolution of proteinuria. Lupus Sci Med 2022;9(1):e000684. doi:https://doi.org/10.1136/lupus-2022-000684
36. Kochi M, Kohagura K, Oshiro N, et al. Association of blood pressure and hyperuricemia with proteinuria and reduced renal function in the general population. Hypertens Res 2023;46(7):1662–1672. doi: https://doi.org/10.1038/s41440-023-01250-w
37. Khoury CC, Chen S, Ziyadeh FN. Pathophysiology of diabetic nephropathy. In: Kimmel PL, Rosenberg ME, eds. Chronic Renal Disease. 2nd ed. Academic Press; 2020:279–296.

Arthritis Care & Research
Vol. 78, No. 6, June 2026, pp 733–742
DOI 10.1002/acr.25686

American College of Rheumatology
Empowering Rheumatology Professionals

Polypharmacy in Adults With Systemic Lupus Erythematosus

Jianing Yang,[1] Matthew E. Growdon,[2] Jessica Fitzpatrick,[2] Charmayne Dunlop-Thomas,[3] Courtney Hoge,[3] Kenneth E. Covinsky,[2] C. Barrett Bowling,[4,5] S. Sam Lim,[3] Jinoos Yazdany,[2] and Laura C. Plantinga[2]

Objective. Estimates of polypharmacy among US adults with systemic lupus erythematosus (SLE)—a relatively young and disproportionately minoritized population—remain sparse. We sought to estimate the prevalence of polypharmacy in SLE and identify the most common medications used.

Methods. For this cross-sectional study, participants were recruited from a population-based cohort of adults with validated SLE in Atlanta, Georgia. Prescription and over-the-counter (OTC) medications were self-reported at the study visit. Polypharmacy was defined as five or more prescription or OTC medications. Estimates of polypharmacy prevalence by key sociodemographic and SLE-related participant characteristics were obtained using crude logistic regression and postestimation marginals.

Results. More than half (56.3%) of participants (n = 451; 15.3% ≥60 years old, 91.8% women, and 81.8% Black) reported polypharmacy. Older age (68.1%, 59.8%, and 43.0% for ages ≥60 years, 40–59 years, and 18–39 years), higher vs lower disease activity (65.8% vs 46.2%) and cumulative SLE-related damage (68.5% vs 42.4%), longer disease duration (62.4% vs 50.0%), and taking three to five vs zero to one immunomodulating medications (79.6% vs 38.0%) were associated with higher age-adjusted prevalence of polypharmacy; prevalence was not statistically significantly different by sex, race, or education. Although hydroxychloroquine (71.4%), glucocorticoids (44.3%), and other immunomodulating drugs (50.3%) were common, polypharmacy was most often driven by other medications, such as antihypertensives (61.9%), nonopioid pain relievers (51.7%), allergy treatments (22.4%), antidepressants (22.2%), and gastric reflux medications (21.7%).

Conclusion. Our results underscore the need to address the burden of medication regimens in this population through individualized medication optimization strategies that account for prescription and OTC medications used by those with SLE.

INTRODUCTION

Polypharmacy, referring to the simultaneous use of multiple medications, is common among older adults[1,2] and those with chronic diseases such as chronic kidney disease,[3] cancer,[4] and diabetes.[5] Even when multiple medications are indicated and appropriate, polypharmacy can still increase the burden of disease self-management and can confer a higher risk of drug–drug interactions, medication nonadherence, health care use (eg, hospitalization and emergency department visits), and even death.[6,7] Polypharmacy may also increase cost-related, logistical, and cognitive burdens on patients and caregivers, particularly for those managing complex regimens,[8] often with multiple administrations per day, different administration routes (eg, pills, injections, inhalers), and/or frequent refills or changes in prescription. Furthermore, over-the-counter (OTC) medications, whether physician-recommended or part of a self-treatment regimen, can contribute to polypharmacy and its risks, but they are likely underreported in the clinical setting.[9]

Presented in part at American College of Rheumatology Convergence, Washington, DC, November 2024.

The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Supported by the National Institute on Aging, NIH (grant R01-AG-061179 to Dr Plantinga) and the Centers for Disease Control and Prevention (grants U01-DP-006488 and U01-DP-0006698 to Dr Lim).

[1]Alameda Health System, Oakland, California; [2]University of California San Francisco; [3]Emory University, Atlanta, Georgia; [4]Durham Veterans Affairs Geriatric Research Education and Clinical Center, Durham, North Carolina; [5]Duke University, Durham, North Carolina.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25686).

Author disclosures and graphical abstract are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25686.

Address correspondence via email to Laura C. Plantinga, PhD, at laura.plantinga@ucsf.edu.

Submitted for publication February 25, 2025; accepted in revised form October 10, 2025.

SIGNIFICANCE & INNOVATIONS

- Polypharmacy is often associated with adverse health-related and quality-of-life outcomes, but the epidemiology of polypharmacy in the setting of systemic lupus erythematosus (SLE) remains understudied.
- In a cross-sectional study of 451 participants recruited from a US population–based adult SLE cohort, we found that more than half (56%) reported polypharmacy (defined as more than five prescription and over-the-counter medications); estimates were even higher with older age, higher SLE activity and disease damage, and longer disease duration.
- Although immunomodulating medications were commonly reported, polypharmacy was more commonly driven by other medications, including antihypertensives (62%), nonopioid pain relievers (52%), allergy treatments (22%), antidepressants (22%), and gastric reflux medications (22%)
- Although further studies are warranted, our results suggest that the burden of medication regimens in SLE is high and that patients with SLE may benefit from individualized, patient-centered, and goal-oriented medication optimization strategies.

Systemic lupus erythematosus (SLE) is a chronic, multisystem autoimmune disease that is treated with immunomodulating medications. Additionally, patients with SLE often have multiple comorbid conditions related to SLE or its treatment, including cardiovascular disease, metabolic disorders (hypertension, dyslipidemia, and obesity), osteoporosis, allergic disorders, malignancy, and mood disorders,[10] all of which are also common in older age. These conditions may necessitate the use of medications beyond those aimed at treating SLE, contributing to polypharmacy in this population. Also, prescribing cascades,[11] in which new medications are prescribed to counter the effects of a disease-targeted treatment (eg, proton pump inhibitors for gastric reflux related to glucocorticoid use), are likely in the setting of SLE. Combined, these factors likely place those with SLE at high risk for polypharmacy.

Despite the high potential for polypharmacy and the clear importance of preventing related poor outcomes—such as higher burden of self-management, risk of medication nonadherence,[12–14] and risk of drug–drug interactions—in SLE, there is limited research exploring the prevalence of polypharmacy and its potential drivers in this relatively young and disproportionately minoritized population. Estimating polypharmacy prevalence and whether it differs by key sociodemographic and clinical variables is an essential first step toward identifying and targeting strategies to prevent polypharmacy and related adverse outcomes. This study aims to evaluate the prevalence of polypharmacy, including OTC medications, overall and by key sociodemographic characteristics (eg, age, race) and SLE-related factors (eg, SLE activity, immunomodulating medication burden), and to identify the SLE-related and other medications contributing to polypharmacy.

PATIENTS AND METHODS

Study population. For the cross-sectional Approaches to Positive, Patient-Centered Experiences of Aging with Lupus (APPEAL) study, we recruited participants from the ongoing population-based Georgians Organized Against Lupus (GOAL) cohort in metropolitan Atlanta, Georgia.[15] GOAL participants are adults (≥18 years) with SLE defined by ≥4 of the 11 revised American College of Rheumatology (ACR) criteria[16] or ≥3 ACR criteria plus a diagnosis of SLE by a board-certified rheumatologist. English-speaking participants who lived in Georgia at the time of recruitment and had sufficient hearing and vision to undergo testing were included. We recruited a total of 451 participants (206 in-person and 245 remote visits[17]) for APPEAL (Supplementary Figure 1). Data were collected through performance tests and questionnaires administered during a one-time, cross-sectional study visit between October 8, 2019 and May 12, 2022, and were also obtained from the annual GOAL survey closest to the APPEAL visit date (median of 81 days from APPEAL visit to closest GOAL survey; 25th percentile = −79 days [GOAL survey preceding APPEAL visit]; 75th percentile = 199 days; Supplementary Figure 1). The protocols for APPEAL and GOAL were approved by the Emory University Institutional Review Board (IRB; IRB00110977 and IRB00003656), and APPEAL data analysis was approved by the University of California San Francisco IRB (23-39650). All APPEAL participants provided informed consent.

Study variables. *Polypharmacy.* Participants were asked what medications they were taking at the time of the APPEAL study visit. We assessed common medications used to treat SLE and its sequelae through yes or no questions (Supplementary Item 1). We also asked participants to list other prescription and OTC medications they were taking. If participants had pill bottles or medication lists, these were used to document medications. We defined polypharmacy as five or more individual prescriptions or OTC medications reported vs zero to four,[18] excluding eye drops, topical agents, inhalants, and supplements. Medications containing two or more active ingredients were counted as a single medication for the estimation of polypharmacy. Prescribed and OTC medications were also categorized for the examination of the most common classes used (Supplementary Table 1).

Other variables. Sociodemographic variables of interest, which were self-reported at the APPEAL visit, included age, sex at birth, race, ethnicity, and educational attainment. Race was reported using a fixed set of categories and further categorized as Black (single or multiple race), White, and other. Ethnicity was

reported via a separate item, also from a fixed set of categories (Hispanic vs non-Hispanic). Educational attainment was categorized as high school graduate equivalency or lower, some college or associate's degree, and college graduate or higher. Insurance type was self-reported during the closest GOAL visit and categorized as Medicare (including dual-eligible), Medicaid, private, other, and no insurance. SLE-related factors included disease activity and damage and the number of immunomodulating medications. Current SLE activity was assessed during the APPEAL visit via the Systematic Lupus Activity Questionnaire (SLAQ; range 0–44, higher scores indicating greater SLE-related disease activity).[19] The Brief Index of Lupus Damage (BILD) score (range 0–46, higher scores indicate greater cumulative SLE-related organ damage)[20] was obtained from linked GOAL survey data. Disease duration was determined from the duration reported at the closest GOAL visit, corrected for the time between the GOAL and APPEAL visit. Immunomodulating medications, reported at the APPEAL visit, included glucocorticoids, hydroxychloroquine, and other SLE treatments (see Supplementary Table 1).

Statistical analysis. Participant characteristics were summarized overall (means and SDs or medians and interquartile ranges, as appropriate) and by polypharmacy status using descriptive statistics and chi-square, *t*, and rank-sum tests, as appropriate, for comparisons. The crude (unadjusted) prevalence of polypharmacy, overall and across characteristics, was estimated using logistic regression models, with prevalence estimates and 95% confidence intervals obtained from postestimation marginals. We also assessed the prevalence of individual classes of medications (Supplementary Table 1) reported by >5% of participants and stratified them by polypharmacy status to evaluate the associations of participant characteristics with polypharmacy. Finally, in sensitivity analyses, we examined (1) age-adjusted prevalence estimates, using logistic regression models that included age as a covariate, with adjusted percentages obtained from postestimation marginals, to assess the robustness of results to potential confounding by age; and (2) prevalence estimates using alternate definitions of polypharmacy, including hyperpolypharmacy (≥10 medications vs <10 medications) and polypharmacy defined using only prescription medications, to assess the robustness of results to the potential misclassification of polypharmacy. Analyses were performed with Stata version 18.5 (StataCorp, College Station, TX) and R version 4.4.1 (R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

Characteristics of study participants. Participants had a mean age of 46.2 years, with 15.3% being ≥60 years old (Table 1). Most reported being women (91.8%) and Black (81.8%). Nearly one-quarter (23.8%) attained a high school degree or less. Median scores for disease activity (SLAQ) and cumulative disease damage (BILD) were 11 and 2, respectively; the median disease duration was 15.0 years (Table 1).

The maximum overall number of immunomodulating medications was five, with 42.6%, 32.4%, and 25.1% reporting zero to one, two, and three to five, respectively. Those reporting polypharmacy were: older (mean age 48.0 vs 43.8 years), were more likely to report having Medicare insurance (38.2% vs 18.3%) and less likely to report having Medicaid insurance (13.4% vs 23.4%), had higher median SLAQ (12 vs 9) and BILD (2 vs 1) scores and longer disease duration (16.8 vs 13.1 years), and were less likely to be on zero to one immunomodulating medications (28.7% vs 60.4%) (Table 1). Although those with polypharmacy were more likely to be men (9.8% vs 6.1%) and Black (83.1% vs 80.2%), these differences were not statistically significant. There were no differences in ethnicity or education attained by polypharmacy status.

Prevalence of polypharmacy overall and by key participant characteristics. Overall, more than half (n = 254; 56.3%) of participants met the criteria for polypharmacy by our primary definition of five or more prescription or OTC medications (Table 1, Figure 1). About one-quarter (n = 60; 23.3%) met the criteria for hyperpolypharmacy (≥10 prescription or OTC medications). When only prescription medications were considered, 207 (45.9%) and 45 (10.0%) participants had polypharmacy and hyperpolypharmacy, respectively (Figure 1).

Polypharmacy prevalence estimates were higher among older participants (68.1% vs 43.0% for the oldest vs youngest age group); men (67.6% vs 55.3% in women); participants with higher disease activity (65.8% vs 46.2%), greater cumulative disease damage (68.5% vs 42.4%), and longer disease duration (62.4% vs 50.0%); and participants taking three to five vs zero to one immunomodulating medications (79.6% vs 38.0%) (Figure 2). Age-adjusted estimates (Supplementary Table 2) were similar to the unadjusted estimates. The prevalence of polypharmacy defined by prescription medications only and the prevalence of hyperpolypharmacy were both lower overall; however, the patterns of prevalence by participant characteristics were similar, except that differences by disease duration were no longer statistically significant (Supplementary Table 3).

Most common medication classes reported by participants with and without polypharmacy. The prevalence of each medication category by polypharmacy status is summarized in Table 2. Immunomodulating medications were common, with 94.9% and 78.1% of those with and without polypharmacy reporting at least one immunomodulating medication (Figure 3A). However, none reported five or more immunomodulating medications, regardless of polypharmacy status; 28.4%, 64.5%, and 95.2% of those with polypharmacy reported 1 or less, 2 or less, and 3 or less immunomodulating medications,

Table 1. Characteristics of study participants*

Characteristic	Overall	Polypharmacy (≥5 Rx and OTC medications)		
		Yes	No	*P* value[a]
n	451	254 (56.3)	197 (43.7)	
Sociodemographic				
Age, mean (SD), years	46.2 (11.8)	48.0 (11.4)	43.8 (12.0)	<0.001
Age category, n (%)				<0.001
18–39 years	128 (28.4)	55 (21.7)	73 (37.1)	
40–59 years	254 (56.3)	152 (59.8)	102 (51.8)	
≥60 years	69 (15.3)	47 (18.5)	22 (11.2)	
Sex at birth, n (%)				0.15
Women	414 (91.8)	229 (90.2)	185 (93.9)	
Men	37 (8.2)	25 (9.8)	12 (6.1)	
Race, n (%)				0.07
Black	369 (81.8)	221 (83.1)	158 (80.2)	
White	52 (11.5)	32 (12.6)	20 (10.2)	
Other	30 (6.7)	11 (4.3)	19 (9.6)	
Ethnicity, n (%)				0.7
Hispanic	25 (5.6)	13 (5.1)	12 (6.1)	
Non-Hispanic	425 (94.4)	240 (94.9)	185 (93.9)	
Education, n (%)				0.4
High school or less	103 (22.8)	52 (20.5)	51 (25.9)	
Some college	119 (26.4)	70 (27.6)	49 (24.9)	
College or more	229 (50.8)	132 (52.0)	97 (49.2)	
Insurance, n (%)				<0.001
Medicare	133 (29.5)	97 (38.2)	36 (18.3)	
Medicaid	80 (17.7)	34 (13.4)	46 (23.4)	
Private	135 (29.9)	75 (29.5)	60 (30.5)	
Other	32 (7.1)	16 (6.3)	16 (8.1)	
None	71 (15.7)	32 (12.6)	39 (19.8)	
SLE-related				
SLAQ score, median (IQR)	11 (6–16)	12 (7–18)	9 (5–14)	<0.001
BILD score, median (IQR)	2 (1–4)	2 (1–4)	1 (0–3)	<0.001
Disease duration, median (IQR), years	15.0 (9.6–22.3)	16.8 (10.5–23.4)	13.1 (8.0–21.1)	0.001
Immunomodulating medications, n (%)				<0.001
0–1	192 (42.6)	73 (28.7)	119 (60.4)	
2	146 (32.4)	91 (35.8)	55 (27.9)	
3–5	113 (25.1)	90 (35.4)	23 (11.7)	

* n = 451 except for disease duration (n = 450), SLAQ (n = 431), body mass index (n = 440), and physical activity (n = 443). BILD, Brief Index of Lupus Damage; IQR, interquartile range; OTC, over-the-counter; Rx, prescription; SLAQ, Systemic Lupus Activity Questionnaire; SLE, systemic lupus erythematosus.
[a] By chi-square, *t*, or rank-sum test, as appropriate.

respectively (Figure 3A). Among those with polypharmacy, nearly all (99.6%) reported ≥2 other medications, almost half (46.3%) reported >5 other medications, and 13.8% reported >10 other prescription and OTC medications (Figure 3B). Patterns of immunomodulating vs other medications by polypharmacy defined using prescription medications only (Supplementary Figure 2) and by hyperpolypharmacy (Supplementary Figure 3) were similar.

In the polypharmacy group, the most commonly reported nonimmunomodulatory medication classes included antihypertensives (80.7%), nonopioid pain medications (68.9%), gastric reflux medications (35.4%), antidepressants (34.6%), and allergy medications (31.1%) (Supplementary Table 4). Similarly, in the hyperpolypharmacy group, the most frequently used nonimmunomodulatory medication classes were antihypertensives (91.7%), nonopioid pain medications (86.7%), gastric reflux medications (61.7%), antidepressants (58.3%), and allergy treatments (40.0%).

DISCUSSION

In this cross-sectional study of adults with prevalent SLE who were recruited from a population-based cohort, more than half (56%) of participants reported polypharmacy; a substantial proportion, about one-quarter (23%), reported hyperpolypharmacy. When OTC medications were excluded, prevalence remained high, at 46% and 10%, respectively. Polypharmacy was more common among older adults, men, and patients with higher disease activity and damage and longer disease duration. Although medications to treat SLE were common, only 1% with polypharmacy reached the threshold of five or more medications based on these treatments alone. Instead, other medications, such as

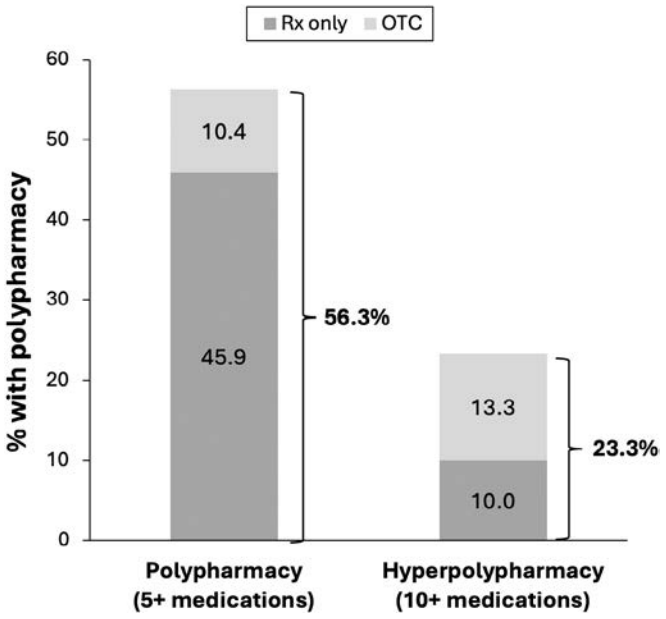


Figure 1. Overall prevalence of polypharmacy in our population of individuals with SLE. OTC, over-the-counter; Rx, prescription.

antihypertensive medications, gastric reflux medications, nonopioid pain medications, and antidepressants contributed substantially to polypharmacy, likely reflecting the significant burdens conferred by disease symptoms, comorbid conditions, and treatment cascade effects in this population.

Our estimate for polypharmacy prevalence far exceeds a recent population estimate for all US adults (20%) and is also approximately 25% higher than that among older US adults (≥65 years; 45%).[2] We found that 60% to 68% of older patients with SLE reported polypharmacy, which was slightly lower than that reported for individuals 50 years and older with SLE in Manitoba (72%); this difference may reflect national differences in health care access or prescribing practices, differences in the source populations (clinic vs study cohort), and differences in how medications were captured (health records vs self-report).[21] Importantly, our estimate among those aged 18 to 39 years (43%) was as high as estimates among older US adults, demonstrating that polypharmacy is prevalent across all age groups in SLE.

It is not surprising that most individuals with SLE would be on multiple medications, given that SLE is a chronic systemic autoimmune disease, often requiring lifelong treatment with immunomodulatory medications. However, our findings show that 95% of participants with polypharmacy were taking three or fewer immunomodulatory medications, suggesting that medications to treat SLE were not a primary driver of polypharmacy in this population. Instead, other medications, including antihypertensives, nonopioid pain medications, gastric reflux medications, and antidepressants, substantially contributed to polypharmacy in our population. This pattern likely reflects the simultaneous burdens of SLE and its symptoms, as well as common comorbid conditions in SLE. It is also likely that these patterns reflect prescribing cascades: for example, treatment with glucocorticoids could lead to hypertension, gastric reflux, and frequent infections, which could lead to treatment with antihypertensives, gastric reflux medications, and anti-infectious agents, respectively.

As expected, we found that older age, higher disease activity and damage, and longer disease duration (all of which may increase the burden of symptoms, comorbid conditions, and treatment side effects and cascades) were associated with higher prevalence of polypharmacy. More surprisingly, whereas previous studies in older adults showed that women are more likely to experience polypharmacy than men,[2] in our cross-sectional study, we found that men had a higher prevalence of polypharmacy than women. This discrepancy, which warrants further investigation, may be due to sex-based differences in SLE burden; sex differences in medication access, use, and/or adherence patterns; and/or chance, given the relatively small proportion of participants who were men.

Our finding that the inclusion of OTC medications resulted in substantially higher polypharmacy prevalence estimates emphasizes the significant contribution of OTC medications such as gastric reflux medications, allergy treatments, and pain relievers to polypharmacy and, particularly, hyperpolypharmacy. OTC medications also likely contribute to patients' out-of-pocket costs. These results underscore the importance of accurately assessing both prescribed and OTC medications, but OTC medications are often not reviewed during clinic visits. In one study comparing electronic health record medication lists to pharmacist-reconciled lists, about half (53%) of electronic health records had discrepancies due to nonreporting of OTC medications.[22] Reviewing these medications during clinic visits may help minimize unnecessary or unwanted medication use (deprescribing), prevent drug–drug interactions, and reduce overall medication burden. Other strategies to address polypharmacy might include pharmacist-led simplification of medication regimens, the use of pill-packaging services (which sort medications by date and time into packets to reduce patient burden and error), or the use of polypills (which reduce the overall number of pills by combining multiple medications into single pills[23]).

Although we cannot distinguish "appropriate" (ie, clinically appropriate and goal concordant) and "inappropriate" polypharmacy[24] with our data, it is important to note that any polypharmacy places a burden on patients. Excessive pill burden,[25] conflicts with personal beliefs about medications and their effectiveness,[13] cognitive decline,[24] fall risk, and higher medical costs[26] are all higher with greater numbers of medications in other populations. The risk of medication interactions and adverse side effects is also higher in the presence of polypharmacy.[27] Medication reconciliation and deprescribing,[28] requiring shared decision-making and careful follow-up, may help reduce the number and doses of prescriptions and OTC medications

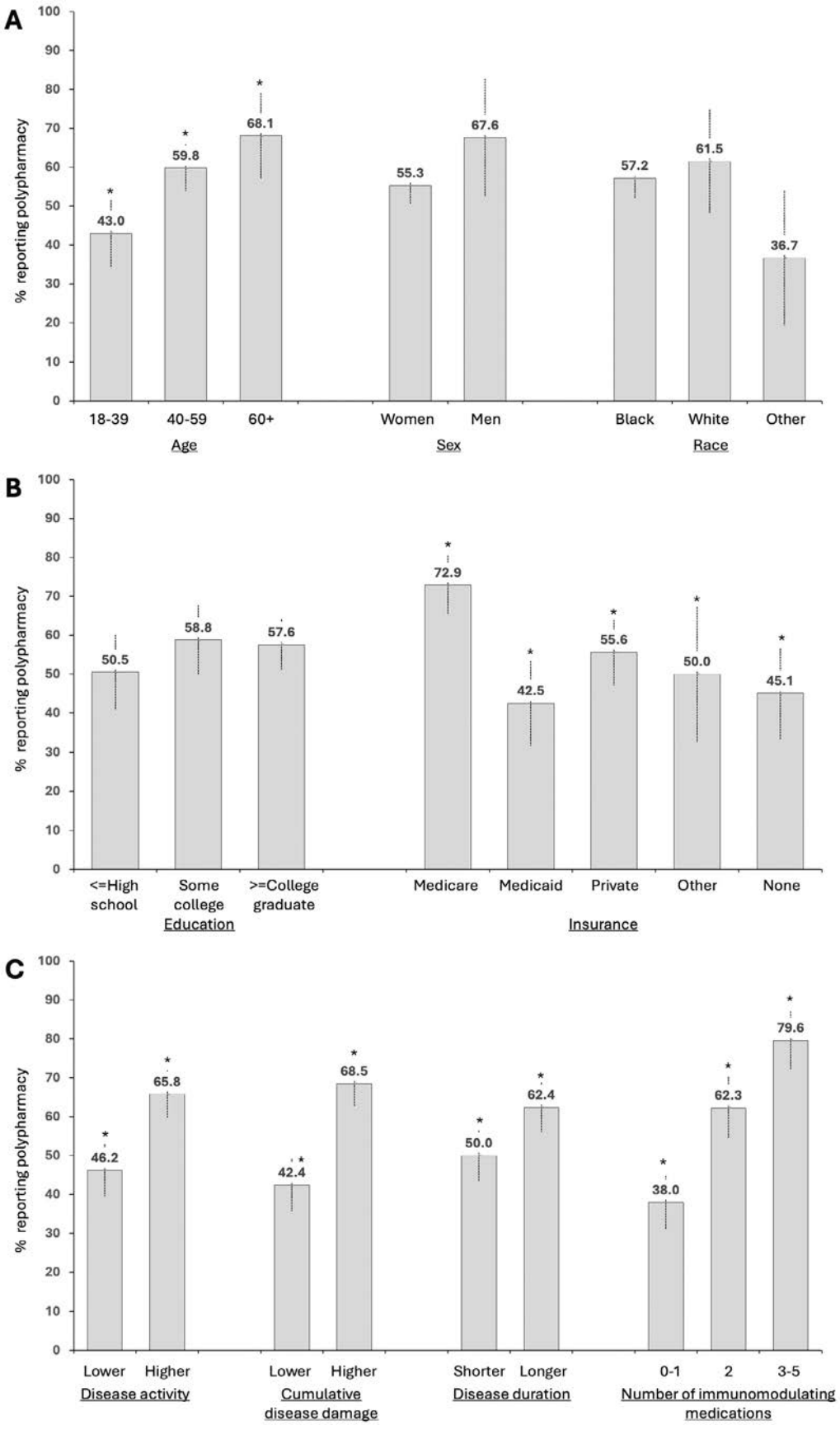


Figure 2. Legend on next page.

Table 2. Most commonly reported (≥5% of all participants) medication classes, overall and by polypharmacy*

Medication class	Overall	Polypharmacy (≥5 prescription and over-the-counter medications)		
		Yes	No	*P* value
Immunomodulating medications, n (%)				
Hydroxychloroquine	322 (71.4)	191 (75.2)	131 (66.8)	0.05
Other[a]	227 (50.3)	165 (65.0)	62 (31.5)	<0.001
Glucocorticoids	200 (44.3)	141 (55.5)	59 (29.9)	<0.001
Other medications, n (%)				
Antihypertensives	279 (61.9)	205 (80.7)	74 (37.6)	<0.001
Nonopioid pain medications[b]	233 (51.7)	175 (68.9)	58 (29.4)	<0.001
Allergy treatments	101 (22.4)	79 (31.1)	22 (11.2)	<0.001
Antidepressants	100 (22.2)	88 (34.6)	12 (6.1)	<0.001
Gastric reflux medications	98 (21.7)	90 (35.4)	8 (4.1)	<0.001
Lipid-lowering agents	86 (19.1)	10 (5.1)	76 (29.9)	<0.001
Opioids	47 (10.4)	44 (17.3)	3 (1.5)	<0.001
Asthma and COPD medications	45 (10.0)	38 (15.0)	7 (3.6)	<0.001
Anticoagulants	38 (8.4)	32 (12.6)	6 (3.0)	<0.001
Anti-infectious agents	34 (7.5)	30 (11.8)	4 (2.0)	<0.001
Diabetes medications	29 (6.4)	25 (9.8)	4 (2.0)	<0.001

* COPD, chronic obstructive pulmonary disease.
[a] Includes methotrexate, leflunomide, cyclophosphamide, cyclosporine, mycophenolate, dapsone, azathioprine, belimumab, rituximab, etanercept, golimumab, tocilizumab, and tofacitinib.
[b] Nonsteroidal anti-inflammatory agents (ibuprofen, naproxen, ketorolac, meloxicam, celecoxib, nabumetone, piroxicam, and indomethacin), acetaminophen, muscle relaxants (cyclobenzaprine, baclofen, tizanidine, cyclobenzaprine, metaxalone, carisoprodol, and orphenadrine), gabapentin, pregabalin, ranolazine, and migraine treatment (rimegepant, ubrogepant, erenumab-aooe, and triptans).

and ensure the maintenance of effective disease control. Deprescribing has also been shown to reduce potentially inappropriate medications and hospitalization among older adults.[28] In a Canadian SLE cohort, an estimated 40% of hospitalizations were related to medications (both adverse drug events and nonadherence-related flares); further, polypharmacy was shown to be associated with a higher risk of hospitalizations due to adverse drug events.[29] To address these issues, future studies should target US SLE populations to examine patient and provider perspectives on polypharmacy, including the acceptability of nonpharmacologic ways to address non–life-threatening but burdensome symptoms (such as fatigue, pain, and gastric reflux)[30]; potential SLE-specific definitions of polypharmacy that account for necessary treatments[31]; the association of polypharmacy with medication-related quality of life, pill and cost[32] burden, medication adherence,[12–14] adverse drug events, and health care use; and the effects of medication reconciliation, deprescribing (including steroid tapering to ≤5 mg/day or elimination within 6 months among stable patients with SLE, as strongly recommended by the 2025 ACR SLE guideline[33]), and other individualized medication optimization strategies on these outcomes in the setting of SLE.

There are several limitations of our study that are not noted above. First, medication use was self-reported, which may have led to underreporting or overreporting (eg, if medications did not align with actual adherence) of polypharmacy. Second, we did not collect information on medication duration, dose, and frequency, which may have led to an overestimation of polypharmacy prevalence, particularly where as-needed medications may have been included in the overall tally. Third, we do not know whether OTC medications were self-initiated or recommended by a rheumatologist or other providers, limiting our ability to assess their clinical appropriateness; however, as noted above, polypharmacy due to clinically appropriate and indicated medications can still be associated with adverse outcomes. Fourth, the temporal gap between some of our variables (eg, BILD score, insurance type) that were taken from the nearest GOAL survey rather than at the time of the APPEAL visit may have contributed to misclassification. Fifth, there is the potential for selection bias due to differences between included and excluded GOAL participants

Figure 2. Prevalence of polypharmacy (≥5 prescription or over-the-counter medications) by selected patient demographics (A), socioeconomic indicators (B), and systemic lupus erythematosus–related factors (C). Systemic lupus erythematosus disease activity was defined by the Systemic Lupus Activity Questionnaire (high = median, low = less than median; median score = 11). Cumulative systemic lupus erythematosus–related disease damage was defined by the Brief Index of Lupus Damage (high = median, low = less than median; median score = 2). Disease duration was also dichotomized by the median (longer = median, shorter = less than median; median duration = 15.0 years). Estimates are postestimation marginals from crude logistic regression, with polypharmacy as the outcome and the selected variables as exposures. *$P < 0.001$ by likelihood ratio test.

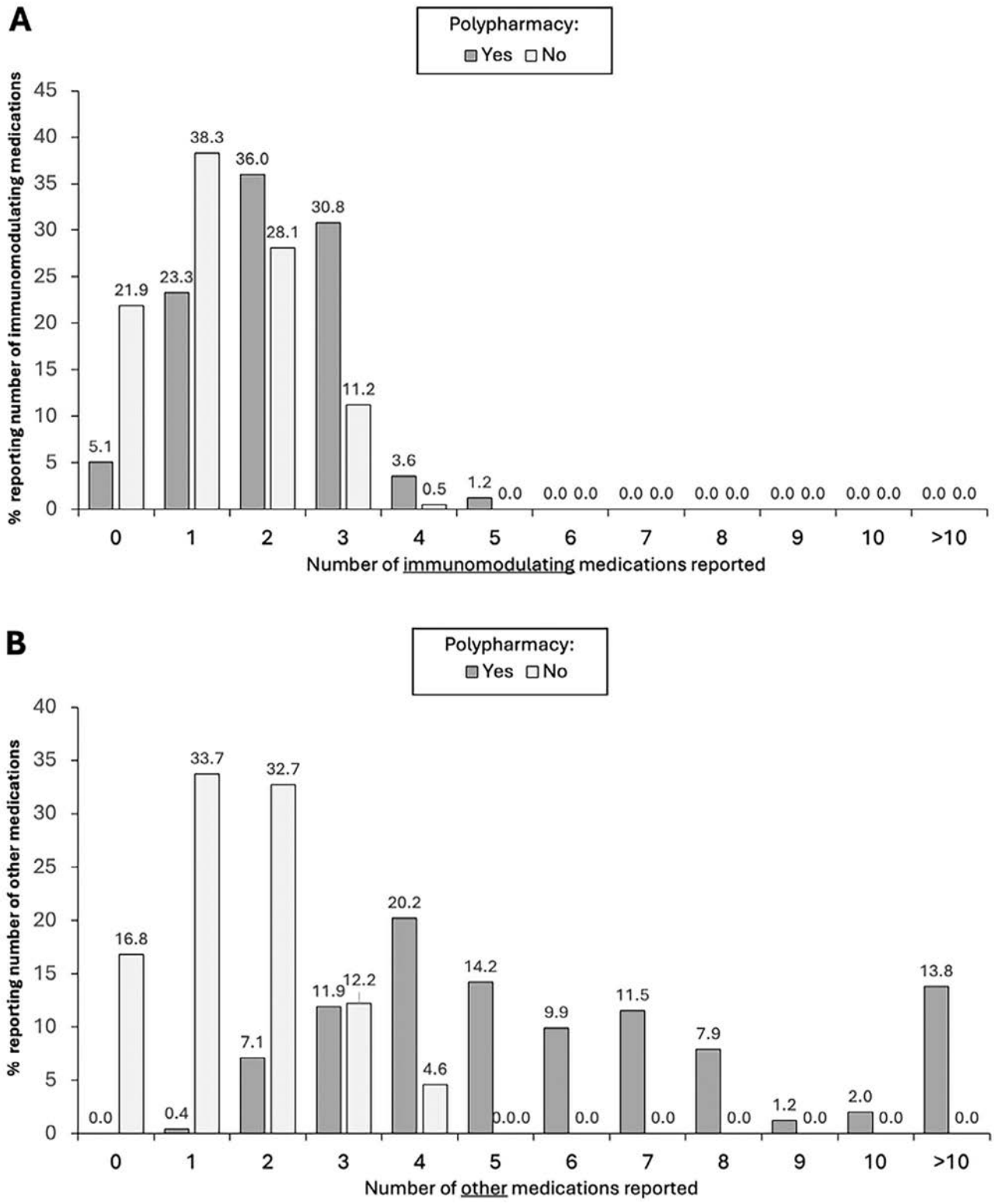


Figure 3. Contribution of immunomodulating (A) and other (B) medications to reported polypharmacy. Shown are the percentages of participants with and without polypharmacy who reported from 0 to >10 immunomodulating and other prescription and over-the-counter medications. Note: no participants reported more than 5 immunomodulating medications.

(Supplementary Figure 1); for example, in a previous study, we showed that APPEAL vs GOAL participants were slightly younger (46 vs 49 years old) and had lower disease activity (median SLAQ score 11 vs 14), characteristics that were shown to be associated with polypharmacy in the current study.[34] Sixth, our study and the parent cohort do not include healthy controls or controls with

other conditions, making it difficult to distinguish the relative contributions of SLE and other conditions to polypharmacy in this population. Finally, our results may not be generalizable to other patients with SLE, including those in other US regions and those outside the United States.

Despite these limitations, this is, to our knowledge, a novel exploration of polypharmacy in a representative US SLE population. Additionally, our study population was primarily Black, reflecting a population disproportionately affected by SLE and related poor outcomes. We estimated a high prevalence of polypharmacy and hyperpolypharmacy across all subgroups, often higher than estimates seen in the general US population of older adults, with even higher prevalence estimates among those who were older or had higher disease burden or longer disease duration. Although future studies are needed to better characterize patient and provider perspectives on polypharmacy and appropriate prescribing, estimate associations of polypharmacy with important outcomes (eg, quality of life, costs, medication adherence, adverse events, and health care use), and explore the effectiveness of individualized medication optimization interventions to prevent and mitigate polypharmacy, these results emphasize the need to address the complexity of medication regimens in this population through patient-centered medication optimization efforts.

ACKNOWLEDGMENTS

We thank the participants of the APPEAL study and Jessie Black, Aita Akharume, Meaza Girmay, and Sydnei Simpson for completing study visits.

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 Plantinga 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. Wang Z, Liu T, Su Q, et al. Prevalence of polypharmacy in elderly population worldwide: a systematic review and meta-analysis. Pharmacoepidemiol Drug Saf 2024;33(8):e5880.
2. Wang X, Liu K, Shirai K, et al. Prevalence and trends of polypharmacy in U.S. adults, 1999-2018. Glob Health Res Policy 2023;8(1):25.
3. Oosting IJ, Colombijn JMT, Kaasenbrood L, et al. Polypharmacy in patients with CKD: a systematic review and meta-analysis. Kidney360 2024;5(6):841–850.
4. Oliveira RF, Oliveira AI, Cruz AS, et al. Polypharmacy and drug interactions in older patients with cancer receiving chemotherapy: associated factors. BMC Geriatr 2024;24(1):557.
5. Johansson KS, Jimenez-Solem E, Petersen TS, et al. Increasing medication use and polypharmacy in type 2 diabetes: the Danish experience from 2000 to 2020. Diabetes Care 2024;47(12):2120–2127.
6. Varghese D, Ishida C, Patel P, et al. Polypharmacy. In: StatPearls. StatPearls Publishing; 2025.
7. Chang TI, Park H, Kim DW, et al. Polypharmacy, hospitalization, and mortality risk: a nationwide cohort study. Sci Rep 2020;10(1):18964.
8. Brysch EG, Cauthon KAB, Kalich BA, et al. Medication regimen complexity index in the elderly in an outpatient setting: a literature review. Consult Pharm 2018;33(9):484–496.
9. Hensrud DD, Engle DD, Scheitel SM. Underreporting the use of dietary supplements and nonprescription medications among patients undergoing a periodic health examination. Mayo Clin Proc 1999;74(5):443–447.
10. Figueroa-Parra G, Meade-Aguilar JA, Hulshizer CA, et al. Multimorbidity in systemic lupus erythematosus in a population-based cohort: the Lupus Midwest Network. Rheumatology (Oxford) 2024;63(11):3056–3064.
11. Wahler RG Jr, Olbrich CL, McCarthy LM. Preventing prescribing cascades: ensuring medication safety for older adults. J Gerontol Nurs 2024;50(9):7–11.
12. Abdul-Sattar AB, Abou El Magd SA. Determinants of medication nonadherence in Egyptian patients with systemic lupus erythematosus: Sharkia Governorate. Rheumatol Int 2015;35(6):1045–1051.
13. Emamikia S, Gomez A, Ådahl T, et al. Factors associated with nonadherence to medications in systemic lupus erythematosus: results from a Swedish survey. Lupus 2024;33(6):615–628.
14. Mehat P, Atiquzzaman M, Esdaile JM, et al. Medication nonadherence in systemic lupus erythematosus: a systematic review. Arthritis Care Res (Hoboken) 2017;69(11):1706–1713.
15. Lim SS, Drenkard C. Understanding lupus disparities through a social determinants of health framework: the Georgians Organized Against Lupus research cohort. Rheum Dis Clin North Am 2020;46(4):613–621.
16. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1997;40(9):1725.
17. Hoge C, Bowling CB, Dunlop-Thomas C, et al. Remote administration of physical and cognitive performance assessments in a predominantly Black cohort of persons with systemic lupus erythematosus. ACR Open Rheumatol 2023;5(9):499–507.
18. Masnoon N, Shakib S, Kalisch-Ellett L, et al. What is polypharmacy? A systematic review of definitions. BMC Geriatr 2017;17(1):230.
19. Yazdany J, Yelin EH, Panopalis P, et al. Validation of the systemic lupus erythematosus activity questionnaire in a large observational cohort. Arthritis Rheum 2008;59(1):136–143.
20. Drenkard C, Yazdany J, Trupin L, et al. Validity of a self-administered version of the brief index of lupus damage in a predominantly African American systemic lupus erythematosus cohort. Arthritis Care Res (Hoboken) 2014;66(6):888–896.
21. Séguin DJ, Peschken CA, Dolovich C, et al. Polypharmacy and potentially inappropriate medication use in older adults with systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2023;75(2):356–364.
22. Stewart AL, Lynch KJ. Identifying discrepancies in electronic medical records through pharmacist medication reconciliation. J Am Pharm Assoc (2003) 2012;52(1):59–66.
23. Akbar MA, Batool A, Khawar M, et al. Polypill strategy: a paradigm shift in cardiovascular disease prevention. Am J Cardiovasc Drugs 2025;25(6):735–737.
24. Chippa V, Roy K. Geriatric cognitive decline and polypharmacy. In: StatPearls. StatPearls Publishing; 2025.
25. Woodford HJ. Polypharmacy in older patients. Br J Hosp Med (Lond) 2024;85(10):1–12.

26. Jandu JS, Mohanaselvan A, Dahal R, et al. Strategies to reduce polypharmacy in older adults. In: StatPearls. StatPearls Publishing; 2025.
27. Novaes PH, da Cruz DT, Lucchetti ALG, et al. The "iatrogenic triad": polypharmacy, drug-drug interactions, and potentially inappropriate medications in older adults. Int J Clin Pharm 2017;39(4):818–825.
28. Carollo M, Crisafulli S, Vitturi G, et al. Clinical impact of medication review and deprescribing in older inpatients: a systematic review and meta-analysis. J Am Geriatr Soc 2024;72(10):3219–3238.
29. Stanciu M, Lee JE, McDonald EG, et al. Medication-related hospitalisations in patients with SLE. Lupus Sci Med 2025;12(1):e001362.
30. Chang A, Winquist NW, Wescott AB, et al. Systematic review of digital and non-digital non-pharmacological interventions that target quality of life and psychological outcomes in adults with systemic lupus erythematosus. Lupus 2021;30(7):1058–1077.
31. Saeed ZI, Ostrominski JW, Aroda VR. Polypharmacy in type 2 diabetes: appropriate or cause for concern? Diabetes Care 2024;47(12):2104–2106.
32. Minhas D, Marder W, Harlow S, et al. Access and cost-related nonadherence to prescription medications among lupus patients and controls: the Michigan Lupus Epidemiology and Surveillance Program. Arthritis Care Res (Hoboken) 2021;73(11):1561–1567.
33. Sammaritano LR, Askanase A, Bermas BL, et al. 2025 American College of Rheumatology (ACR) guideline for the treatment of systemic lupus erythematosus. Arthritis Care Res (Hoboken) Published online November 3, 2025. doi:https://doi.org/10.1002/acr.25690
34. Plantinga LC, Bowling CB, Hoge C, et al. Physical performance in a diverse, population-based cohort of individuals with systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2024;76(4):517–525.

Arthritis Care & Research
Vol. 78, No. 6, June 2026, pp 743–748
DOI 10.1002/acr.25701

Role of Prednisone and 25-Hydroxyvitamin D on Bone Mineral Density and Osteoporosis in Systemic Lupus Erythematosus

Nima Madanchi,[1] Andrea Fava,[1] Daniel W. Goldman,[1] Laurence S. Magder,[2] and Michelle Petri[1]

Objective. We evaluated the role of 25-hydroxyvitamin D (25[OH]D), prednisone, and other risk factors for bone mineral density (BMD) loss and osteoporosis in systemic lupus erythematosus (SLE).

Methods. We calculated the association between 25(OH)D levels and other potential risk factors and BMD measures (spine T scores) and osteoporosis (defined as a T score ≤ –2.5) in a lupus cohort of 1,003 patients. Generalized estimating equations were used to assess the statistical significance of observed differences, accounting for patients with multiple BMD measures.

Results. Univariate analysis showed that older age, lower body mass index (BMI), mean past SLE activity (Safety of Estrogens in Lupus Erythematosus National Assessment–Systemic Lupus Erythematosus Disease Activity Index > 2), mean daily prednisone dose at any level, and a history of low C3 or C4 were associated with osteoporosis. Older age, lower BMI, mean 25(OH)D levels, mean daily prednisone dose (again at any dose), and prior low C3 were associated with lower spine T scores. In the multivariate analysis for osteoporosis, older age, lower BMI, low mean 25(OH)D levels, and prednisone dose >5 mg/day remained predictors. Discontinuation of prednisone within one year before the next BMD measurement led to higher mean spine T scores.

Conclusion. Prednisone dose, even at 5 mg/day, and low mean 25(OH)D levels were found to be modifiable predictors of osteoporosis. Mean (reflecting supplementation), rather than initial, 25(OH)D levels were significantly associated with BMD. Some of the bone loss from prior prednisone use appears to be reversible.

INTRODUCTION

Systemic lupus erythematosus (SLE) leads to a higher risk of low bone mineral density (BMD) at all measured sites compared to expected.[1,2] The risk of osteopenia is 25% to 74%[3,4] and that of osteoporosis is 1.4% to 68%[5,6] in SLE. Fracture occurrence in patients with SLE is increased nearly five-fold compared with healthy women of a similar age.[7] The association of SLE and osteoporosis is particularly complex, as a recent genomic study found highly overlapping differential gene expressions in SLE and osteoporosis and a close relationship between the molecular mechanisms of these two entities.[8]

Older age, low body mass index (BMI) and weight, SLE duration, organ damage score, renal dysfunction, high C-reactive peptide or erythrocyte sedimentation rate, and cumulative dose of glucocorticoid have all been previously reported to be risk factors of low BMD in SLE.[2,9,10] Low vitamin D level is a risk factor for low BMD in the general population.[11] No previous study has explored the relationship between initial versus mean serum vitamin D levels and low BMD in SLE.

In this study, we evaluated the association between SLE disease activity, prednisone, 25-hydroxyvitamin D (25[OH]D) level (initial and over time), and other risk factors with bone density loss and osteoporosis in the Hopkins Lupus Cohort.

PATIENTS AND METHODS

Study design. The Hopkins Lupus Cohort includes patients with confirmed SLE classification based on the Systemic Lupus International Collaborating Clinics (SLICC) 2012 classification criteria[12] who are living in the Baltimore (Maryland, United States) area. Patient race and ethnicity was self-reported, using a fixed set of categories using NIH terminology. The cohort was approved by the Johns Hopkins University School of

The Hopkins Lupus Cohort was supported by the NIH (grants R01-AR-069572 [to Dr Petri] and R01-DK-134625 [to Dr Fava]).
[1]Johns Hopkins University, Baltimore, Maryland; [2]University of Maryland, Baltimore.
Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25701.

Address correspondence via email to Michelle Petri, MD, MPH, at mpetri@jhmi.edu.
Submitted for publication April 24, 2025; accepted in revised form October 29, 2025.

SIGNIFICANCE & INNOVATIONS

- Osteoporosis in systemic lupus erythematosus is driven by glucocorticoids. However, the role of glucocorticoids and other risk factors such as vitamin D on bone mineral density is not understood.
- This study is the first to look at multiple ways to measure glucocorticoid therapy and the first to look at longitudinal measures of 25-hydroxyvitamin D.
- Any mean prednisone dose of 5 mg or higher increased the risk of osteoporosis significantly.
- Bone mineral density improves with glucocorticoid taper.

Medicine institutional review boards (PRN number NA_00039294). Patients' rights, safety, and well-being were protected based on the principles of the Declaration of Helsinki. All patients gave written informed consent. The participants' health and private data were kept confidential. Patients in the cohort were seen at quarterly clinic visits, which included measures of serum 25(OH)D levels since 2009. BMD was measured periodically using dual x-ray absorptiometry (DXA) scans done for clinical care.

Inclusion and exclusion criteria. Patients aged 40 years or older were included. Patients were excluded if there were no 25(OH)D level or bone density scan results.

Patient population. Of the 2,860 patients in our cohort at the time of this analysis, there were 3,007 DXA measures from 1,003 different patients. In total, 320 patients had a single measure, 451 had two to four measures, and 232 had five or more. BMD measures were made due to clinical indications.

Measurements. We measured 25(OH)D levels by chemiluminescence immunoassay. We used the spine T scores in the DXA scans (Hologic Horizon DXA System) for BMD. T scores were calculated by comparing the BMD result with the BMD results from healthy 25- to 35-year-old adults of the same sex and ethnicity. Osteoporosis was defined as a T score ≤ –2.5.

Statistical analysis. For each BMD measure of each patient, we defined potential time-varying predictors based on measurements from cohort visits preceding that BMD assessment. For example, for 25(OH)D, we considered two possible predictors: the average 25(OH)D levels in all previous cohort visits and the 25(OH)D level measured at the first cohort visit. We compared predictor groups with respect to the mean spine T score and the proportion of patients with a T score ≤ –2.5. The statistical significance of observed differences was quantified using *P* values based on a generalized estimating equation (GEE) model that accounts for correlation between repeated measures on the same patient. Multivariate logistic regression based on GEE was used to estimate the association between predictors and osteoporosis adjusting for age, prednisone dose, race, BMI, and history of low complement.

RESULTS

There were 918 female and 85 male patients. There were 377 Black patients, 564 White patients, and 62 patients of other races (29 East Asian patients, 33 not specified). A total of 92.6% had a high school diploma or a higher degree. The mean age at the diagnosis of SLE was 38.0 years. A total of 37.6% were past smokers, 6.0% had a prior alcohol abuse history, and 5.2% had a prior illicit drug use history. Regarding the SLICC classification criteria, 45.2% had acute cutaneous lupus, 19.9% had discoid lupus, 53.4% had photosensitivity, 56.7% had oral or nasal ulcers, 76.5% had arthritis, 97.1% had antinuclear antibody, 48.2% had serositis, 9.9% had neurologic involvement, 70.2% had hematologic involvement, and 98.4% had immunologic involvement.

Table 1 shows the mean lumbar spine T scores and the prevalence of osteoporosis in different patient subgroups. Older age, lower BMI, mean Safety of Estrogens in Lupus Erythematosus National Assessment–Systemic Lupus Erythematosus Disease Activity Index (SELENA-SLEDAI) of >2 (including serologies), higher mean daily prednisone dose, and prior low C3 or C4 levels predicted osteoporosis in the univariate analysis. Older age, lower BMI, mean 25(OH)D levels, mean daily prednisone dose, and prior low C3 (but not low C4) levels predicted lower spine T scores in the univariate analysis. Cumulative dose of prednisone was also a predictor of osteoporosis ($P = 0.0019$) and comparable to mean prednisone dose as a predictor (data not shown).

Table 2 shows the multivariate analysis for osteoporosis in different subgroups, controlling for age, race, BMI, prednisone, and history of low complement. Older age, Black race, low BMI, mean 25(OH)D levels, and mean prednisone dose >5 mg/day were independent predictors of osteoporosis. A mean prednisone dose of less than 5 mg/day was still associated with 1.35 times higher risk of osteoporosis, although the association did not reach statistical significance.

We examined other factors in the univariate and multivariate models. The association with menopause status could not be assessed with precision due to the fact that only a small percentage (5%) of the BMD measures were made in premenopausal women, and this was highly correlated with age. Osteoporosis and mean T score were not associated with calcium and/or vitamin D supplementation, which we recorded from 2009 onwards. We use bisphosphonates routinely for osteoporosis but could not analyze as we did not collect data on the start date. Hip osteoporosis was less common than lumbar spine and increased with higher prednisone doses, but not significantly.

Table 1. Univariate analysis: prevalence of osteoporosis and mean T score in patient subgroups*

Predictor	DXA measures, n	Osteoporosis, n (%)	*P* value[a]	T score, mean (SD)	*P* value[a]
Age, y			0.0001		<0.0001
40–49	1,080	83 (8)		−0.56 (1.32)	
50–59	1,049	109 (10)		−0.85 (1.34)	
60–69	677	106 (16)		−0.92 (1.54)	
70+	201	31 (15)		−0.72 (1.71)	
Sex			0.3600		0.3600
Female	2,786	308 (11)		−0.77 (1.40)	
Male	221	21 (10)		−0.60 (1.63)	
Race			0.1400		0.3300
Black patients	1,116	132 (12)		−0.67 (1.52)	
White patients	1,722	166 (10)		−0.78 (1.34)	
Other	169	31 (18)		−1.09 (1.37)	
Body mass index			0.0001		<0.0001
<18	76	17 (22)		−1.67 (1.18)	
18–25	1,056	177 (17)		−1.18 (1.29)	
25–30	885	79 (9)		−0.65 (1.39)	
30–35	538	35 (7)		−0.30 (1.47)	
25+	430	16 (4)		−0.26 (1.34)	
Mean SELENA-SLEDAI[b]			0.0028		0.3100
<2	1,428	118 (8)		−0.70 (1.39)	
2–4	929	132 (14)		−0.83 (1.48)	
4+	404	57 (14)		−0.84 (1.40)	
Mean clinical SELENA-SLEDAI[b]			0.3400		0.7500
<2	2,061	218 (11)		−0.78 (1.41)	
2–4	622	81 (13)		−0.75 (1.46)	
4+	78	8 (10)		−0.49 (1.47)	
Mean 25 (OH)D, ng/mL[b]			0.7000		0.0025
<20	91	16 (18)		−0.76 (1.77)	
20–30	273	28 (10)		−0.61 (1.42)	
30–40	531	52 (10)		−0.65 (1.42)	
40–50	648	67 (10)		−0.67 (1.47)	
50+	410	53 (13)		−0.99 (1.39)	
First 25 (OH)D, ng/mL			0.1200		0.0890
<20	463	60 (13)		−0.70 (1.50)	
20–30	455	45 (10)		−0.69 (1.42)	
30–40	540	47 (9)		−0.60 (1.44)	
40–50	321	42 (13)		−0.99 (1.41)	
50+	174	22 (13)		−0.83 (1.51)	
Mean prednisone dose, mg/day[b]			0.0006		<0.0001
None	875	58 (7)		−0.52 (1.36)	
<5	902	87 (10)		−0.78 (1.38)	
5–10	657	101 (15)		−0.89 (1.52)	
10+	329	62 (19)		−1.10 (1.41)	
Prior anti-dsDNA[c]			0.2200		0.5200
N	1,339	142 (11)		−0.75 (1.38)	
Y	1,423	166 (12)		−0.77 (1.47)	
Prior low C3[c]			0.0290		0.0097
N	1,431	125 (9)		−0.59 (1.45)	
Y	1,331	182 (14)		−0.94 (1.37)	
Prior low C4[c]			0.0096		0.3500
N	1,657	158 (10)		−0.72 (1.43)	
Y	1,105	149 (13)		−0.83 (1.41)	

* 25(OH)D, 25-hydroxyvitamin D; BMD, bone mineral density; dsDNA, double-stranded DNA; DXA, dual x-ray absorptiometry; GEE, generalized estimating equation; SELENA-SLEDAI, Safety of Estrogens in Lupus Erythematosus National Assessment–Systemic Lupus Erythematosus Disease Activity Index.

[a] *P* values based on GEE to account for repeated observations on the same person.

[b] Mean SELENA-SLEDAI, 25(OH)D level, and prednisone dose were defined as the mean from all cohort visits before the visit at which BMD was measured. Clinical SELENA-SLEDAI means that C3, C4, and anti-dsDNA are omitted from the calculation of the SELENA-SLEDAI.

[c] Prior anti-dsDNA, prior low C3, and prior low C4 were defined as the occurrence at any cohort visit before the visit at which BMD was measured.

Table 2. Multivariate analysis: odds ratios for osteoporosis in different subgroups controlling for potential confounding variables*

Predictor	Adjusted odds ratio (95% CI)[a]	Adjusted *P* values[a]
Age, y		
40–49	1.00 (ref group)	–
50–59	2.02 (1.50, 2.73)	<0.0001
60–69	2.61 (1.80, 3.76)	<0.0001
70+	2.45 (1.53, 3.91)	0.0002
Sex		
Female	1.00 (ref group)	–
Male	0.64 (0.30, 1.36)	0.2400
Race		
Black patients	1.53 (1.02, 2.30)	0.0390
White patients	1.00 (ref group)	–
Other	1.63 (0.75, 3.55)	0.2100
Body mass index		
<18	1.30 (0.75, 2.26)	0.3500
18–25	1.00 (ref group)	–
25–30	0.62 (0.48, 0.82)	0.0006
30–35	0.48 (0.33, 0.71)	0.0003
35+	0.32 (0.20, 0.53)	<0.0001
Mean clinical SELENA-SLEDAI[b]		
<2	1.00 (ref group)	–
2–4	1.14 (0.77, 1.69)	0.5200
4+	1.18 (0.61, 2.27)	0.6300
Mean 25 (OH)D level, ng/mL[b]		
<20	1.00 (ref group)	–
20–30	0.58 (0.34, 1.00)	0.0510
30–40	0.57 (0.31, 1.05)	0.0690
40–50	0.46 (0.25, 0.87)	0.0160
50+	0.54 (0.27, 1.11)	0.0930
First 25 (OH)D level, ng/mL		
<20	1.00 (ref group)	–
20–30	0.62 (0.34, 1.13)	0.1200
30–40	0.52 (0.29, 0.93)	0.0290
40–50	0.77 (0.40, 1.47)	0.4300
50+	0.67 (0.31, 1.46)	0.3100
Mean prednisone dose, mg/day[b]		
None	1.00 (ref group)	–
<5	1.35 (0.84, 2.17)	0.2100
5–10	2.04 (1.25, 3.31)	0.0041
10+	2.60 (1.56, 4.34)	0.0002
Prior anti-dsDNA[c]		
N	1.00 (ref group)	–
Y	0.86 (0.61, 1.22)	0.4100
Prior low C3[c]		
N	1.00 (ref group)	–
Y	0.97 (0.68, 1.38)	0.8600
Prior low C4[c]		
N	1.00 (ref group)	–
Y	1.25 (0.85, 1.84)	0.2500

* 25(OH)D, 25-hydroxyvitamin D; BMD, bone mineral density; CI, confidence interval; dsDNA, double-stranded DNA; GEE, generalized estimating equation; SELENA-SLEDAI, Safety of Estrogens in Lupus Erythematosus National Assessment–Systemic Lupus Erythematosus Disease Activity Index.
[a] Adjusted for age, race, body mass index, prednisone dose, low C3, and low C4 using a GEE logistic regression.
[b] Mean clinical SELENA-SLEDAI, 25(OH)D level, and prednisone dose were defined as the mean from all cohort visits before the visit at which BMD was measured. The clinical SLEDAI removes low complement and anti-DNA points.
[c] Prior anti-dsDNA, prior low C3, and prior low C4 were defined as the occurrence at any cohort visit before the visit at which BMD was measured.

Table 3 addresses the question of whether reducing prednisone use led to an improvement in BMD. We classified patients based on their mean prednisone use in the year before the BMD measurement and before that as well. We found that those who had a history of prednisone use and discontinued their prednisone within one year before the next BMD measurement had a lower risk of osteoporosis ($P = 0.095$) and a higher mean BMD ($P = 0.038$) than those who did not discontinue. However, the risk was not as low as those who never took prednisone in the past.

DISCUSSION

Our study confirmed the general population findings that age increases osteoporosis, and higher BMI is protective. The SLE specific risk factors found were prednisone >5 mg/day and low initial and mean 25(OH)D levels. Finally, reduction in prednisone appeared to reduce the risk of osteoporosis at the next BMD measurement.

There was a dose-response for prednisone dose and osteoporosis in our multivariate model. EULAR updated its 2019 recommendations for the management of SLE in 2023 and lowered the target maintenance prednisone dose from 7.5 to 5 mg/day or lower.[13] We also, however, found an increased risk among those with a mean prednisone dose of less than 5 mg/day (although not statistically significant). As osteoporotic fractures are one of the most common items found in the SLICC/American College of Rheumatology (ACR) Damage Index,[14] reduction in prednisone use remains the most important modifiable risk factor for osteoporosis. Encouragingly (Table 3), reduction in prednisone use appeared to lead to improvement at the next BMD measurement, indicating that bone loss in SLE is partially reversible. Studies on Cushing syndrome after treatment also showed improvement in BMD, including improvement after many years,[15,16] emphasizing the potential reversibility of glucocorticoid-induced bone loss.

Vitamin D is a sterol hormone (rather than a vitamin) with multiple metabolic effects. Vitamin D insufficiency and deficiency are extremely common in SLE. In our past analysis, 25(OH)D supplementation helped proteinuria, but the benefit plateaued at a 25(OH)D level of 40 ng/mL.[15]

We found an association between osteoporosis risk and SLE disease activity in the univariate analysis only. Past studies did not find such an association with SLE disease activity.[10,17,18] This association appeared to be explained by the immunologic components of SELENA-SLEDAI, because when they were removed to create "clinical SLEDAI," the association disappeared (Table 1). When we adjusted for prednisone, the relationship between the immunologic components (low complement and anti–double-stranded DNA) and osteoporosis risk were not significant, suggesting that the association observed in the univariate analyses between low complement and osteoporosis was due to confounding by prednisone use. One study previously reported

Table 3. Estimated relationship across the timing of prednisone use, prednisone reduction, and BMD*

Mean prednisone dose more than a year before BMD measurement	Mean prednisone dose in the year before the BMD measurement	Number (%) with osteoporosis	Spine T score, mean (SD)
>5 mg/day	None (n = 276)	33 (12.0)[a]	−0.70 (1.53)[b,c]
	Any (n = 619)	113 (18.3)[a,b,d]	−1.10 (1.43)[b,d,e]
<5 mg/day	None (n = 1,392)	102 (7.3)[a,d]	−0.62 (1.36)[d,c]
	Any (n = 259)	36 (13.9)[b]	−0.93 (1.47)[e]

* BMD, bone mineral density.
[a] $P = 0.0950$.
[b] $P = 0.0380$.
[c] $P = 0.0730$.
[d] $P < 0.0001$.
[e] $P = 0.0061$.

that lower complement levels were associated with osteoporosis.[9] In contrast, another study reported higher C3 levels in patients with SLE with low BMD.[19] Our previous 1995 study (with less follow up and smaller number of patients) showed low C4 was associated with low BMD.[20] However, C3-lacking mice had higher BMD.[21] We now conclude that SLE disease activity and low complement levels are not associated with osteoporosis in SLE.

In our cohort, patients with SLE with low 25(OH)D levels are advised to start either daily 2,000 IU or weekly 50,000 IU. Thus the "mean" levels in our study reflect supplementation (although adherence remains an issue).

The role of vitamin D supplementation in the primary prevention of low BMD and fractures in the general population, however, remains controversial. A clinical trial in 2004 reported that vitamin D and calcium supplementation reduced osteoporotic fractures in community-dwelling elderly people with known vitamin D deficiency.[22] A meta-analysis in 2016 reported the use of calcium plus vitamin D supplements produced a statistically significant 15% reduced risk of total fractures and a 30% reduced risk of hip fractures in both community-dwelling and institutionalized middle-aged to older adults.[23] However, two more recent meta-analyses found the use of supplementation with calcium, vitamin D, or both was not associated with a significant difference in the risk of fractures in community-dwelling older adults without known vitamin D deficiency, osteoporosis, or prior fracture.[24,25] These studies do not generalize well though to SLE, as patients with SLE are generally younger and vitamin D insufficiency and deficiency are almost universal in SLE.[26,27]

Regarding the primary prevention of fractures, the United States Preventive Services Task Force in 2018 recommended against daily supplementation with 400 IU or less of vitamin D and 1,000 mg or less of calcium in community-dwelling postmenopausal women. The task force concluded that the current evidence was insufficient to assess the balance of the benefits and harms of vitamin D and calcium supplementation, alone or combined, in men and premenopausal women.[28] A more recent randomized clinical trial in 2020 reported that in healthy vitamin D–sufficient individuals aged 55 to 70 years, treatment with high-dose vitamin D compared to low-dose vitamin D resulted in volumetric BMD loss in women but not in men.[29]

The strength of this study is the large number of patients with SLE followed in one center with the recording of multiple known risk factors for osteoporosis longitudinally. However, there are unavoidable limitations. Patients under 40 years were excluded. BMD measurements were done for clinical care and not by protocol. Thus, the prevalence of osteoporosis reported in this study applies to those with a clinical indication for a BMD and not to the general SLE population. The cohort is primarily composed of Black and White patients, so the results are not generalizable to other races. Although confounding by indication is always a possibility, the relationship between disease activity and BMD was very small after adjusting for prednisone dose and other variables in the multivariate model. Also, the relationship between prednisone dose and BMD persisted after adjustments in the multivariate models and has been seen by other researchers, so we believe it is real.

We recommend avoiding prednisone use at any level in SLE to reduce the risk of bone loss and osteoporosis. For the first time, the complex relationship of 25(OH)D with BMD has been elucidated. Optimal 25(OH)D levels should be at 40 to 50 ng/mL

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 Petri 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. Gu C, Zhao R, Zhang X, et al. A meta-analysis of secondary osteoporosis in systemic lupus erythematosus: prevalence and risk factors. Arch Osteoporos 2019;15(1):1.

2. Almehed K, Forsblad d'Elia H, Kvist G, et al. Prevalence and risk factors of osteoporosis in female SLE patients-extended report. Rheumatology (Oxford) 2007;46(7):1185–1190.
3. Kalla AA, Fataar AB, Jessop SJ, et al. Loss of trabecular bone mineral density in systemic lupus erythematosus. Arthritis Rheum 1993; 36(12):1726–1734.
4. Mok CC, Mak A, Ma KM. Bone mineral density in postmenopausal Chinese patients with systemic lupus erythematosus. Lupus 2005; 14(2):106–112.
5. Boyanov M, Robeva R, Popivanov P. Bone mineral density changes in women with systemic lupus erythematosus. Clin Rheumatol 2003; 22(4-5):318–323.
6. Uaratanawong S, Deesomchoke U, Lertmaharit S, et al. Bone mineral density in premenopausal women with systemic lupus erythematosus. J Rheumatol 2003;30(11):2365–2368.
7. Ramsey-Goldman R, Dunn JE, Huang CF, et al. Frequency of fractures in women with systemic lupus erythematosus: comparison with United States population data. Arthritis Rheum 1999;42(5):882–890.
8. Han J, Ren G, Xu Z, et al. Exploring the relationship between systemic lupus erythematosus and osteoporosis based on bioinformatics. Lupus 2022;31(2):163–177.
9. Jung JY, Choi ST, Park SH, et al. Prevalence of osteoporosis in patients with systemic lupus erythematosus: a multicenter comparative study of the World Health Organization and fracture risk assessment tool criteria. Osteoporos Sarcopenia 2020;6(4):173–178.
10. Becker A, Fischer R, Scherbaum WA, et al. Osteoporosis screening in systemic lupus erythematosus: impact of disease duration and organ damage. Lupus 2001;10(11):809–814.
11. Bischoff-Ferrari HA, Dietrich T, Orav EJ, et al. Positive association between 25-hydroxy vitamin D levels and bone mineral density: a population-based study of younger and older adults. Am J Med 2004;116(9):634–639.
12. Petri M, Orbai A-M, Alarcón GS, et al. Derivation and validation of the Systemic Lupus International Collaborating Clinics classification criteria for systemic lupus erythematosus. Arthritis Rheum 2012;64(8): 2677–2686.
13. 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.
14. Gladman DD, Goldsmith CH, Urowitz MB, et al. The Systemic Lupus International Collaborating Clinics/American College of Rheumatology (SLICC/ACR) Damage Index for systemic lupus erythematosus international comparison. J Rheumatol 2000;27(2):373–376.
15. Fütő L, Tőke J, Patócs A, et al. Skeletal differences in bone mineral area and content before and after cure of endogenous Cushing's syndrome. Osteoporos Int 2008;19(7):941–949.
16. van Houten P, Netea-Maier R, Wagenmakers M, et al. Persistent improvement of bone mineral density up to 20 years after treatment of Cushing's syndrome. Eur J Endocrinol 2021;185(2):241–250.
17. Bultink IE, Lems WF, Kostense PJ, et al. Prevalence of and risk factors for low bone mineral density and vertebral fractures in patients with systemic lupus erythematosus. Arthritis Rheum 2005;52(7):2044–2050.
18. Pineau CA, Urowitz MB, Fortin PJ, et al. Osteoporosis in systemic lupus erythematosus: factors associated with referral for bone mineral density studies, prevalence of osteoporosis and factors associated with reduced bone density. Lupus 2004;13(6):436–441.
19. Correa-Rodríguez M, Pocovi-Gerardino G, Callejas-Rubio J-L, et al. Disease damage accrual and low bone mineral density in female patients with systemic lupus erythematosus. Biol Res Nurs 2021; 23(4):575–583.
20. Petri M. Musculoskeletal complications of systemic lupus erythematosus in the Hopkins Lupus Cohort: an update. Arthritis Care Res 1995; 8(3):137–145.
21. MacKay DL, Kean TJ, Bernardi KG, et al. Reduced bone loss in a murine model of postmenopausal osteoporosis lacking complement component 3. J Orthop Res 2018;36(1):118–128.
22. Larsen ER, Mosekilde L, Foldspang A. Vitamin D and calcium supplementation prevents osteoporotic fractures in elderly community dwelling residents: a pragmatic population-based 3-year intervention study. J Bone Miner Res 2004;19(3):370–378.
23. Weaver CM, Alexander DD, Boushey CJ, et al. Calcium plus vitamin D supplementation and risk of fractures: an updated meta-analysis from the National Osteoporosis Foundation. Osteoporos Int 2016;27(1): 367–376.
24. Zhao JG, Zeng XT, Wang J, et al. Association between calcium or vitamin D supplementation and fracture incidence in community-dwelling older adults: a systematic review and meta-analysis. JAMA 2017;318(24):2466–2482.
25. Kahwati LC, Weber RP, Pan H, et al. Vitamin D, calcium, or combined supplementation for the primary prevention of fractures in community-dwelling adults: evidence report and systematic review for the US Preventive Services Task Force. JAMA 2018;319(15):1600–1612.
26. Petri M, Bello KJ, Fang H, et al. Vitamin D in systemic lupus erythematosus: modest association with disease activity and the urine protein-to-creatinine ratio. Arthritis Rheum 2013;65(7):1865–1871.
27. Ruiz-Irastorza G, Egurbide MV, Olivares N, et al. Vitamin D deficiency in systemic lupus erythematosus: prevalence, predictors and clinical consequences. Rheumatology (Oxford) 2008;47(6):920–923.
28. Vitamin D, Calcium, or Combined Supplementation for the Primary Prevention of Fractures in Community-Dwelling Adults: Preventive Medication. U.S. Preventive Services Task Force. Published April 17, 2018. https://www.uspreventiveservicestaskforce.org/uspstf/recommendation/vitamin-d-calcium-or-combined-supplementation-for-the-primary-prevention-of-fractures-in-adults-preventive-medication#:~:text=Recommendation%20Summary&text=The%20USPSTF%20recommends%20against%20daily,community%2Ddwelling%2C%20postmenopausal%20women.
29. Burt LA, Billington EO, Rose MS, et al. Adverse effects of high-dose vitamin D supplementation on volumetric bone density are greater in females than males. J Bone Miner Res 2020;35(12):2404–2414.

Arthritis Care & Research
Vol. 78, No. 6, June 2026, pp 749–760
DOI 10.1002/acr.25692

American College of Rheumatology
Empowering Rheumatology Professionals

Sarilumab in Polyarticular-Course Juvenile Idiopathic Arthritis: Dose-Finding and 1-Year Analysis of a Phase 2b, Open-Label, Multicenter Study

Fabrizio De Benedetti, MD, PhD,[1] Inmaculada C. Penadés, MD,[2] Irina Nikishina, MD, PhD,[3] Ivan Foeldvari, MD,[4] Alberto J. Spindler, MD,[5] Anna Kozlova, MD,[6] Nadina Rubio-Pérez, MD,[7] Pierre Quartier, MD,[8,9] Zbigniew Żuber, MD,[10] Raul Barria, MD,[11] Daniel Clemente, MD, PhD,[12] Gabriel V. Cornejo, MD,[13] Katherine Marzan, MD,[14] Nancy Liu, PhD,[15] Christine Xu, PhD,[15] Stephen DiMartino, MD, PhD,[16] Angeliki Giannelou, MD,[16#] Fei Cao-Ghoul, MD,[17#] Bolanle Akinlade, MD, MBA,[16] and Lydie Baret-Cormel, MD[17]

Objective. This study assessed sarilumab in treating patients with polyarticular-course juvenile idiopathic arthritis (pcJIA).

Methods. This phase 2b, open-label study (NCT02776735) consisted of three sequential parts (each with a core-treatment and extension phase). During part 1, three doses were assessed in two weight groups (group A/B: ≥30–60 kg/≥10 to <30 kg) to select the optimal dose with regard to pharmacokinetics, safety and efficacy were evaluated in latter parts. During the extension phase of part 1, patients initially assigned to the selected optimal dose continued on this dose; the remaining patients were switched to this selected dose. Patients in parts 2 and 3 received the selected dose from baseline. The primary endpoint was pharmacokinetic exposure (area under the serum concentration versus time curve during a dose interval τ of 2 weeks [$AUC_{0-\tau}$], maximum serum concentration observed [C_{max}], and concentrations observed before treatment administration during repeated dosing [C_{trough}]). Safety and efficacy were assessed.

Results. The mean age of treated patients (N = 101; 76.2% female) was 9.4 years. Of the evaluated doses in part 1, dose 2 (group A/B: 3.0/4.0 mg/kg every 2 weeks [q2w]) was selected. In patients receiving the selected dose from baseline (n = 73), C_{max}, $AUC_{0-14\ days}$ and C_{trough} at steady state in group A/B were 27.1/40.4 mg/L, 276/395-day*mg/L, and 9.57/14.4 mg/L, respectively. At week 48, the JIA–American College of Rheumatology 30%, 50%, 70%, and 90% improvement criteria rates were 100%, 100%, 93.8%, and 76.6%, respectively. Adverse events were reported in 70 of 73 patients (95.9%). Twenty-seven patients (37%) experienced grade 3/4 neutropenia; there were no associated infections. No deaths occurred.

Conclusion. In patients with pcJIA receiving the selected dose from baseline, pharmacokinetic exposure was comparable to a dose of 200 mg q2w for adults with rheumatoid arthritis. Clinically relevant improvements were observed in disease activity, with safety being consistent with the known profile of sarilumab.

INTRODUCTION

Juvenile idiopathic arthritis (JIA) includes all arthritis with onset before age 16 years and persistence for at least 6 weeks in the absence of any defined cause.[1] JIA includes seven subtypes defined by the International League of Associations for Rheumatology (ILAR). It is the most common childhood rheumatic disease with an estimated global prevalence and incidence

[#]At the time of study conduct and content finalization.

ClinicalTrials.gov identifier: NCT02776735.
Supported by Sanofi and Regeneron Pharmaceuticals, Inc.

[1]Ospedale Pediatrico Bambino Gesù IRCCS, ERN-RITA Center, Rome, Italy; [2]Hospital Universitario y Politécnico la Fe, Valencia, Spain; [3]V. A. Nasonova Research Institute of Rheumatology, Moscow, Russia; [4]Am Schoen Klinik Eilbek, Hamburg, Germany; [5]Centro Médico Privado de Reumatologia, Tucumán, Argentina; [6]Centre of Pediatric Hematology, Oncology and Immunology, Moscow, Russia; [7]Hospital Universitario "Dr. J. E. González," Universidad Autónoma de Nuevo León, Monterrey, Mexico; [8]Necker-Enfants Malades Hospital, Assistance Publique-Hopitaux de Paris, Paris, France; [9]Université Paris-Cité, Paris, France; [10]Andrzej Frycz Modrzewski Krakow University, Krakow, Poland; [11]Bioreuma, Concepción, Chile; [12]Hospital Infantil Universitario Niño Jesús, Madrid, Spain; [13]Clinica de Reumatologia y Enfermedades Autoinmunes, Guadalajara, Jalisco, Mexico; [14]Children's Hospital Los Angeles, Keck School of Medicine, University of Southern California, Los Angeles, California; [15]Sanofi, Morristown, New Jersey; [16]Regeneron Pharmaceuticals, Inc, Tarrytown, New York; [17]Sanofi, Paris, France.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25692).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25692.

Address correspondence via email to Fabrizio De Benedetti, MD, PhD, at fabrizio.debenedetti@opbg.net.

Submitted for publication January 22, 2025; accepted in revised form September 16, 2025.

SIGNIFICANCE & INNOVATIONS

- The 1-year results of this phase 2b study indicate that sarilumab, an interleukin-6 (IL-6) receptor inhibitor, administered at doses of 3.0 mg/kg every 2 weeks (q2w) and 4.0 mg/kg q2w for patients with polyarticular-course juvenile idiopathic arthritis (pcJIA) weighing at least 30 kg and at least 10 kg to less than 30 kg, respectively, yielded exposures that were comparable to the approved dose (200 mg q2w) of sarilumab in adult patients with rheumatoid arthritis (RA).
- There was a rapid and clinically relevant improvement in the signs and symptoms of pcJIA among patients as early as week 2 after initiation of sarilumab treatment, which was sustained until week 48.
- The safety profile of sarilumab observed in this study was consistent with its known profile in adult patients with RA and was similar to that of IL-6 inhibition in children and adults.

ranging from 3.8 to 400/100,000 and 1.6 to 23/100,000, respectively.[2–5]

Polyarticular-course JIA (pcJIA) is an umbrella term that includes rheumatoid factor (RF)–positive polyarticular JIA (pJIA), RF-negative pJIA, and extended oligoarticular JIA (eoJIA). Patients with pcJIA are at great risk of disability.[6] Conventional synthetic disease-modifying antirheumatic drugs (csDMARDs) are used as a first-line therapy for pcJIA. Per the 2019 American College of Rheumatology (ACR) guidelines, biologic DMARDs (bDMARDs) may be added to csDMARDs if treatment response is insufficient. The guidelines also recommend bDMARDs as the initial therapy in patients with high-risk factors that include involvement of a high-risk joint (eg, cervical spine, wrist, and hip) or high disease activity and/or those judged by their physician to be at significant risk of joint damage.[7] However, a recent real-world analysis (predominantly included patients with pcJIA) demonstrated that despite multiple approved treatments, 45% to 52% of patients had chronically uncontrolled disease despite treatment with bDMARDs.[8] This highlights an unmet clinical need in this population for newer treatment options to meet treatment goals of remission (ie, absence of disease activity) or, at least, of rapid and effective reduction in disease activity.[9]

Elevated levels of interleukin-6 (IL-6), a pleiotropic cytokine involved in chronic inflammation, have been reported in the serum and synovial fluid of patients with pcJIA, correlating with disease activity and degree of disability.[10–14] Sarilumab, a human monoclonal antibody, blocks the binding of IL-6 to both membrane-bound and soluble IL-6 receptor α (sIL-6Rα). It is approved for treating adult patients with moderately to severely active rheumatoid arthritis (RA),[15] adult patients with polymyalgia rheumatica,[15] and pediatric patients with active pJIA.[15,16] Targeting IL-6 with tocilizumab (a humanized monoclonal antibody to the IL-6R) has shown to be effective in patients with pcJIA.[17–20]

This open-label study (NCT02776735) investigated sarilumab as a treatment for pcJIA in children and adolescents aged 2 to 17 years. The core-treatment phase of dose-finding (part 1) aimed to identify the optimal sarilumab dose regimen (ie, dose); parts 2 and 3 further evaluate pharmacokinetics (PK), pharmacodynamics (PD), efficacy, and safety of the selected sarilumab dose in a larger patient group. Herein, we present results from the core-treatment phase of part 1 and the 1-year analysis (after last patient enrollment) of the selected dose in patients across all three parts of the study.

METHODS

Study design

This multicenter, phase 2b, open-label study was conducted across 15 countries and divided into three sequential parts. Each part included an initial 12-week core-treatment phase followed by an extension phase (Supplementary Figure S1A). To minimize exposure to ineffective doses, only patients who achieved at least a JIA-ACR 30% improvement criteria rate (JIA-ACR 30) response at week 12 were allowed to enter the extension phase in parts 1 and 2. However, as all patients achieved a JIA-ACR 30 response in the core-treatment phase of parts 1 and 2, this criterion was removed for part 3. The duration of the extension phase was 144 weeks for parts 1 and 2 and 84 weeks for part 3.

The clearance of monoclonal antibodies increases in a nonlinear manner relative to increasing body weight.[21] Therefore, given the range of body weights and ages in the pcJIA population, weight-adjusted dosages were considered in two groups (group A/B: ≥30–60 kg/≥10 to <30 kg). In part 1, three subcutaneous doses were evaluated per body weight group for a total of six planned patients per dose and weight group. Doses 1, 2, and 3 were proposed using PK modeling and simulation (Xu C, Zhang M, Liu Y, et al; unpublished observations)[22] to provide similar PK exposure to the doses of 150 mg every 2 weeks (q2w), 200 mg q2w, and 150 mg every week (qw), respectively, observed for adults with RA.[23] Patients in groups A/B received 2.0/2.5 mg/kg q2w (dose 1), or 3.0/4.0 mg/kg q2w (dose 2), or 2.0/2.5 mg/kg qw (dose 3). Dose escalation was conducted stepwise, starting with the lowest dose in the higher-weight group (group A). Initiation of group B was permitted after data review of at least three patients enrolled with the same dose in group A after at least 6 weeks of treatment (Supplementary Figure S1B). Escalation to higher doses within each weight group was permitted following a review of safety, efficacy, PK, and PD data, once the first three patients enrolled in the current weight and dose group had completed 6 weeks of treatment. Enrollment in dose 3 in group B was permitted after data review from all other weight and dose groups. Dose-escalation decisions were made by the

dose-escalation committee (DEC; made up of sponsors' representatives [experts in the field of medicine, pharmacovigilance, PK, and biostatistics] and the principal investigator). Safety was monitored throughout the study by an independent data-monitoring committee (DMC). After the 12-week core phase, patients in parts 1 and 2, who had a JIA-ACR 30 response at week 12 entered the 144-week extension phase, and patients in part 3 entered the 84-week extension phase. Based on health authority recommendations, additional patients at the selected dose were enrolled in parts 2 and 3 of the core-treatment phase to provide sufficient precision PK/PD parameters, PK-PD relationship assessments (part 2), and descriptive safety and efficacy analysis (part 3).

The study was conducted in accordance with consensus ethics principles including the Declaration of Helsinki and the International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use guidelines for Good Clinical Practice and all applicable laws, rules, and regulations. The study was approved by an independent ethics committee or institutional review board at each trial site (Supplementary Table S2).

Patients

Eligible patients were aged 2 to 17 years with a diagnosis of RF-positive or RF-negative pJIA or eoJIA per the ILAR criteria[1] with at least five active joints per ACR definition for "active arthritis" (swelling within joint not because of deformity or limitation of motion with pain or tenderness). Patients had to be inadequate responders to current treatment and considered candidates for bDMARDs per the investigators' judgment. Written informed consent was obtained from all patients/guardians, based on local regulation and guidelines.

Key exclusion criteria included prior treatment with an anti–IL-6/IL-6R antagonist, JAK inhibitor (within 4 weeks) or other biologics (within five half-lives) of first sarilumab dose, and use of parenteral/intra-articular glucocorticoid (GC) injection within 4 weeks before study baseline. Patients receiving nonsteroidal anti-inflammatory drugs, oral GCs, or csDMARD treatment were to be on a stable dose before entering the study (Supplementary Table S3).

Study endpoints

The primary endpoint was sarilumab PK exposure (maximum serum concentration [C_{max}] and area under serum concentration vs time curve [AUC_{0-c}]) following the first dose and serum concentration observed before treatment administration during repeated dosing (C_{trough}) from baseline to week 12. Secondary endpoints included PD, including serum concentration of high-sensitivity C-reactive protein (hs-CRP), IL-6, and sIL-6Rα; efficacy; and safety. Efficacy was determined by the proportion of patients achieving, at week 12, a JIA-ACR 30%, 50%, 70%, or 90% improvement criteria rates (JIA-ACR 30/50/70/90) response; change from baseline in JIA-ACR components; and change from baseline in juvenile arthritis disease activity score with 27-joint count (JADAS-27) with hs-CRP.

To provide corroboratory evidence of the treatment effects, more recent efficacy assessments and definitions of inactive disease for pcJIA from the 2019 ACR guidelines[7] were included before database lock: change in JADAS 10-joint count hs-CRP; clinical JADAS-27 (cJADAS-27); cJADAS-10 at weeks 6, 12, 24, and 48; and the proportion of patients with no active joints, JADAS-10 inactive disease (JADAS-10 ≤ 2.7), clinical inactive disease (CID) per 2004 Wallace criteria (defined as the absence of active joints and, of active uveitis, with a CRP level < 10 mg/L and no disease activity as per the physician's global assessment), or CID per Wallace criteria without systemic GC use and clinical remission (inactive disease per Wallace criteria for 6 consecutive months prior to the assessment visits) at week 48. Subgroup analysis was performed for JIA-ACR70 and cJADAS inactive disease activity (cJADAS-10 ≤ 2.5) at week 48 in patients receiving the selected dose from baseline with different JIA subtypes, with/without prior bDMARDs, and concomitant csDMARD use at baseline.

Statistical analysis

In the core-treatment phase of part 1, six patients per dose and weight group for a total of 36 patients were planned to be enrolled. No formal power/sample size calculations were performed for this assessment. However, this sample size was considered achievable and sufficient to describe PK and safety parameters. In part 2, an additional 12 patients per weight group were planned to enroll to achieve a total of 18 patients per weight group at the selected dose for sufficient precision to estimate PK parameters with the 95% confidence interval (CI) within 60% to 140% of the geometric mean. Considering a 15% dropout rate in the core-treatment phase, approximately 30 patients were planned for enrollment to the selected dose in part 2 to achieve a total of 60 patients in parts 1 and 2, including 36 patients at the selected dose. Approximately 28 patients were planned for enrollment to the selected dose in part 3 to achieve a total of 100 treated patients for the entire study, as recommended by the health authority.

A population PK (PopPK) model of pooled data was developed to describe the PK profile of sarilumab. Descriptive statistics for the first dose, week 12, and steady state exposures of interests (eg, C_{trough}, C_{max}, and AUC_t) was reported by dose regimen and by different covariate categories. Descriptive statistics were also used to summarize PD and safety data. For efficacy endpoints, data for participants who achieved a JIA-ACR 30/50/70/90/100 response for each dose, overall, and by body weight group were summarized by visit using counts,

proportions, and 95% CIs. Efficacy endpoints up to 1 year were analyzed using a "data as observed while on treatment approach" (no imputation of missing data). Additionally, during the 12-week core-treatment phase, JIA-ACR responses were analyzed using two additional approaches for handling treatment discontinuations and missing data (nonresponder imputation and last observation carried forward). Overall score and change from baseline in JADAS-27, JADAS-10, cJADAS-27, and cJADAS-10 were summarized by visit (including number, mean, SE, SD, median, minimum, and maximum) for each dose and body weight group. Statistical analysis was performed using SAS version 9.0 or later (SAS Institute Inc).

Data availability

Qualified researchers may request access to patient-level data and related study documents, including the clinical study report, study protocol with any amendments, blank case report form, statistical analysis plan, and dataset specifications. Patient-level data will be anonymized, and study documents will be redacted to protect the privacy of the trial participants. Further details on Sanofi's data-sharing criteria, eligible studies, and process for requesting access can be found at https://www.vivli.org/.

RESULTS

Patients

Of the 102 patients enrolled in all parts of the study, 101 received sarilumab and 86 patients completed at least 52 weeks of treatment (Figure 1). The overall median treatment duration was 1,078 days with a cumulative treatment exposure of 214.4 patient-years (PY) (Supplementary Table S4).

Most treated patients were female (76.2%) and had RF-negative pJIA (66.3%) with a mean age of 9.4 years (Table 1). At baseline, patients had high disease activity, with a mean cJADAS-10 score of 19.7, a mean active joint count of 15.8, and a mean hs-CRP level of 12.2 mg/L. Overall, at baseline, 85.1% and 19.8% patients were receiving csDMARDs and systemic GCs, respectively; 22.8% patients had received prior bDMARDs.

Core-treatment phase of part 1

In part 1, 42 patients were enrolled and 34 completed the 12-week core-treatment phase. PK evaluation of all three doses demonstrated nonlinear PK with target-mediated drug disposition of sarilumab. Sarilumab PK exposure in patients who received doses 1, 2, and 3 was comparable to the sarilumab doses of 150 mg q2w, 200 mg q2w, and 150 mg qw, respectively, for adult patients with RA; PK exposure was comparable between weight groups for all three doses. A decrease in mean hs-CRP concentration and an increase in mean sIL-6Rα concentration were greater with doses 2 and 3 than with dose 1 (Supplementary Figure S2). All three doses proved efficacious for decreasing signs and symptoms (as assessed by JIA-ACR 30 response) and disease activity (as assessed by JADAS-27 CRP) in both weight groups, with no dose cohort prematurely stopped because of a lack of efficacy. However, JIA-ACR responses were achieved less rapidly with dose 1 compared with doses 2 and 3 (Supplementary Figure S3). This was also evident when missing data were handled with a nonresponder imputation approach or with last observation carried forward (Supplementary Table S5). The incidence of adverse events (AEs) was comparable among all doses; however, protocol-specified extensive hematology monitoring identified a higher incidence of grade 3/4 neutropenia (absolute neutrophil count [ANC] < 1,000/mm^3) with dose 3 (Supplementary Table S6). None of these events were associated with infections or severe infections and all resolved within a few days. Hence, based on the totality of PK, PD, safety, and efficacy results of the core-treatment phase of part 1, dose 2 (thereafter designated as the selected dose) was selected for further evaluation. Following the selection of this dose, 20 patients who had been receiving doses 1 or 3 during the core-treatment phase of part 1 had their dose adjusted to the selected dose. Of 14 patients who received the selected dose in the core-treatment phase of part 1, 13 patients continued the selected dose during the extension phase. In the subsequent parts 2 and 3, 31 and 28 patients were treated with the selected dose, respectively. Hence, 73 patients were treated with the selected dose from baseline, and 93 patients had any exposure to the selected dose during the study. Eight patients discontinued treatment before the dose adjustment (Figure 1).

One-year analysis

PK. PK evaluation in patients on the selected dose from baseline up to week 48 demonstrated comparable PK exposure of sarilumab in both weight groups. The observed C_{trough} in patients on the selected dose was 2.24 mg/L to 3.10 mg/L at week 2 (after the first dose), 7.49 mg/L to 9.47 mg/L at week 12 (after completion of the core-treatment phase), and 11.6 mg/L to 4.2 mg/L at week 48. PK parameters of sarilumab estimated from the PopPK model after the first dose and at the steady state are provided in Table 2. C_{max}, $AUC_{0–14\ days}$ and C_{trough} at the steady state in group A/B was 27.1/40.4 mg/L, 276/395-day*mg/L, and 9.57/14.4 mg/L, respectively. Based on the PopPK model, in patients who received the selected dose from baseline, the accumulation ratios were approximately two- to four-fold as compared with the first dose for C_{trough}, C_{max}, and $AUC_{0–14\ days}$, respectively. In patients who had their dose adjusted to the selected dose, the observed C_{trough} reached levels comparable with those seen in patients originally assigned

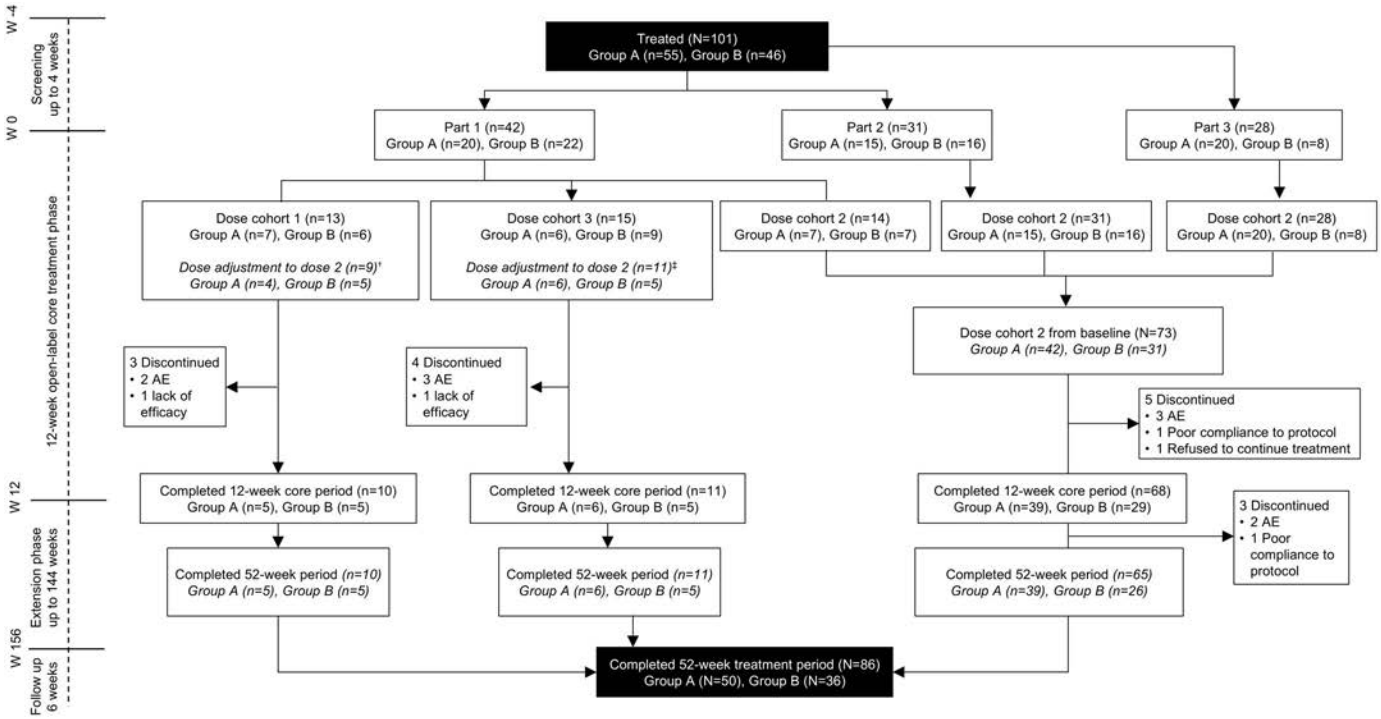


Figure 1. Patient disposition. Group A's weight group was at least 30 kg to 60 kg for patients enrolled in part 1 and at least 30 kg for patients enrolled in parts 2 and 3; Group B's weight group was 10 to less than 30 kg for all three parts. Dose 1 was 2.0 mg/kg every 2 weeks for the patients of group A and 2.5 mg/kg every 2 weeks for the patients of group B during the core-treatment phase of part 1; the patients were then switched to receive the selected dose 2. Dose 2 was 3.0 mg/kg every 2 weeks for the patients of group A and 4.0 mg/kg every 2 weeks for the patients of group B throughout the study period. Dose 3 was 2.0 mg/kg every week for the patients of group A and 2.5 mg/kg every week for the patients of group B during core-treatment phase of part 1; the patients were then switched to receive the selected dose 2. †The timing of the dose adjustment ranged from week 84 to week 120. One patient of group A discontinued treatment (because of an AE) after a 52-week treatment period before dose adjustment. ‡The timing of the dose adjustment ranged from week 32 to week 72. AE, adverse event.

to the selected dose after 12 weeks from the time of dose adjustment (data not shown).

PD. Median hs-CRP concentrations in patients on the selected dose from baseline decreased in both weight groups from week 2 onward and remained at a low concentration level (<2 mg/L) through week 48 (Supplementary Figure S4A). Median IL-6 and mean sIL-6Rα concentration increased from baseline to week 12 (Supplementary Figure S4B and C).

Efficacy. JIA-ACR responses (as observed while on treatment) in patients on the selected dose from baseline improved from week 2 onward (Figure 2A). JIA-ACR 30/50/70/90 response rates were observed in 100.0%/95.6%/80.9%/45.6% of the patients at week 12, respectively. JIA-ACR response rates continued to improve until week 48 with 100%/100%/93.8%/76.6% patients achieving JIA-ACR 30/50/70/90 response, respectively. Consistent with JIA-ACR responses, all JIA-ACR components (joints with active arthritis, joints with limited motion, physician and patient global visual analog scale, Childhood Health Assessment Questionnaire Disability Index, and hs-CRP) improved from week 2 onward (Supplementary Table S7). A decrease in disease activity as measured by JADAS-27 CRP was also observed from week 2 onward. Mean change in the JADAS-27 CRP score from baseline was −17.46 at week 12 and −20.75 at week 48 (Supplementary Figure S5). A decrease in the mean (SE) of cJADAS-27 CRP and cJADAS-10 was observed from baseline through week 48 (Figure 2B). Similar findings were observed for mean change in JADAS-10 CRP, cJADAS-10, and cJADAS-27 scores from baseline and mean JADAS-27 CRP and JADAS-10 scores (Figure 2C; Supplementary Figure S5A and B). At week 48, 60.9% and 51.6% patients on the selected dose from baseline achieved CID (per Wallace criteria) with no GC use and clinical remission, respectively (Figure 2D).

No differences were found in efficacy (JIA-ACR 70 and cJADAS ≤ 2.5) at week 48 following sarilumab treatment between patients with and without RF-positive JIA. Patients who had previously failed a bDMARD showed efficacy (JIA-ACR 70 and cJADAS ≤ 2.5) comparable to that of those who had not. Further, the use of csDMARDs did not appear to have an impact on the efficacy results, JIA-ACR 70, and cJADAS inactive disease (cJADAS-10 ≤ 2.5) (Figure 3).

Safety. Among 73 patients receiving the selected dose from baseline, 70 patients (95.9%) experienced at least one AE (Table 3). No deaths occurred during the study. The most frequently observed AEs were infections and neutropenia. The most

Table 1. Demographic and baseline clinical characteristics of the study population*

Parameters	Dose cohort 1[a]	Dose cohort 2 from baseline[b]	Dose cohort 3[c]	Any dose		
	Weight groups A + B (n = 13)[d]	Weight groups A + B (n = 73)[d]	Weight groups A + B (n = 15)[d]	Weight group A (n = 55)[d]	Weight group B (n = 46)[d]	Weight groups A + B (N = 101)[d]
Age, mean (SD), years	9.6 (4.3)	9.5 (4.7)	8.3 (5.2)	12.7 (3.0)	5.4 (2.9)	9.4 (4.7)
Sex (female), n (%)	10 (76.9)	58 (79.5)	9 (60.0)	43 (78.2)	34 (73.9)	77 (76.2)
Disease characteristics						
JIA subtype, n (%)						
RF-positive pJIA	4 (30.8)	13 (17.8)	2 (13.3)	13 (23.6)	6 (13.0)	19 (18.8)
RF-negative pJIA	9 (69.2)	48 (65.8)	10 (66.7)	32 (58.2)	35 (76.1)	67 (66.3)
eoJIA	0	12 (16.4)	3 (20.0)	10 (18.2)	5 (10.9)	15 (14.9)
Duration of pcJIA, mean (SD), years	3.4 (4.5)	2.5 (3.3)	2.5 (2.8)	3.4 (4.2)	1.7 (1.8)	2.6 (3.4)
Any prior bDMARD, n (%)	5 (38.5)	14 (19.2)	4 (26.7)	10 (23.8)	4 (12.9)	23 (22.8)
Baseline disease features						
cJADAS-10, mean (SD)	19.5 (4.4)	20.1 (4.1)	17.8 (4.9)	19.8 (4.0)	19.5 (4.7)	19.7 (4.3)
Active joint count, mean (SD)	15.8 (10.5)	16.1 (9.4)	14.7 (9.6)	18.1 (10.3)	13.2 (7.7)	15.8 (9.5)
hs-CRP, mean (SD), mg/L	5.7 (12.4)	14.5 (36.8)	6.5 (10.4)	9.6 (19.0)	15.3 (42.7)	12.2 (32.0)
Baseline treatments						
Any csDMARD, n (%)	11 (84.6)	62 (84.9)	13 (86.7)	49 (89.1)	37 (80.4)	86 (85.1)
MTX, n (%)	11 (84.6)	59 (80.8)	13 (86.7)	46 (83.6)	37 (80.4)	83 (82.2)
Systemic GC, n (%)	4 (30.8)	10 (13.7)	6 (40.0)	13 (23.6)	7 (15.2)	20 (19.8)
Prednisone-equivalent dose, mean (SD), mg/kg/day	0.26 (0.17)	0.20 (0.14)	0.19 (0.16)	0.18 (0.16)	0.26 (0.12)	0.21 (0.15)

* bDMARD, biologic disease-modifying antirheumatic drug; cJADAS-10, clinical juvenile arthritis disease activity score 10 joint count; csDMARD, conventional synthetic DMARD; eoJIA, extended oligoarticular JIA; GC, glucocorticoid; hs-CRP, high-sensitivity C-reactive protein; JIA, juvenile idiopathic arthritis; MTX, methotrexate; pcJIA, polyarticular-course JIA; pJIA, polyarticular JIA; RF, rheumatoid factor.
[a] Dose 1 was 2.0 mg/kg every 2 weeks for the patients of group A and 2.5 mg/kg every 2 weeks for the patients of group B during the core-treatment phase of part 1; the patients were then switched to receive the selected dose 2.
[b] Dose 2 was 3.0 mg/kg every 2 weeks for the patients of group A and 4.0 mg/kg every 2 weeks for the patients of group B throughout the study period.
[c] Dose 3 was 2.0 mg/kg every week for the patients of group A and 2.5 mg/kg every week for the patients of group B during core-treatment phase of part 1; the patients were then switched to receive the selected dose 2.
[d] Group A's weight group was at least 30 kg to 60 kg for patients enrolled in part 1 and at least 30 kg for patients enrolled in parts 2 and 3; Group B's weight group was 10 to less than 30 kg for all three parts.

Table 2. Sarilumab pharmacokinetics parameters (estimated by the population pharmacokinetic model) in serum following first administration and at steady state after repeated administration of three doses of sarilumab*

Parameters	Dose cohort 1[a]		Dose cohort 2 from baseline[b]		Dose cohort 3[c]	
	Weight group A[d]	Weight group B[d]	Weight group A[d]	Weight group B[d]	Weight group A[d]	Weight group B[d]
Following first dose administration	(n = 7)	(n = 6)	(n = 42)	(n = 31)	(n = 6)	(n = 9)
C_{max}, mg/L	7.61 (2.69)	10.3 (1.65)	15.9 (5.45)	24.0 (3.93)	11.3 (1.84)	10.4 (2.84)
$AUC_{0-\tau}$, day mg/L	54.2 (15.9)	69.2 (20.2)	138 (43.7)	205 (42.6)	62.0 (11.6)	55.3 (17.1)
C_{trough}, mg/L	0.481 (0.379)	0.419 (0.240)	2.38 (1.80)	3.62 (2.33)	6.98 (2.22)	5.64 (2.46)
At steady state following repeated dose administration	(n = 5)	(n = 5)	(n = 39)	(n = 24)	(n = 5)	(n = 5)
C_{max}, mg/L	11.1 (4.82)	14.0 (2.97)	27.1 (11.6)	40.4 (7.77)	37.6 (8.52)	28.9 (5.32)[e]
$AUC_{0-\tau}$, day mg/L	90.7 (30.8)	110 (40.9)	276 (121)	395 (101)	470 (122)	352 (64.8)[e]
C_{trough}, mg/L	1.36 (0.872)	1.39 (1.44)	9.57 (5.84)	14.4 (9.81)	28.0 (9.06)	21.6 (5.16)[e]

* All values are presented as mean (SD). $AUC_{0-\tau}$, area under the serum concentration versus time curve during a dose interval τ of 2 weeks; C_{max}, maximum serum concentration observed; C_{trough}, concentrations observed before treatment administration during repeated dosing.
[a] Dose 1 was 2.0 mg/kg every 2 weeks for the patients of group A and 2.5 mg/kg every 2 weeks for the patients of group B during the core-treatment phase part 1; the patients were then switched to receive the selected dose 2.
[b] Dose 2 was 3.0 mg/kg every 2 weeks for the patients of group A and 4.0 mg/kg every 2 weeks for the patients of group B throughout the study period.
[c] Dose 3 was 2.0 mg/kg every week for the patients of group A and 2.5 mg/kg every week for the patients of group B during the core-treatment phase part 1; the patients were then switched to receive the selected dose 2.
[d] Group A's weight group was at least 30 kg to 60 kg for the patients enrolled in part 1 and at least 30 kg for patients enrolled in parts 2 and 3. Group B's weight group was 10 to less than 30 kg for all three parts.
[e] Two patients with one or two missing doses (at week 28 or weeks 28 and 29) were included in the summary at steady state for the dose cohort 3 group B. After excluding these two patients, the mean (SD) of C_{max}, $AUC_{0-\tau}$, and C_{trough} was 29.8 (7.19) mg/L, 360 (89.2) mg·day/L, and 20.9 (7.19) mg/L in the other three patients, respectively.

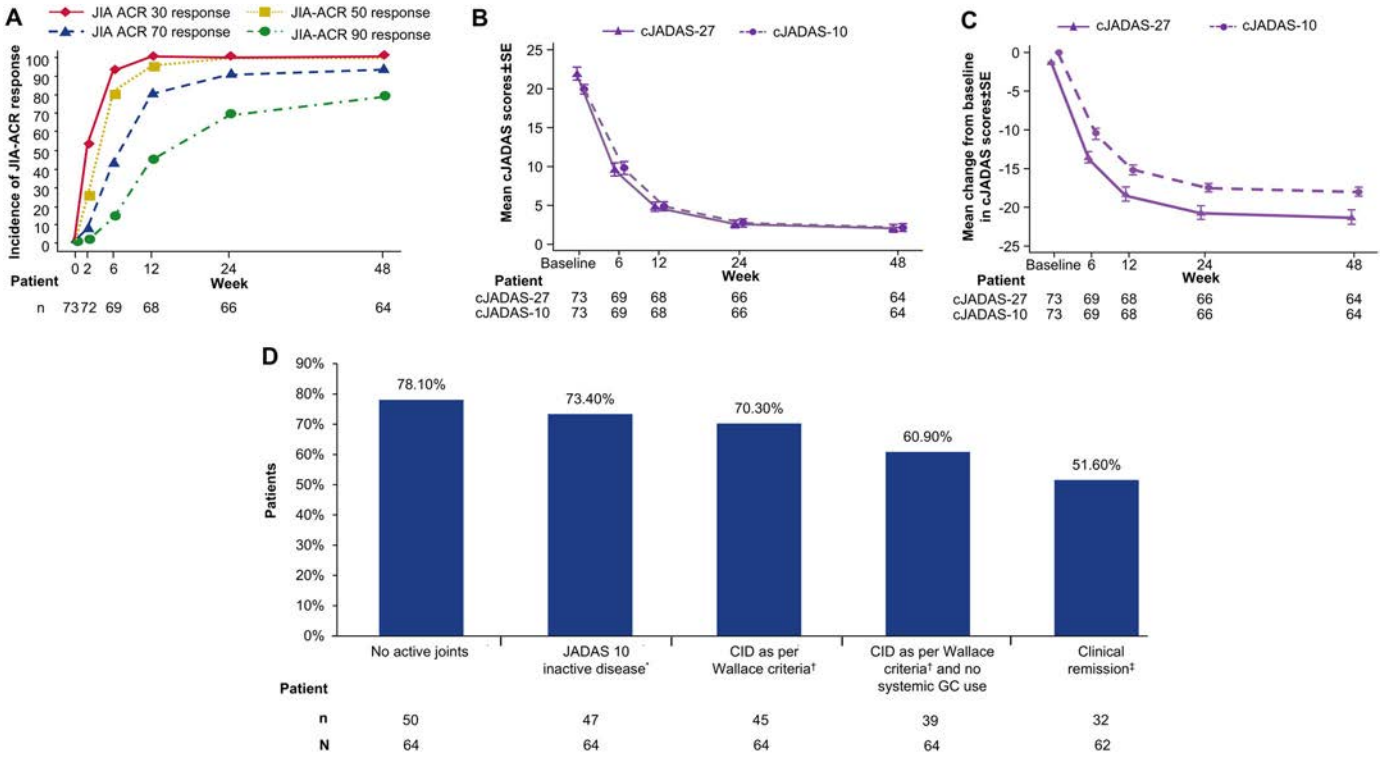


Figure 2. (A) Proportion of the patients achieving JIA-ACR 30/50/70/90 response (as observed while on treatment) up to week 48. (B) Change from baseline in cJADAS scores (as observed while on treatment) up to week 48. (C) Mean cJADAS scores (as observed while on treatment) up to week 48. (D) Proportions of patients with disease activity reduction criteria (observed cases) at week 48 in patients who received selected dose 2 from baseline. *JADAS-10 ≤ 2.7. †The Wallace criteria for CID are the absence of active joints and active uveitis, a normal C-reactive protein level (<10 mg/L), and no disease activity per the Physician's Global Assessment of Disease Activity (visual analog score ≤ 1.0). ‡Clinical remission was defined as CID per the Wallace criteria for 6 consecutive months prior to assessment visits. ACR 30/50/70/90, American College of Rheumatology 30%, 50%, 70%, and 90% improvement criteria rates; CID, clinically inactive disease; cJADAS, clinical juvenile arthritis disease activity score; GC, glucocorticoid; JIA, juvenile idiopathic arthritis; N, number of patients assessed for the parameters; n, number of patients with the outcomes.

frequent infections were nasopharyngitis, upper respiratory tract infections, gastroenteritis, and rhinitis. Six patients experienced nine serious AEs (SAEs). Two patients experienced two infectious SAEs (acute sinusitis and bone tuberculosis); the incidence of infectious SAEs was 1.2/100 PY (95% CI 0.15–4.50). All SAEs resolved during the study period except the event of bone tuberculosis (reported as Pott's disease) that occurred on day 4 in a patient who had negative tuberculosis test at screening that led to permanent treatment discontinuation. No SAE was considered related to the study treatment per the investigator's judgment. AEs in patients with any exposure to the selected dose (n = 93) are provided in Supplementary Table S8.

In patients receiving the selected dose from baseline, grade 3/4 neutropenia was identified in 27 patients (37%) (Group A = 10/42 [23.8%]; Group B = 17/31 [54.9%]) (Supplementary Table S9). Grade 3/4 neutropenia mostly occurred during the first 12 weeks of treatment and were reversible (median duration 22 days) (Supplementary Figure S6). In the overall population (N = 101), 85.7% of patients with normal ANC had infections compared with 85.4% with grade 3/4 neutropenia (Supplementary Table S10). The incidence of infections during neutropenic and nonneutropenic periods were 215.49 infections/100 PY (95% CI 117.81–361.55) and 161.95 infections/100 PY (95% CI 145.51–179.73), respectively (Supplementary Table S11). No serious or opportunistic infections occurred during the neutropenic period.

In patients with any exposure to the selected dose, 13 patients (14.0%) had at least one injection-site reaction, of which 12 (12.9%) were related to study treatment. All were of mild intensity and transient in nature with most resolving within 24 hours. Injection-site erythema (n = 9; 9.7%) was most frequently reported.

In patients who received the selected dose, positive and persistent antidrug antibody (ADA) response was detected in three patients (4.3%) and two patients (2.9%), respectively. There was no incidence of treatment-boosted ADAs.

DISCUSSION

The experimental doses of sarilumab tested in the core-treatment phase of part 1 were proposed using the modeling and simulation of sarilumab PK data from adult patients with RA

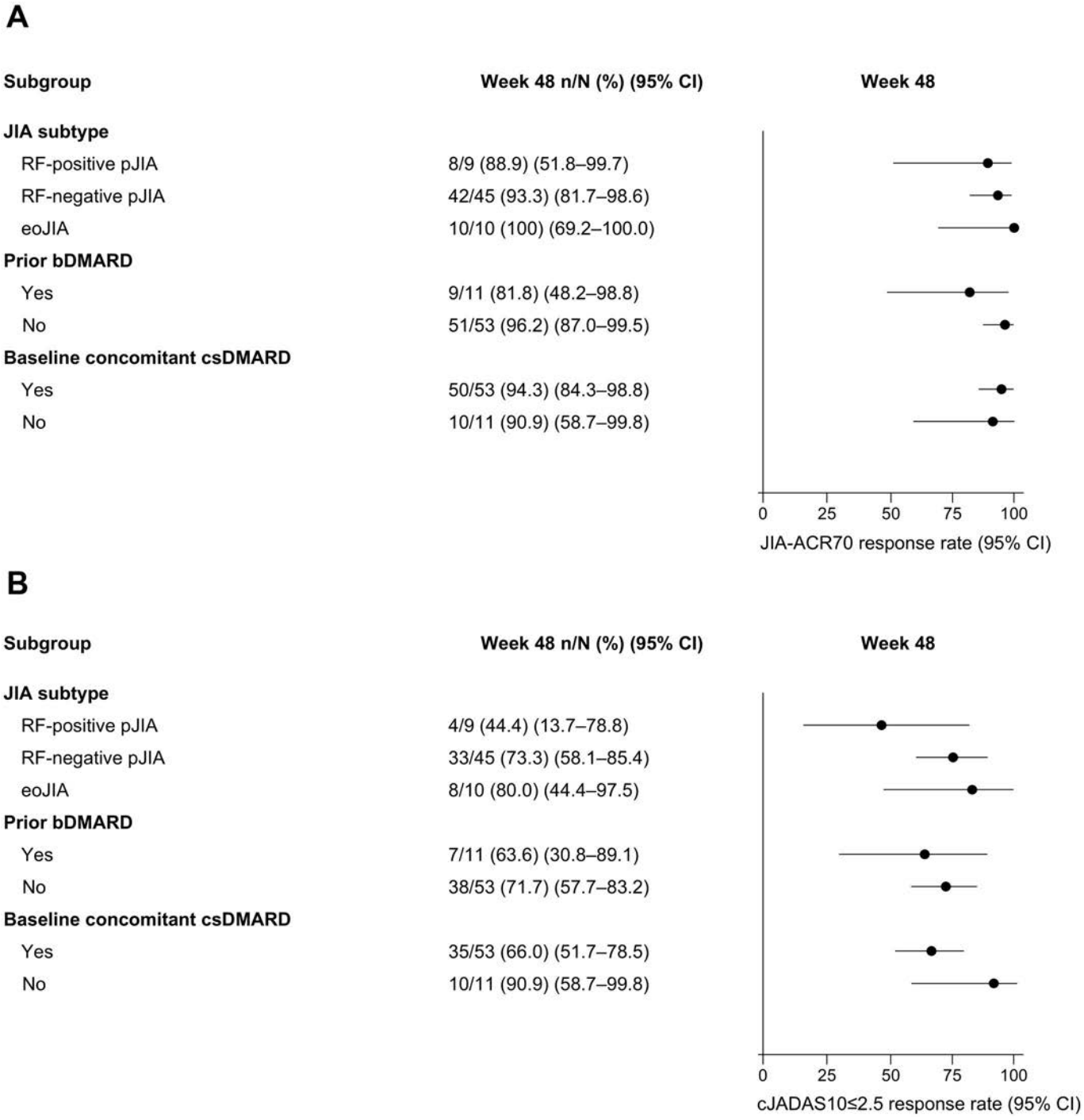


Figure 3. (A) JIA-ACR70 response rates and (B) cJADAS10 ≤ 2.5 in different subgroups of patients who received the selected dose 2 from baseline. ACR70, American College of Rheumatology 70% improvement criteria rate; bDMARD, biologic disease-modifying antirheumatic drug; CI, confidence interval; cJADAS, clinical juvenile arthritis disease activity score; csDMARD, conventional synthetic DMARD; eoJIA, extended oligoarticular juvenile idiopathic arthritis; pJIA, polyarticular JIA; RF, rheumatoid factor.

(Xu C, Zhang M, Liu Y, et al; unpublished observations).[22] The results of the core-treatment phase of part 1 demonstrated that sarilumab exhibited nonlinear PK with target-mediated drug disposition, and systemic exposure to sarilumab was similar between the two weight groups with each evaluated dose. Body weight–adjusted doses yielded comparable exposure in both weight groups and were comparable to the fixed doses tested in adults with RA. However, PD effects (greater suppression of CRP and greater increase of sIL-6Rα) and efficacy (with a rapid and numerically high level of response, ie, JIA-ACR) were more

Table 3. Investigator-reported AEs in patients who received selected dose 2 from baseline during the AE period*

Patients with AEs	Dose cohort 2 from baseline[a]		
	Weight group A (n = 42)[b]	Weight group B (n = 31)[b]	Weight groups A + B (n = 73)[b]
Any AEs	40 (95.2) [292.9]	30 (96.8) [542.5]	70 (95.9) [397.2]
Infectious AEs	30 (71.4) [91.9]	28 (90.3) [193.7]	58 (79.5) [134.5]
Serious AEs[c]	3 (7.1) [4.3]	3 (9.7) [7.5]	6 (8.2) [5.6]
Infectious serious AEs	1 (2.4) [1.1]	1 (3.2) [1.5]	2 (2.7) [1.2]
AEs leading to death	0	0	0
AE leading to treatment discontinuation	1 (2.4) [1.1]	6 (19.4) [8.9]	7 (9.6) [4.4]
Infections (bone tuberculosis)	0	1 (3.2) [1.5]	1 (1.4) [0.6]
Neutropenia	1 (2.4) [1.1]	4 (12.9) [6.0]	5 (6.8) [3.1]
Increase ALT	0	1 (3.2) [1.5]	1 (1.4) [0.6]
AEs of special interest	3 (7.1) [5.3]	10 (32.3) [26.8]	13 (17.8) [14.3]
Grade 4 neutropenia or any neutropenia leading to permanent treatment discontinuation[d]	3 (7.1) [5.3]	7 (22.6) [22.4]	10 (13.7) [12.5]
Increased ALT leading to permanent treatment discontinuation	0	1 (3.2) [1.5]	1 (1.4) [0.6]
Opportunistic infection (herpes zoster)	0	1 (3.2) [1.5]	1 (1.4) [0.6]
Tuberculosis	0	1 (3.2) [1.5]	1 (1.4) [0.6]
Common AEs with an overall incidence of ≥10%			
Nasopharyngitis	11 (26.2) [22.5]	16 (51.6) [51.6]	27 (37.0) [35.5]
Neutropenia	10 (23.8) [24.6]	15 (48.4) [108.8]	25 (34.2) [59.8]
Upper respiratory tract infections	5 (11.9) [9.6]	3 (9.7) [8.9]	8 (11.0) [9.3]
Gastroenteritis	4 (9.5) [5.3]	8 (25.8) [14.9]	12 (16.4) [9.3]
Rhinitis	4 (9.5) [4.6]	3 (9.7) [8.9]	7 (9.6) [6.2]

* Values are the n (%) [N_E per 100 patient-years]. AE, adverse event; ALT, alanine aminotransferase; N_E, exposure-adjusted event rate (time available for the entire study).
[a] Dose 2 was 3.0 mg/kg every 2 weeks for the patients of group A and 4.0 mg/kg every 2 weeks for the patients of group B throughout the study period.
[b] Group A's weight group was at least 30 kg to 60 kg for the patients enrolled in part 1 and at least 30 kg for patients enrolled in parts 2 and 3. Group B's weight group was 10 to less than 30 kg for all three parts.
[c] Six patients experienced nine serious AEs; each serious AE was reported once, and none of them were related to the study drug. The reported serious AEs included bone tuberculosis, pancreatic pseudocyst, acute pancreatitis, inguinal hernia, tonsillar hypertrophy, ligament rupture, meniscus injury, juvenile idiopathic arthritis, and acute sinusitis.
[d] All those with neutropenia recovered within a median duration of 14 days after treatment discontinuation.

pronounced with doses 2 and 3 versus dose 1. Notably, although AEs were balanced for all three doses, grade 3/4 neutropenia were numerically higher with dose 3 than the other doses. Hence, following the review of PK, PD, efficacy, and safety data during the core-treatment phase of part 1, dose 2 was selected as the optimal sarilumab dose for further evaluation. This dose achieved similar exposure to and mirrored the PD and efficacy profile of the equivalent dose of 200 mg q2w in adults with RA.[23–25] The q2w dosing schedule of the selected dose reduces injection burden, which is particularly relevant in children and allows for longer time for ANC recovery.

The present study enrolled a population with severe (a significant proportion of the patients had failed at least one bDMARD) and long-lasting disease (>2.5 years). Furthermore, at baseline, patients had high disease activity as documented by widespread arthritis (mean number of active joints of 16) and by the high proportion of patients receiving concomitant csDMARDs or systemic GCs at baseline. The results of 1-year analysis showed improvement in JIA-ACR response from week 2 onward. Importantly, in patients receiving the selected dose from baseline, clinically meaningful improvement in sign and symptoms (JIA-ACR 70 response) was noted in 80.9% of patients at week 12, which improved to 93.6% at week 48. A previous study with tocilizumab reported a JIA-ACR 70 response in 62.2% and 64.6% patients at weeks 16 and 40, respectively.[17] The high affinity of sarilumab to the IL-6Rα receptor (dissociation constant values for monomeric and dimeric IL-6Rα receptor were 61.9 and 12.8 pM, respectively) might be a possible explanation for higher efficacy of sarilumab for pcJIA treatment.[26] At week 48, approximately two-thirds of the patients achieved CID (assessed by JADAS and the Wallace criteria), with most successfully discontinuing systemic GCs. More than half of the patients achieved clinical remission. These observations show that sarilumab treatment for pcJIA meets two important considerations emphasized by current international guidelines: (1) achieve clinical remission or at least minimal/low disease activity and (2) avoid long-term GC use.[7,27]

In this study, improvement in the JIA-ACR 70 response and cJADAS-10 ≤ 2.5 was consistent across all JIA subtypes, including RF-positive pJIA, which is often more difficult to treat.[28] It is well-documented that patients with failed bDMARD therapies for pcJIA are less likely to achieve good clinical outcomes with subsequent targeted therapies.[29–31] Notably, in this study, the proportion of patients achieving a JIA-ACR 70 response and

cJADAS10 ≤ 2.5 were comparable with patients with or without a prior failure of a bDMARD. Furthermore, both the efficacy outcomes consistently improved irrespective of concomitant csDMARD treatment. This is clinically relevant as a significant proportion of patients (~50%) develop csDMARD intolerance.[32] Our study was not powered to assess differences in these subgroups; therefore, these results should be interpreted with caution, and validation from real-world studies is needed.

The selected dose of sarilumab was well-tolerated. The safety profile was consistent with the known safety profile of anti–IL-6 agents in adults and children, with no new safety concerns.[17,24,25,33] Most reported AEs were infections and neutropenia. The incidence of infectious AEs was 146.6/100 PY. The infections were mainly limited to the upper airway, which are common in pediatric populations,[34] and were balanced across weight groups. The incidence of infectious SAEs was 1.2/100 PY, which is low compared with the occurrence rates, which were 4.9 and 5.2/100 PY for other anti–IL-6R agents.[17,20]

Per the study protocol, ANC was systematically assessed leading to the detection of grade 3/4 neutropenia in 37% of patients. In adult patients treated with sarilumab or tocilizumab, the ANC nadir occurred a few days after first treatment administration, after which ANC values recovered, consistent with class effects.[35] Importantly, most patients with grade 3/4 neutropenia belonged to the lower-weight group. An evident trend toward increased rates of neutropenia in younger, lower-weight patients with pcJIA treated with tocilizumab was observed.[36,37] No definitive mechanism is available to explain why younger patients are more prone to neutropenia; however, in general, a greater prevalence of neutropenia has been reported in children.[38] All grade 3/4 neutropenia resolved within a few days. Incidence rates of infections during neutropenic or nonneutropenic periods were comparable; notably, no serious infections occurred during the neutropenic periods. This observation is consistent with data obtained from children and adults treated with tocilizumab or sarilumab, supporting the conclusion that neutropenia induced by IL-6 inhibition does not increase risk of infection.[17,33,35,36,39]

The limitations of this study include the open-label design with no control arm of placebo or an active comparator. This was because of the ethical concerns of exposing children to a noneffective treatment, given the potential chronic and severe nature of the disease and the presence of available data on IL-6 inhibition in adults and children with arthritis.[17,20,40] These issues were reflected in the study design, with a sequential approach for the tested doses, the involvement of a DMC and DEC to monitor safety and efficacy of tested doses, and stringent hematology monitoring throughout the study. This was also in relation to the strategy for developing the drug using the pediatric extrapolation concept introduced by the US Food and Drug Administration, which allows the extrapolation of efficacy from adults to children for a similar disease course and treatment response.

Sarilumab exhibited nonlinear PK with target-mediated drug disposition in patients with pcJIA. A dose of 3.0 mg/kg q2w in patients at least 30 kg and 4.0 mg/kg q2w in patients at least 10 kg to less than 30 kg achieved exposures similar to those of the approved dose of 200 mg q2w in adults with RA and was selected as the optimal sarilumab dose in pcJIA. The 1-year analysis results reiterated PK and PD findings in a larger cohort and showed that the dose was efficacious for pcJIA treatment with more than half of the patients achieving CID off GCs. No new safety signals were detected. Events of neutropenia were frequent and well-established for IL-6 inhibitors and were not associated with infections.

ACKNOWLEDGMENTS

The authors thank the patients who participated in this study, the co-investigators (listed in Supplementary Table S1), the DMC members (Dr Ross Petty, Dr Laura Schanberg, and Dr Ronald Laxer), and the staff at participating centers for their contributions to the study.

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 De Benedetti 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

Medical writing assistance for this manuscript was provided by Arohi Sarang, MPharm; Sambuddha Das, PhD; and Kavita Garg, PhD, CMPP, from Sanofi. The study was sponsored by Sanofi and Regeneron Pharmaceuticals, Inc. Authors affiliated with Sanofi and Regeneron Pharmaceuticals, Inc. at the time of the study had a role in the study design, analysis, or interpretation of the data as well as in critically revising the manuscript, as stated in the authors contributions. Publication of this article was contingent on the approval of Sanofi and Regeneron Pharmaceuticals, Inc.

REFERENCES

1. 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.
2. Thierry S, Fautrel B, Lemelle I, et al. Prevalence and incidence of juvenile idiopathic arthritis: a systematic review. Joint Bone Spine 2014; 81(2):112–117.
3. Oberle EJ, Harris JG, Verbsky JW. Polyarticular juvenile idiopathic arthritis - epidemiology and management approaches. Clin Epidemiol 2014;6:379–393.
4. Dave M, Rankin J, Pearce M, et al. Global prevalence estimates of three chronic musculoskeletal conditions: club foot, juvenile idiopathic

arthritis and juvenile systemic lupus erythematosus. Pediatr Rheumatol Online J 2020;18(1):49.

5. Harrold LR, Salman C, Shoor S, et al. Incidence and prevalence of juvenile idiopathic arthritis among children in a managed care population, 1996-2009. J Rheumatol 2013;40(7):1218–1225.
6. Webb K, Wedderburn LR. Advances in the treatment of polyarticular juvenile idiopathic arthritis. Curr Opin Rheumatol 2015;27(5): 505–510.
7. Ringold S, Angeles-Han ST, Beukelman T, et al. 2019 American College of Rheumatology/Arthritis Foundation guideline for the treatment of juvenile idiopathic arthritis: therapeutic approaches for non-systemic polyarthritis, sacroiliitis, and enthesitis. Arthritis Care Res (Hoboken) 2019;71(6):717–734.
8. Brunner HI, Schanberg LE, Kimura Y, et al; PRCSG Advisory Council and the CARRA Registry Investigators. New medications are needed for children with juvenile idiopathic arthritis. Arthritis Rheumatol 2020; 72(11):1945–1951.
9. Ravelli A, Consolaro A, Horneff G, et al. Treating juvenile idiopathic arthritis to target: recommendations of an international task force. Ann Rheum Dis 2018;77(6):819–828.
10. De Benedetti F, Robbioni P, Massa M, et al. Serum interleukin-6 levels and joint involvement in polyarticular and pauciarticular juvenile chronic arthritis. Clin Exp Rheumatol 1992;10(5):493–498.
11. Spîrchez M, Samaşca G, Iancu M, et al. Relation of interleukin-6, TNF-alpha and interleukin-1alpha with disease activity and severity in juvenile idiopathic arthritis patients. Clin Lab 2012;58(3-4):253–260.
12. De Benedetti F. Targeting interleukin-6 in pediatric rheumatic diseases. Curr Opin Rheumatol 2009;21(5):533–537.
13. Mangge H, Kenzian H, Gallistl S, et al. Serum cytokines in juvenile rheumatoid arthritis. Correlation with conventional inflammation parameters and clinical subtypes. Arthritis Rheum 1995;38(2): 211–220.
14. de Jager W, Hoppenreijs EP, Wulffraat NM, et al. Blood and synovial fluid cytokine signatures in patients with juvenile idiopathic arthritis: a cross-sectional study. Ann Rheum Dis 2007;66(5):589–598.
15. Kevzara. Prescribing information. Accessed May 15, 2025. Sanofi; 2024. https://www.accessdata.fda.gov/drugsatfda_docs/label/2024/761037s015lbl.pdf
16. European Medicines Agency. Kevzara summary of product characteristics. Accessed May 15, 2025. https://www.ema.europa.eu/en/documents/product-information/kevzara-epar-product-information_en.pdf
17. Brunner HI, Ruperto N, Zuber Z, et al; Paediatric Rheumatology International Trials Organisation PRINTO; Pediatric Rheumatology Collaborative Study Group (PRCSG). Efficacy and safety of tocilizumab in patients with polyarticular-course juvenile idiopathic arthritis: results from a phase 3, randomised, double-blind withdrawal trial. Ann Rheum Dis 2015;74(6):1110–1117.
18. Opoka-Winiarska V, Żuber Z, Alexeeva E, et al. Long-term, interventional, open-label extension study evaluating the safety of tocilizumab treatment in patients with polyarticular-course juvenile idiopathic arthritis from Poland and Russia who completed the global, international CHERISH trial. Clin Rheumatol 2018;37(7):1807–1816.
19. Malattia C, Ruperto N, Pederzoli S, et al; Paediatric Rheumatology International Trials Organisation (PRINTO) and the Pediatric Rheumatology Collaborative Study Group (PRCSG). Tocilizumab may slow radiographic progression in patients with systemic or polyarticular-course juvenile idiopathic arthritis: post hoc radiographic analysis from two randomized controlled trials. Arthritis Res Ther 2020;22(1):211.
20. Brunner HI, Ruperto N, Zuber Z, et al; Paediatric Rheumatology International Trials Organisation (PRINTO) and the Pediatric Rheumatology Collaborative Study Group (PRCSG). Efficacy and safety of tocilizumab for polyarticular-course juvenile idiopathic arthritis in the open-label two-year extension of a phase III trial. Arthritis Rheumatol 2021; 73(3):530–541.
21. Gill KL, Machavaram KK, Rose RH, et al. Potential sources of inter-subject variability in monoclonal antibody pharmacokinetics. Clin Pharmacokinet 2016;55(7):789–805.
22. Xu C, Zhang M, Liu Y, et al. Modeling, Simulation, and extrapolation to optimize pediatric development plan: Sarilumab Polyarticular-Course juvenile arthritis application. J Pharmacokinet Pharmacodyn 2026.
23. Huizinga TW, Fleischmann RM, Jasson M, et al. Sarilumab, a fully human monoclonal antibody against IL-6Rα in patients with rheumatoid arthritis and an inadequate response to methotrexate: efficacy and safety results from the randomised SARIL-RA-MOBILITY Part A trial. Ann Rheum Dis 2014;73(9):1626–1634.
24. Fleischmann R, van Adelsberg J, Lin Y, et al. Sarilumab and non-biologic disease-modifying antirheumatic drugs in patients with active rheumatoid arthritis and inadequate response or intolerance to tumor necrosis factor inhibitors. Arthritis Rheumatol 2017;69(2): 277–290.
25. Genovese MC, Fleischmann R, Kivitz AJ, et al. Sarilumab plus methotrexate in patients with active rheumatoid arthritis and inadequate response to methotrexate: results of a phase III Study. Arthritis Rheumatol 2015;67(6):1424–1437.
26. Xu C, Rafique A, Potocky T, et al. Differential binding of sarilumab and tocilizumab to IL-6Rα and effects of receptor occupancy on clinical parameters. J Clin Pharmacol 2021;61(5):714–724.
27. Ringold S, Weiss PF, Beukelman T, et al; American College of Rheumatology. 2013 update of the 2011 American College of Rheumatology recommendations for the treatment of juvenile idiopathic arthritis: recommendations for the medical therapy of children with systemic juvenile idiopathic arthritis and tuberculosis screening among children receiving biologic medications. Arthritis Rheum 2013;65(10):2499–2512.
28. Guzman J, Oen K, Tucker LB, et al; ReACCh-Out investigators. The outcomes of juvenile idiopathic arthritis in children managed with contemporary treatments: results from the ReACCh-Out cohort. Ann Rheum Dis 2015;74(10):1854–1860.
29. Triaille C, Quartier P, De Somer L, et al. Patterns and determinants of response to novel therapies in juvenile and adult-onset polyarthritis. Rheumatology (Oxford) 2024;63(3):594–607.
30. Kearsley-Fleet L, Heaf E, Davies R, et al; BCRD and BSPAR-ETN study groups. Frequency of biologic switching and the outcomes of switching in children and young people with juvenile idiopathic arthritis: a national cohort study. Lancet Rheumatol 2020;2(4):e217–e226.
31. Mannion ML, Amin S, Balevic S, et al. Comparative effectiveness of a second tumor necrosis factor inhibitor versus a non-tumor necrosis factor biologic in the treatment of patients with polyarticular-course juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2024;76(8): 1090–1098.
32. 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.
33. Burmester GR, Lin Y, Patel R, et al. Efficacy and safety of sarilumab monotherapy versus adalimumab monotherapy for the treatment of patients with active rheumatoid arthritis (MONARCH): a randomised, double-blind, parallel-group phase III trial. Ann Rheum Dis 2017; 76(5):840–847.
34. Jin X, Ren J, Li R, et al. Global burden of upper respiratory infections in 204 countries and territories, from 1990 to 2019. EClinicalMedicine 2021;37:100986.
35. Emery P, Rondon J, Parrino J, et al. Safety and tolerability of subcutaneous sarilumab and intravenous tocilizumab in patients with rheumatoid arthritis. Rheumatology (Oxford) 2019;58(5):849–858.

36. Pardeo M, Wang J, Ruperto N, et al; Paediatric Rheumatology International Trials Organisation (PRINTO) and the Pediatric Rheumatology Collaborative Study Group (PRCSG). Neutropenia during tocilizumab treatment is not associated with infection risk in systemic or polyarticular-course juvenile idiopathic arthritis. J Rheumatol 2019; 46(9):1117–1126.
37. European Medicines Agency. RoActemra (tocilizumab) assessment report. Published September 7, 2023. Accessed May 15, 2025. https://www.ema.europa.eu/en/documents/withdrawal-report/withdrawal-assessment-report-roactemra_en.pdf
38. Hsieh MM, Everhart JE, Byrd-Holt DD, et al. Prevalence of neutropenia in the U.S. population: age, sex, smoking status, and ethnic differences. Ann Intern Med 2007;146(7):486–492.
39. Moots RJ, Sebba A, Rigby W, et al. Effect of tocilizumab on neutrophils in adult patients with rheumatoid arthritis: pooled analysis of data from phase 3 and 4 clinical trials. Rheumatology (Oxford) 2017;56(4):541–549.
40. Burmester GR, Strand V, Kivitz AJ, et al. Long-term safety and efficacy of sarilumab with or without background csDMARDs in rheumatoid arthritis. Rheumatology (Oxford) 2023;62(10):3268–3279.

Arthritis Care & Research
Vol. 78, No. 6, June 2026, pp 761–770
DOI 10.1002/acr.25700

American College of Rheumatology
Empowering Rheumatology Professionals

Pharmacokinetics, Pharmacodynamics, and Safety of Subcutaneous Belimumab in Pediatric Patients With Systemic Lupus Erythematosus: A Multicenter, Open-Label Trial

Hermine I. Brunner,[1] Jordi Antón,[2] Inmaculada Calvo-Penadés,[3] Richard Dimelow,[4] Gerd Horneff,[5,6] Sylvia Kamphuis,[7] Rebecca Marino,[8] Andre van Maurik,[4] Kirsten Minden,[9] Masaaki Mori,[10,11] Josephine Ocran-Appiah,[12] Christel Wilkinson,[4] Yuichi Yamasaki,[13] and Deepak Assudani,[14] for the Pediatric Rheumatology Collaborative Study Group (PRCSG) and Paediatric Rheumatology International Trials Organisation (PRINTO) Network investigators

Objective. This study aimed to characterize the pharmacokinetics, pharmacodynamics, safety, and exploratory efficacy of subcutaneous belimumab in pediatric patients with active systemic lupus erythematosus (SLE) receiving standard therapy.

Methods. This single-arm, multicenter, open-label trial (GSK study 200908; ClinicalTrials.gov identifier: NCT04179032) used three-weight-band subcutaneous dosing of belimumab 200 mg every week (qw) for pediatric patients weighing ≥50 kg, every 10 days for pediatric patients weighing 30 to <50 kg, and every 2 weeks (q2w) for pediatric patients weighing 15 to <30 kg. The pharmacokinetic profile was characterized by observed concentration at week 12 and population pharmacokinetics (popPK) estimates derived from concentrations over 52 weeks. Pharmacodynamics, safety, and exploratory efficacy (≥4-point reduction from the baseline Safety of Estrogens in Lupus Erythematosus National Assessment - Systemic Lupus Erythematosus Disease Activity Index [SELENA-SLEDAI] score) were descriptively evaluated. Alternative two-weight-band subcutaneous dosing (≥40 kg: 200 mg qw; 15 to <40 kg: 200 mg q2w) was simulated to predict pharmacokinetics for this regimen.

Results. Patients weighing 30 to <50 kg had lower observed belimumab concentrations than those weighing ≥50 kg (week 12 geometric mean 82.8 vs 134 μg/mL), but popPK analyses predicted the three weight bands to be generally consistent and in alignment with established adult subcutaneous and pediatric intravenous exposures. The pharmacodynamic and safety profile was consistent with known belimumab effects. By week 52, 18 of 22 patients (81.8%) had a ≥4-point reduction from baseline SELENA-SLEDAI scores. Hypothetical two-weight-band dosing simulations predicted consistent exposure across weight bands and in line with established exposures in SLE.

Conclusion. Exposure following subcutaneous belimumab administration in pediatric patients is consistent with approved usage; these findings, along with consistent safety and efficacy data, support subcutaneous belimumab use for pediatric patients with SLE.

ClinicalTrials.gov identifier: NCT04179032.

Supported by GSK.

[1]Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio; [2]Hospital Sant Joan de Déu, Universitat de Barcelona, Barcelona, Spain; [3]Hospital Universitario y Politécnico La Fe, València, Spain; [4]GSK, Stevenage, United Kingdom; [5]Asklepios Klinik Sankt Augustin GmbH, Sankt Augustin, Germany; [6]University Hospital of Cologne, Cologne, Germany; [7]Erasmus University Medical Center, Rotterdam, the Netherlands; [8]GSK, Durham, North Carolina; [9]Charité Universitätsmedizin Berlin, Berlin, Germany; [10]Institute of Science Tokyo, Tokyo, Japan; [11]St Marianna University School of Medicine, Kawasaki, Japan; [12]GSK, Philadelphia, Pennsylvania; [13]Kagoshima University Hospital, Kagoshima, Japan; [14]GSK, London, United Kingdom.

Please refer to GSK weblink to access GSK's data sharing policies and seek anonymized patient-level data as applicable via the link https://www.gsk-studyregister.com/en/.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25700).

Author disclosures and graphical abstract are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25700.

Address correspondence via email to Hermine I. Brunner, MD, MSc, MBA, at Hermine.brunner@cchmc.org.

Submitted for publication March 14, 2025; accepted in revised form October 15, 2025.

SIGNIFICANCE & INNOVATIONS

- This was the first trial to assess subcutaneous belimumab, a B-cell modulator that selectively inhibits B-lymphocyte stimulator, in pediatric patients with systemic lupus erythematosus (SLE) receiving standard therapy.
- Our findings of broadly consistent exposure across pediatric weight ranges, in line with established adult subcutaneous and pediatric intravenous exposures, along with consistent safety and efficacy data, support subcutaneous belimumab use for pediatric patients with SLE.

INTRODUCTION

Childhood-onset systemic lupus erythematosus (cSLE) is a complex autoimmune disease. Patients often have major organ involvement and accrue irreversible organ damage from disease activity and medication side effects.[1–5] cSLE is associated with an approximate 18-fold increased mortality risk compared with the general population.[6] Compared with adult-onset SLE (aSLE), cSLE is associated with a higher risk of kidney involvement, and patients with cSLE are more likely to have received dialysis or a transplant than those with aSLE.[6–8] Because of the increased disease severity in cSLE, more intensive pharmacological management is required compared with aSLE.[7,9] In clinical practice, this may mean high cumulative glucocorticoid exposure,[2,9] which, considering the recognized risk of steroid-associated organ damage and childhood growth delay, is a concern.[10–13] As such, steroid-sparing therapeutic strategies are needed to help prevent flares and minimize organ damage accrual in cSLE.[2,4,9,14]

Belimumab is a human monoclonal antibody and B-cell modulator that targets the central immunopathogenic pathway in lupus, selectively inhibiting soluble B-lymphocyte stimulator and reducing autoreactive B cells that drive disease activity.[15,16] Intravenous (IV) belimumab has been approved since 2019 as add-on therapy for patients at least five years of age with active, autoantibody-positive SLE and since 2022 for active lupus nephritis.[17–20] A phase 2 randomized controlled trial, PLUTO (GSK study BEL114055; ClinicalTrials.gov identifier: NCT01649765), established the efficacy and safety profile of IV belimumab for pediatric patients with cSLE.[17,18,21] To date, real-world evidence in cSLE is limited; however, emerging data support belimumab's effectiveness in reducing glucocorticoid exposure and suggest potential for improvements in disease activity.[22–24]

Belimumab has also been available in a subcutaneous (SC) formulation for adults with active SLE in combination with standard therapy since 2017, providing an alternative route of administration to the IV route.[17,25] This study aimed to characterize the pharmacokinetics (PK), pharmacodynamics (PD), safety, and exploratory efficacy of repeat doses of SC belimumab in pediatric patients who are also receiving standard therapy for cSLE.

PATIENTS AND METHODS

Study design. This was a single-arm, multicenter, open-label trial of belimumab 200 mg SC in pediatric patients with active SLE receiving standard therapy (GSK study number 200908; ClinicalTrials.gov identifier: NCT04179032; trial protocol available from ClinicalTrials.gov[26]). This study was conducted at 11 sites in 7 countries (Argentina, Germany, Japan, Mexico, the Netherlands, Spain, and the United States). The trial comprised an initial 12-week, open-label treatment phase (part A) with an optional 40-week, open-label continuation phase for any patient who completed part A (part B; weeks 12–52). An additional optional access extension phase for eligible patients who completed part B is ongoing.

Three SC belimumab dose groups were defined by baseline body weight, with the proposed weight thresholds (15 to <30 kg, 30 to <50 kg, and ≥50 kg) approximately corresponding to the age groups 5 to <9 years, 9 to <14 years, and 14 years and older, according to the Centers for Disease Control and Prevention clinical growth charts.[27] Informed by PK modeling from prior studies of belimumab in cSLE and aSLE,[28,29] patients weighing 15 to <30 kg were to receive belimumab 200 mg SC every 2 weeks (q2w) (no patients enrolled in this weight group), patients weighing 30 to <50 kg received belimumab 200 mg SC every 10 days (q10d), and patients weighing ≥50 kg received belimumab 200 mg SC every week (qw), per the approved adult dose. Belimumab dosage was not changed during part A (weeks 1–12). During part B (weeks 12–52), dosing was adjusted if the patient's weight category changed from baseline.

Exposure, PD, safety, and exploratory efficacy data were collected through part A and part B (52 weeks), with follow-up assessments at 8 weeks (exposure, PD, and safety) and 16 weeks (safety only) after the final dose of belimumab for those not continuing to the optional access extension phase. In part A, trough PK samples were taken on days 8, 15, 29, 57, and 85 for patients weighing ≥50 kg or 15 to <30 kg, whereas patients weighing 30 to <50 kg provided samples on days 11, 21, 31, 61, and 81. Approximately 28 patients were required to be enrolled in part A to ensure that 24 patients were evaluable at week 12 to characterize PK. The sample size of 24 participants and the specified sampling schedule allowed estimation of central clearance and volume of distribution with a 95% confidence interval within 60% and 140% of their point estimates,[30] with a power of 99.9% and 99.7%, respectively.

Patients. Patients were male or female, were 5 to 17 years of age, and weighed ≥15 kg at enrollment. All patients had four or more of the 1997 revised American College of Rheumatology criteria for SLE at screening or previously[31] and had active SLE disease, defined as a Safety of Estrogens in Lupus Erythematosus National Assessment - Systemic Lupus Erythematosus Disease Activity Index (SELENA-SLEDAI) score ≥6 at screening, and

documented antinuclear antibody titer ≥1:80, and/or an anti–double-stranded DNA (anti-dsDNA) antibody level ≥30 IU/mL.

Patients were required to have a protocol-defined stable SLE standard therapy for ≥30 days before the first dose of belimumab, with no new SLE therapies added during this time period. Patients who had previous treatment with belimumab, B cell–targeted therapy or abatacept, or any biologic investigational agent within 364 days of belimumab initiation were excluded. Those who had certain medical conditions or a history of them were excluded, including those with severe central nervous system lupus, renal replacement therapy within 90 days of belimumab initiation, an estimated glomerular filtration rate <30 mL/min/1.73 m^2, and acute severe nephritis. A full list of the eligibility criteria is provided in Supplementary Table 1.

Patients provided age-appropriate assent, and their parent or legally appointed representative provided written informed consent before the initiation of any study procedures or study-specific data collection. Patients who turned 18 years of age during the study provided their own informed consent.

Demographic parameters, including date of birth, sex, race, and ethnicity, were captured at screening. Race and ethnicity were self-reported from a fixed set of categories, and more than one category could be chosen.

End points. The primary end points were observed belimumab concentrations at week 12, and PK parameters using population PK (popPK) model-derived estimates for area under the curve (AUC) over a dosing interval at steady state (AUC_{ss}); average belimumab concentration at steady state ($Cavg_{ss}$; AUC_{ss} normalized by the dosing interval); maximum belimumab concentration at steady state ($Cmax_{ss}$); and minimum belimumab concentration at steady state ($Cmin_{ss}$).

The secondary end points were changes from baseline in anti-dsDNA, complement component 3 (C3), C4, IgA, IgG, IgM, and B-cell subset levels at weeks 12 and 52, and the incidence of adverse events (AEs), serious AEs (SAEs), and AEs of special interest (AESIs) during the study. The percentage of patients with a change in status from baseline anti-dsDNA positivity, low C3 or C4 levels, and normal/high Ig levels were also assessed as exploratory analyses.

Efficacy was an exploratory end point, assessed as the proportion of patients with a ≥4-point reduction in their baseline SELENA-SLEDAI score at weeks 12 and 52. A decrease in prednisone use from baseline at weeks 12 and 52 was an additional exploratory analysis.

Data analysis. Descriptive statistics were used to summarize the observed belimumab concentrations at week 12 and steady-state exposures derived from popPK analysis. Approximate steady state in PK by week 12 was expected based on previous terminal half-life estimates of 16.3 days in pediatric patients and 18 days for adults.[28,29] To derive popPK exposures, a popPK data set was compiled from all SC belimumab-treated patients in this study and all IV belimumab–treated pediatric patients from the previous PLUTO study.[21] A model was fitted to this popPK data set, informed by previous popPK models developed for adult SC and pediatric IV dosing.[28,29] The model was two-compartmental with first order absorption, distribution, and elimination (Supplementary Table 2). Belimumab exposures were extrapolated for patients weighing 15 to ≤30 kg, as no patients were enrolled in this weight band. For descriptive comparison, popPK-derived belimumab exposure distributions from adult patients receiving belimumab 200 mg SC in the BLISS-SC study (GSK Study BEL112341; ClinicalTrials.gov identifier: NCT01484496)[25] were calculated, as well as popPK-derived exposure distributions for pediatric patients receiving belimumab 10 mg/kg IV in the PLUTO study (GSK study BEL114055; ClinicalTrials.gov identifier: NCT01649765).[21]

PD was descriptively assessed by percentage change from baseline levels of anti-dsDNA, C3, C4, IgA, IgG, IgM, and B-cell subsets at weeks 12 and 52. The proportion of patients at week 52 with a change from baseline from positive anti-dsDNA to negative anti-dsDNA, low C3 or C4 levels to normal/high levels, and normal/high IgA, IgG, or IgM levels to low levels was described. Anti-dsDNA positivity was defined as ≥30 IU/mL, a normal/high C3 level was defined as ≥90 mg/dL, and a normal/high C4 was defined as ≥13 mg/dL. Normal/high and low Ig levels were defined by the lower limit of normal for a patient's age.

The safety and tolerability profile was evaluated descriptively using the incidence of treatment-emergent AEs, SAEs, AESIs, and AEs leading to treatment discontinuation throughout the study, based on the Medical Dictionary for Regulatory Activities Version 25.1 classification.[32] AESIs included malignancies, postinjection systemic reactions (including hypersensitivity reactions), infections of special interest (opportunistic infections, herpes zoster, tuberculosis, sepsis), depression, suicide/self-injury, and death. AEs were evaluated as mild, moderate, or severe by the investigators.

Ethics. This study was conducted in compliance with the Declaration of Helsinki and was approved by the relevant national, regional, or investigational center ethics committees or institutional review boards (Supplementary Table 3), in accordance with International Council for Harmonisation of Technical Requirements of Pharmaceuticals for Human Use Good Clinical Practice and applicable country-specific requirements. Results are reported following applicable components of the Consolidated Standards of Reporting Trials guidelines.

Simulated two-weight-band SC dosing regimen. Weight-based dosing was used in this study to accommodate for marked differences in the weight of children of different ages. Because the q10d dosing regimen investigated in this study may be more difficult to adhere to than a qw or q2w regimen, which would have a consistent day of the week of dosing, a hypothetical

SLE pediatric population with two-weight-band qw or q2w SC dosing was simulated to predict steady-state PK parameters for this alternative regimen using the popPK model described previously. The simulated dosing regimen was 200 mg SC qw for patients weighing ≥40 kg or 200 mg SC q2w for patients weighing 15 to <40 kg. Further information can be found in the Supplementary Methods and Supplementary Table 2.

RESULTS

Patient population. Of 28 patients screened, 25 were enrolled in the study, making up the intent-to-treat population (Supplementary Figure 1). Those not enrolled either did not meet the eligibility criteria (n = 2) or were excluded based on investigator's discretion (n = 1). Patients were enrolled between November 2019 and September 2021. All 25 patients completed part A of the study up to week 12, and 92.0% (n = 23/25) completed part B up to week 52. Of the two patients who withdrew from the study before week 52, one withdrew due to AEs, and one withdrew at the investigator's discretion. Of patients completing part B, 47.8% (n = 11/23) have entered the optional access extension phase of the study, which is ongoing.

Baseline characteristics are summarized in Table 1. Most patients were female (n = 21/25; 84.0%) and White (n = 16/25; 64.0%), with a mean age at screening of 14 (SD 2.1) years. No patients of the lowest weight category (15 to <30 kg) were enrolled. Most patients were anti-dsDNA antibody positive at baseline (n = 16/25; 64.0%). Baseline SELENA-SLEDAI score components are detailed in Supplementary Table 4.

The mean duration of belimumab exposure was 357.8 (SD 43.0) days through week 52. Among patients weighing 30 to <50 kg, the mean number of injections over 52 weeks was 33.8 (SD 6.59), with a minimum of 15 and a maximum of 39. Among patients weighing ≥50 kg, the mean number of injections over 52 weeks was 51.2 (SD 2.23), with a minimum of 46 and a maximum of 53.

PK. At week 12 and through 52 weeks, mean belimumab concentrations were generally lower in the 30 to <50 kg cohort compared with the ≥50 kg cohort (Table 2; Supplementary Figure 2). However, popPK model simulations of exposure versus body weight demonstrated that steady-state exposures for the three-weight-band SC regimens are expected to be consistent across the pediatric weight range, with data extrapolated to the lowest weight band (Figure 1A). PopPK-derived exposures for pediatric IV belimumab patients in the PLUTO study and adult SC belimumab patients in the BLISS-SC study are reported in Supplementary Table 5 and Figure 1 for comparative purposes, showing similar average concentrations at steady state ($Cavg_{ss}$) across the pediatric weight range.

Table 1. Patient baseline characteristics (ITT population)*

	All patients (N = 25)
Female, n (%)	21 (84.0)
Age, mean (SD), y	14.0 (2.1)
Age range, min–max, y	10–17
Ethnicity, n (%)	
Hispanic or Latino	11 (44.0)
Not Hispanic or Latino	14 (56.0)
Race, n (%)	
American Indian or Alaska Native	3 (12.0)
Asian	4 (16.0)
Black or African American	1 (4.0)
Multiracial	1 (4.0)
Native Hawaiian or other Pacific Islander	0
White	16 (64.0)
Body weight cohorts, n (%)	
≥50 kg	13 (52.0)
30 to <50 kg	12 (48.0)
15 to <30 kg[a]	0
Body weight range, min–max, kg	34.5–78.5
SELENA-SLEDAI score, mean (SD)	9.5 (3.0)
Laboratory measures	
Anti-dsDNA antibody positive (≥30 IU/mL), n (%)	16 (64.0)
Anti-dsDNA antibody levels, median (IQR),[b] IU/mL	524.0 (94.5–915.5)
ANA positive (titer ≥1:80), n (%)	24 (96.0)
Anti-dsDNA and/or ANA positive, n (%)	24 (96.0)
Low C3 levels (<90 mg/dL), n (%)	11 (44.0)
C3 levels, mean (SD), mg/dL	95.9 (23.7)
Low C4 levels (<13 mg/dL), n (%)	16 (64.0)
C4 levels, mean (SD), mg/dL	12.5 (6.9)
Low IgA (less than lower limit of normal), n (%)	0
Low IgG (less than lower limit of normal), n (%)	0
Low IgM (less than lower limit of normal), n (%)	4 (16.0)
Total B cells, median (IQR), $CD19^+$ cells/mL	233,545.0 (143,582.0–445,690.0)[c]
Concomitant medication	
Glucocorticoids, n (%)	21 (84)
Average daily prednisone[d] dose category, n (%)	
0 mg/day	4 (16)
>0 to ≤7.5 mg/day	9 (36)
>7.5 mg/day	12 (48)
Antimalarials, n (%)	25 (100)
Immunosuppressants, n (%)	21 (84)

* ANA, antinuclear antibody; anti-dsDNA, anti–double-stranded DNA; IQR, interquartile range; ITT, intent-to-treat; SELENA-SLEDAI, Safety of Estrogens in Lupus Erythematosus National Assessment - Systemic Lupus Erythematosus Disease Activity Index.
[a] No patients weighing 15 to <30 kg were enrolled.
[b] Among patients who were positive for anti-dsDNA antibody (n = 16).
[c] n = 18.
[d] Glucocorticoids were converted to prednisone-equivalent doses.

PD and biomarkers. Among those patients who were anti-dsDNA antibody positive at baseline (n = 16), anti-dsDNA antibody levels generally decreased by week 52, with a median change of −58.9% (interquartile range [IQR] −83.00% to −13.16%) at week 52 (Figure 2A). At week 52, 7.1% (n = 1/14)

Table 2. Observed week 12 and popPK-derived steady-state exposures of study participants and simulated steady-state exposures of a pediatric patient population for the two-weight band SC regimen*

	Pediatric patients (Study 200908)[a]			Simulated pediatric patients[b]	
	Total (N = 25)	≥50 kg (n = 13)	30 to <50 kg (n = 12)	≥40 kg (n = 4,484)	15 to <40 kg (n = 5,516)
Belimumab dosage	200 mg SC qw/q10d	200 mg SC qw	200 mg SC q10d	200 mg SC qw	200 mg SC q2w
Geometric mean (CV%), range or prediction interval					
Week 12 concentration, μg/mL	106 (49.3), 27.3–224	134 (34.7), 65.3–224	82.8 (49.3), 27.3–155	Not calculated	Not calculated
$Cmin_{ss}$, μg/mL	112 (41.7), 38.7–203	138 (26.2), 89.0–203	90.1 (43.9), 38.7–159	119 (38.6), 56.8–246	73.7 (44.6), 31.8–172
$Cavg_{ss}$, μg/mL	124 (37.3), 47.9–214	146 (25.2), 95.6–214	103 (39.8), 47.9–173	129 (36.8), 63.4–256	94.8 (37.8), 47.6–198
$Cmax_{ss}$, μg/mL	131 (34.9), 54.1–220	151 (24.7), 99.4–220	111 (37.6), 54.1–182	134 (35.9), 67.7–263	110 (34.8), 58.0–217
AUC_{ss}, μg·day/mL	1,027 (32.1), 479–1,733	1,024 (25.2), 669–1,498	1,031 (39.8), 479–1,733	899 (36.8), 444–1,794	1,328 (37.8), 665–2,768

* AUC_{ss}, area under the curve over a dosing interval at steady state; $Cavg_{ss}$, average concentration at steady state (AUC_{ss} normalized by the dosing interval); $Cmax_{ss}$, maximum concentration at steady state; $Cmin_{ss}$, minimum concentration at steady state; CV, coefficient of variation; popPK, population pharmacokinetics; q2w, every 2 weeks; q10d, every 10 days; qw, every week; SC, subcutaneous.
[a] Geometric mean (CV%) is presented with range.
[b] Geometric mean (CV%) is presented with prediction interval.

of patients who were anti-dsDNA antibody positive at baseline were anti-dsDNA antibody negative (Supplementary Figure 3A). C3 levels generally increased across the 52 weeks for those with low levels at baseline (n = 11), with a median increase of +23.5% (IQR –2.96% to +29.94%) (Figure 2B). Likewise, for patients with low C4 at baseline (n = 16), a median increase of +67.5% (IQR +31.00% to +92.86%) at week 52 was observed (Figure 2C). At week 52, 45.5% (n = 5/11) of patients with low C3 levels at baseline had normal/high C3 levels, and 50.0% (n = 7/14) of patients with low C4 levels at baseline had normal/high C4 levels (Supplementary Figure 3B and C).

Median percentage changes in Igs at week 52 were as follows: IgA, –13.1% (IQR –22.93% to –5.27%); IgG, –12.3% (IQR –22.82% to –5.82%); and IgM, –30.8% (IQR –41.45% to –25.06%) (Supplementary Figure 4). In patients with normal/high Ig levels at baseline, IgA and IgG levels at week 52 were above the lower limit of normal, whereas 16.7% (n = 3/18) of patients had reduced IgM levels to below the lower limit of normal (Supplementary Figure 3D–F). Total B cells had a median change of –61.3% (IQR –74.47% to –14.08%) from baseline levels at week 52, $CD20^+$ B cells had a median change of –61.7% (IQR –75.04% to –11.74%), and naïve B cells had a median change of –72.1% (IQR –85.15% to –50.53%) (Supplementary Figure 5A–C). Circulating memory B cells had a large initial median increase from baseline at week 2 of +120.8% (IQR +49.22% to +180.95%; Supplementary Figure 5D). Levels subsequently decreased from that initial surge but remained at +27.1% (IQR –33.02% to +111.19%) above baseline at week 52. Plasmablasts had a median change of –70.1% (IQR –80.22% to +68.32%) from baseline at week 52 (Supplementary Figure 5E).

Safety. Throughout the study, 88.0% (n = 22/25) of patients had at least one treatment-emergent AE (Table 3). The most common organ system class involved was infections and infestations, reported in 72.0% (n = 18/25) of patients. The most common AEs by preferred term were COVID-19 infection (n = 9/25, 36.0%; all mild or moderate severity), injection-site pain (n = 4/25, 16.0%; all mild severity), leukopenia (n = 4/25, 16.0%; all mild severity), and neutropenia (n = 4/25, 16.0%; all mild or moderate severity). One patient discontinued belimumab treatment due to decreased lymphocyte, neutrophil, and white blood cell counts. One SAE was reported: hospitalization for a COVID-19 infection. Of the AESIs, there were three patients who experienced mild post-injection systemic reactions (drug hypersensitivity, erythema, and injection-site urticaria; n = 1 for each); all patients recovered. There were no serious post-injection systemic reaction AESIs, and no post-injection systemic reactions led to the discontinuation of belimumab. There were no infections of special interest or cases of depression or suicide/self-injury. No deaths occurred during the reported study period.

Efficacy. At week 12, 66.7% (16/24) of patients achieved a ≥4-point reduction from the baseline SELENA-SLEDAI score; this increased to 81.8% (18/22) at week 52. Among patients in the ≥50 kg cohort (qw dosing), 75.0% (9/12) achieved a ≥4-point reduction from the baseline SELENA-SLEDAI score at week 12, increasing to 92.3% (n = 12/13) at week 52. In the 30 to <50 kg cohort (q10d dosing), 58.3% (7/12) of patients achieved a ≥4-point reduction from the baseline SELENA-SLEDAI score at week 12, with 66.7% (n = 6/9) meeting this criterion at week 52. Among patients using prednisone at baseline, 33.3% (7/21) had decreased prednisone use at week 12, and 52.6% (10/19) had decreased use by week 52.

Simulated two-weight-band SC dosing regimen. Belimumab exposures were generally consistent across the

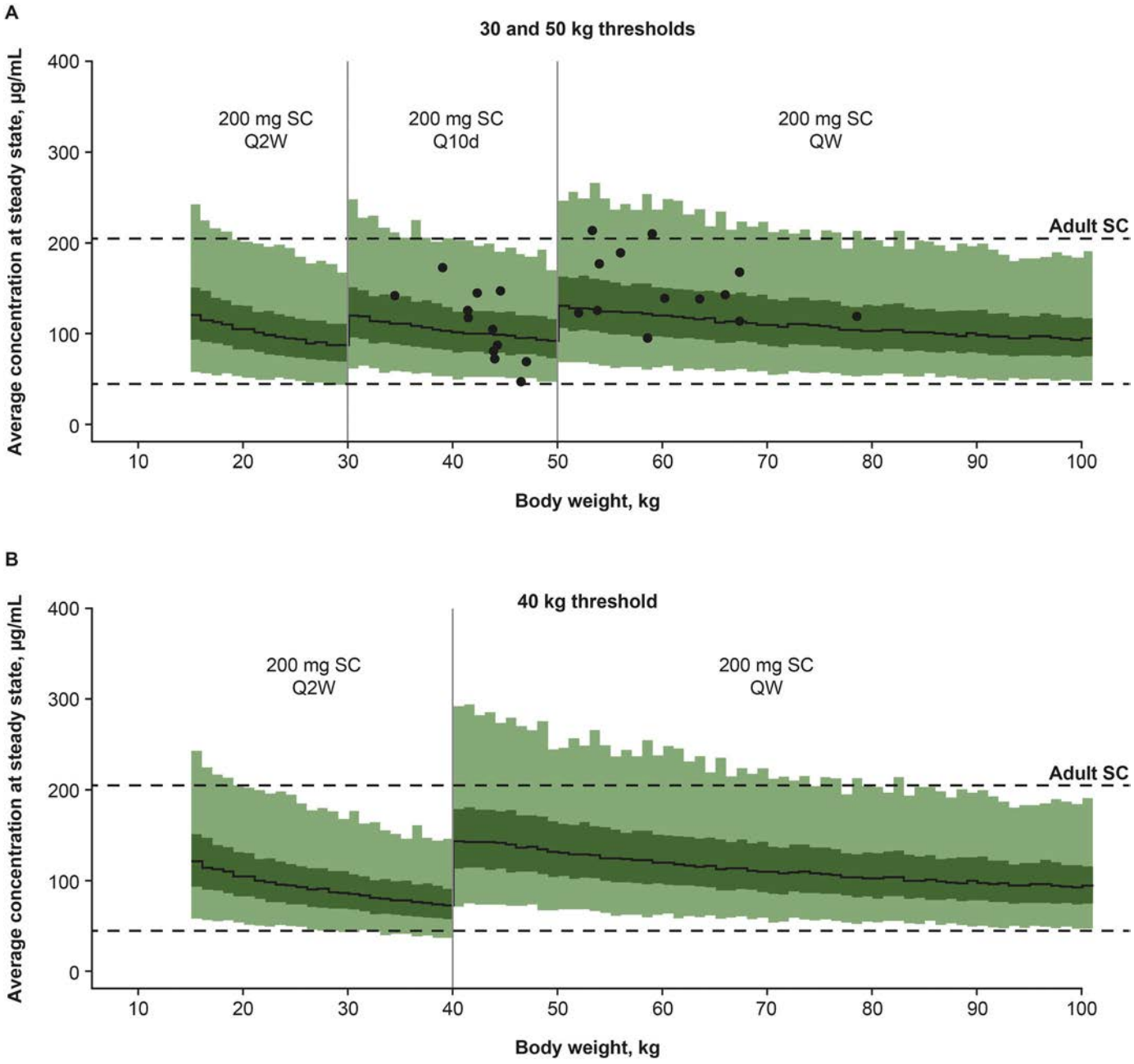


Figure 1. Predicted $Cavg_{ss}$ versus body weight for (A) the three-weight-band SC regimen studied and (B) the alternative two-weight-band SC regimen. (A) Individual predicted $Cavg_{ss}$ of patients in the study 200908 (N = 25) are shown as black data points. (A and B) Median predicted $Cavg_{ss}$ is shown as a solid black line, interquartile range of $Cavg_{ss}$ is shown in dark green, and 95% prediction interval of $Cavg_{ss}$ is shown in light green. Weight thresholds are shown with a gray vertical line. One thousand pediatric patients were simulated per 1 kg body weight interval. The $Cavg_{ss}$ 95% prediction interval of adult patients (Adult SC) in the BLISS-SC study (GSK Study BEL112341; ClinicalTrials.gov identifier: NCT01484496) is superimposed with horizontal dashed lines. No patients were enrolled in the 15 to <30 kg weight group in study 200908. $Cavg_{ss}$, average concentration at steady state (area under the curve at steady state normalized by the dosing interval); qw, every week; q2w, every 2 weeks, q10d, every 10 days; SC, subcutaneous.

weight ranges for the simulated two-weight-band SC dosing regimen (Figure 1B; Table 2). The expected range of steady-state exposure ($Cavg_{ss}$) was slightly larger for the simulated two-weight-band SC dosing regimen than for the three-weight-band SC dosing regimen investigated in the present study (Figure 1).

PK simulations demonstrated that when switching from IV to SC belimumab for the two-weight-band SC regimen, exposure is maintained when the first SC dose is administered one to four weeks after the last IV infusion (Supplementary Figure 6). In particular, steady-state exposure can be expected from the first SC dose when administered one or two weeks after the last IV dose.

Table 3. Summary of number of patients with treatment-emergent AEs throughout the study*

	Total (N = 25)
≥1 treatment-emergent AE, n (%)	22 (88.0)
AEs by SOC and preferred term,[a] n (%)	
Infections and infestations	18 (72.0)
COVID-19	9 (36.0)
Nasopharyngitis	3 (12.0)
Upper respiratory tract infection	3 (12.0)
Viral infection	2 (8.0)
General disorders and administration site conditions	8 (32.0)
Injection-site pain	4 (16.0)
Injection-site erythema	2 (8.0)
Blood and lymphatic system disorders	7 (28.0)
Leukopenia	4 (16.0)
Neutropenia	4 (16.0)
Lymphopenia	3 (12.0)
Anemia	2 (8.0)
Investigations	5 (20.0)
Neutrophil count decreased	2 (8.0)
Urine protein/creatinine ratio increased	2 (8.0)
White blood cell count decreased	2 (8.0)
Skin and subcutaneous tissue disorders	5 (20.0)
Erythema	2 (8.0)
Musculoskeletal and connective tissue disorders	4 (16.0)
Myalgia	2 (8.0)
≥1 treatment-related AE,[b] n (%)	14 (56.0)
Mild	10 (40.0)
Moderate	4 (16.0)
Severe	0
≥1 SAE, n (%)	1 (4.0)
Infections and infestations	1 (4.0)
COVID-19	1 (4.0)
≥1 AE resulting in treatment discontinuation	1 (4.0)
AESI, n (%)	
All malignancies[c]	0
Post-injection systemic reactions[c,d]	3 (12.0)
All infections of special interest[c,e]	0
Depression (including mood disorders and anxiety)	0
Suicide/self-injury[f]	0
Death	0

* AE, adverse event; AESI, AE of special interest; MedDRA, Medical Dictionary for Regulatory Activities; SAE, serious AE; SOC, system organ class.
[a] For SOC with any preferred terms occurring in more than one patient; preferred terms occurring in more than one patient reported.
[b] Patients counted by maximum AE severity experienced.
[c] Per custom MedDRA query (version 25.1).
[d] All post-injection systemic reactions were mild and did not lead to discontinuation of belimumab, and all patients recovered.
[e] Includes opportunistic infections, herpes zoster, tuberculosis, and sepsis.
[f] Per GSK adjudication.

DISCUSSION

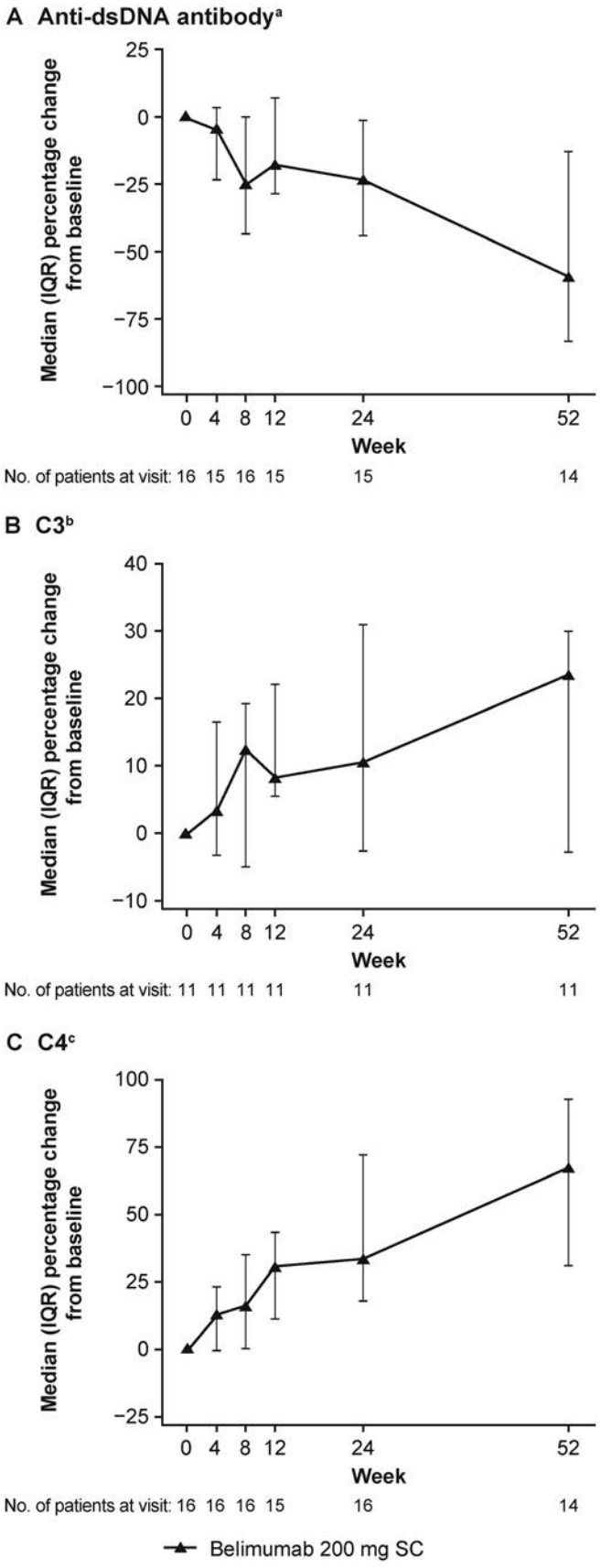
This study of three-weight-band SC belimumab dosing in pediatric patients with SLE receiving standard of care therapy indicates consistent exposure across the weight bands, using extrapolation to the lowest weight band, with no new safety concerns. PD and efficacy responses in this open-label study were consistent with the general trends observed in previous belimumab studies in adults and children.[21,25,33] Our PK and PD results of SC belimumab in this pediatric population support the expansion of current treatment options for this population with often severe SLE manifestations and few steroid-sparing therapeutic options.

The BLISS-SC trial first established the efficacy and safety of belimumab 200 mg SC qw in an adult population with moderate-to-severe SLE.[25] The dosing regimens in the present study were selected to achieve a similar average exposure as observed with the recommended SC dosing regimen in adult patients with SLE. This SC target exposure is also consistent with that obtained with IV belimumab in adults or pediatric patients,[28] for which similar clinical responses have been demonstrated in pediatric and adult patients.[34]

Given the marked changes in weight during childhood, weight-based dosing SC regimens were used. To accommodate children 5 to 17 years of age, and to ensure younger children with a lower body weight were not over-exposed, the three-weight-band dosing regimen was used. Adherence was generally good, with an average of 33.8 injections in the q10d dosing group and 51.2 in the weekly dosing group, including the two patients who withdrew from the study.

Observed belimumab concentrations at week 12 and through 52 weeks were generally somewhat lower in patients weighing 30 to <50 kg, who received belimumab q10d, compared with those weighing 50 kg or heavier, who received belimumab qw. However, the average concentration at steady state ($Cavg_{ss}$) remained comparable in both weight strata in popPK analysis, supporting the rationale for the chosen body-weight-dependent dosing regimen used in the study. Although only trough samples were collected in the present study, frequent sampling over the first 12 weeks to steady state provided sufficient information for a precise characterization of absorption, with systemic PK parameters predominantly informed from the pediatric IV belimumab PLUTO study. The IV PLUTO study included patients with body weights as low as 17 kg, which facilitates extrapolation of data to the lower weight band in this study. The popPK model-derived exposures of pediatric patients receiving SC belimumab resemble those from the BLISS-SC adult trial, with similar average concentrations also seen for pediatric patients receiving IV belimumab in the PLUTO study.[21,25] Fat-free mass was tested for inclusion in the popPK model in place of body weight as a predictor of belimumab PK, which resulted in a small but non-significant improvement in the fit to the pediatric PK data. Ultimately, body weight was selected in preference of fat-free mass because it had practical advantage for a model evaluating the exposure–body weight relationship.

The PD results of the present study demonstrate a profile consistent with that observed in other studies of belimumab.[21,25,33,35] Observed decreases in IgG, total B cell,



and naïve B cell levels by week 52 in this study were consistent with the pediatric IV belimumab PLUTO study,[21] without increased rates of serious infection. Observed decreases in the number of plasmablasts aligned with the known PD profile of belimumab[35]; however, plasmablast changes were not reported for the pediatric IV belimumab PLUTO study for comparison. Fluctuation in circulating memory B-cell levels, as reported in the present study, is expected with belimumab treatment and is believed to be due to memory B-cell redistribution from lymphoid organs and inflamed tissues into circulation.[36]

The safety and tolerability of belimumab in pediatric patients observed throughout this study is consistent with the known safety profile of belimumab, and no new safety concerns were identified.[21,25] A greater proportion of belimumab-treated patients in this study had an AE classified under infections or infestations (72.0%) compared with the pediatric IV belimumab PLUTO study (56.6%) or the adult SC belimumab BLISS-SC study (55.4%).[21,25] Notably, the most common AE in this study was attributable to COVID-19, as was the only SAE. The present study took place during the COVID-19 pandemic, whereas both the BLISS-SC and PLUTO studies were conducted before the pandemic. Additionally, the small sample size of this study should be considered when interpreting these results.

Most patients experienced an improvement in their cSLE disease activity (SELENA-SLEDAI score ≥4-point reduction from baseline) through 52 weeks with belimumab in combination with standard therapy in the present study. Although efficacy was only an exploratory end point in this open-label study, our results are consistent with the efficacy observed in previous SLE studies for the approved belimumab doses in adults (SC) and children (IV).[21,25] Although differences in the proportion of patients in each weight band achieving a ≥4-point reduction in the SELENA-SLEDAI score at weeks 12 and 52 were observed in this study, the small cohort sizes and differences in baseline SELENA-SLEDAI scores likely influenced these results.

A two-weight-band dosing regimen with qw or q2w dosing may be easier to implement in a home setting than a three-weight-band regimen, as the two-weight-band regimen avoids the 10-day dosing interval for patients weighing between 30 and <50 kg and would therefore allow for a consistent day of the week of administration. Based on the results of simulated PK modeling, qw dosing for patients weighing ≥40 kg and q2w dosing for patients weighing 15 to <40 kg yields belimumab exposures that

Figure 2. Median (IQR) percentage change in (A) anti-dsDNA, (B) C3, and (C) C4 levels from baseline to week 52 (ITT population). [a]In patients who were anti-dsDNA antibody positive (≥30 IU/mL) at baseline. [b]In patients with low C3 levels (<90 mg/dL) at baseline. [c]In patients with low C4 levels (<13 mg/dL) at baseline. anti-dsDNA, anti–double-stranded DNA; IQR, interquartile range; ITT, intent-to-treat; SC, subcutaneous.

are consistent across the pediatric weight range and in line with approved adult SC or pediatric IV use. Through PK simulations, the present analysis also suggests that patients can switch from IV to SC administration and maintain comparable exposure levels. The simulation found SC initiation one to two weeks after the last IV dose to be optimal to attain approximately steady-state exposures from the first SC dose, which is consistent with the labeled use of SC belimumab with initiation one to four weeks after the last IV dose.[18]

Given the exploratory nature of some end points, the small number of patients in the study, and the open-label study design with no placebo group, comparisons with the results of previous studies should be interpreted with caution. This is further compounded by differences in methodologies for handling missing data. Notably, dosing recommendations for children weighing less than 15 kg cannot be rendered at this time, and belimumab is not currently approved in this weight range.

In conclusion, this study demonstrates that weight-based stratified dosing of SC belimumab in pediatric patients with cSLE achieves exposure consistent with the established SC adult and IV pediatric doses, with no new safety concerns. These findings allow for the extension of SC belimumab to pediatric patients (aged five years and older) with cSLE who weigh more than 15 kg. This provides clinicians with a potentially glucocorticoid-sparing treatment option for SC use to help manage cSLE and improve patient outcomes.

ACKNOWLEDGMENTS

The authors would like to thank the participating patients and their families, clinicians, and study investigators. Dr Mori would like to acknowledge the support of Kazuko Yamazaki, St Marianna University School of Medicine.

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

GSK was involved in designing the study; contributed to the collection, analysis, and interpretation of the data; supported the authors in the development of the manuscript; and funded the medical writing assistance provided by Claire Barron, MSc, Fishawack Indicia Ltd, UK, part of Avalere Health. All authors, including those employed by GSK, approved the content of the submitted manuscript and were involved in the decision to submit the manuscript for publication.

REFERENCES

1. Groot N, Shaikhani D, Teng YKO, et al. Long-term clinical outcomes in a cohort of adults with childhood-onset systemic lupus erythematosus. Arthritis Rheumatol 2019;71(2):290–301. doi:https://doi.org/10.1002/art.40697
2. Watson L, Leone V, Pilkington C, et al; UK Juvenile-Onset Systemic Lupus Erythematosus Study Group. Disease activity, severity, and damage in the UK juvenile-onset systemic lupus erythematosus cohort. Arthritis Rheum 2012;64(7):2356–2365. doi:https://doi.org/10.1002/art.34410
3. Holland MJ, Beresford MW, Feldman BM, et al; PRINTO and PRCSG Investigators. Measuring disease damage and its severity in childhood-onset systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2018;70(11):1621–1629. doi:https://doi.org/10.1002/acr.23531
4. Hanif M, Sarker C, Beresford MW, et al. OP0246 Factors associated with damage accrual in childhood systemic lupus erythematosus (cSLE): corticosteroid regimens and maintenance of low disease activity [abstract]. Ann Rheum Dis 2024;83(suppl 1):227–228. doi:https://doi.org/10.1136/annrheumdis-2024-eular.1166
5. Hiraki LT, Benseler SM, Tyrrell PN, et al. Clinical and laboratory characteristics and long-term outcome of pediatric systemic lupus erythematosus: a longitudinal study. J Pediatr 2008;152(4):550–556. doi:https://doi.org/10.1016/j.jpeds.2007.09.019
6. Ambrose N, Morgan TA, Galloway J, et al; UK JSLE Study Group. Differences in disease phenotype and severity in SLE across age groups. Lupus 2016;25(14):1542–1550. doi:https://doi.org/10.1177/0961203316644333
7. Hersh AO, von Scheven E, Yazdany J, et al. Differences in long-term disease activity and treatment of adult patients with childhood- and adult-onset systemic lupus erythematosus. Arthritis Rheum 2009;61(1):13–20. doi:https://doi.org/10.1002/art.24091
8. Huang X, Jia N, Xiao F, et al. Differences in the clinical manifestations and mortality of systemic lupus erythematosus onset in children and adults: a systematic review and meta-analysis. Int Arch Allergy Immunol 2022;183(1):116–126. doi:https://doi.org/10.1159/000518321
9. Brunner HI, Vadhariya A, Dickson C, et al. Treatment patterns in paediatric and adult patients with SLE: a retrospective claims database study in the USA. Lupus Sci Med 2023;10(2):e000817. doi:https://doi.org/10.1136/lupus-2022-000817
10. Al Sawah S, Zhang X, Zhu B, et al. Effect of corticosteroid use by dose on the risk of developing organ damage over time in systemic lupus erythematosus-the Hopkins Lupus Cohort. Lupus Sci Med 2015;2(1):e000066. doi:https://doi.org/10.1136/lupus-2014-000066
11. Rygg M, Pistorio A, Ravelli A, et al; Paediatric Rheumatology International Trials Organisation (PRINTO). A longitudinal PRINTO study on growth and puberty in juvenile systemic lupus erythematosus. Ann Rheum Dis 2012;71(4):511–517. doi:https://doi.org/10.1136/annrheumdis-2011-200106
12. Davidson JE, Fu Q, Rao S, et al. Quantifying the burden of steroid-related damage in SLE in the Hopkins Lupus Cohort. Lupus Sci Med 2018;5(1):e000237. doi:https://doi.org/10.1136/lupus-2017-000237
13. Urowitz MB, Gladman DD, Ibañez D, et al. Effect of disease activity on organ damage progression in systemic lupus erythematosus: University of Toronto Lupus Clinic cohort. J Rheumatol 2021;48(1):67–73. doi:https://doi.org/10.3899/jrheum.190259
14. Smith EMD, Aggarwal A, Ainsworth J, et al; cSLE T2T International Task Force. Towards development of treat to target (T2T) in childhood-onset systemic lupus erythematosus: PReS-endorsed overarching principles and points-to-consider from an international task force. Ann Rheum Dis 2023;82(6):788–798. doi:https://doi.org/10.1136/ard-2022-223328

15. Baker KP, Edwards BM, Main SH, et al. Generation and characterization of LymphoStat-B, a human monoclonal antibody that antagonizes the bioactivities of B lymphocyte stimulator. Arthritis Rheum 2003;48(11):3253–3265. doi:https://doi.org/10.1002/art.11299
16. Petri M, Stohl W, Chatham W, et al. Association of plasma B lymphocyte stimulator levels and disease activity in systemic lupus erythematosus. Arthritis Rheum 2008;58(8):2453–2459. doi:https://doi.org/10.1002/art.23678
17. Levy RA, Gonzalez-Rivera T, Khamashta M, et al. 10 Years of belimumab experience: what have we learnt? Lupus 2021;30(11):1705–1721. doi:https://doi.org/10.1177/09612033211028653
18. GSK. Highlights of prescribing information: Benlysta (belimumab) [package insert]. 2024. Accessed October 2025. https://gskpro.com/content/dam/global/hcpportal/en_US/Prescribing_Information/Benlysta/pdf/BENLYSTA-PI-MG-IFU.PDF
19. GSK. GSK announces US FDA approval of Benlysta (belimumab) for pediatric patients with active lupus nephritis. 2022. Accessed October 2024. https://us.gsk.com/en-us/media/press-releases/gsk-announces-us-fda-approval-of-benlysta-belimumab-for-pediatric-patients-with-active-lupus-nephritis/
20. GSK. Annex I summary of product characteristics. European Medicines Agency. 2024. Accessed October 2025. https://www.ema.europa.eu/en/documents/product-information/benlysta-epar-product-information_en.pdf
21. 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. doi:https://doi.org/10.1136/annrheumdis-2020-217101
22. Hashimoto S, Fujita M, Aizawa T, et al. Early add-on use of belimumab as induction therapy for patients with childhood-onset systemic lupus erythematosus. Lupus 2024;33(5):543–544. doi:https://doi.org/10.1177/09612033241238867
23. Roberts JE, Burn C, Sadun RE, et al. Real-world use and outcomes of belimumab in childhood-onset lupus: A single-center retrospective study. Lupus 2023;32(9):1111–1116. doi:https://doi.org/10.1177/09612033231187752
24. Chen F, Zheng Y, Chen X, et al. Belimumab in childhood systemic lupus erythematosus: A review of available data. Front Immunol 2022;13:940416. doi:https://doi.org/10.3389/fimmu.2022.940416
25. Stohl W, Schwarting A, Okada M, et al. Efficacy and safety of subcutaneous belimumab in systemic lupus erythematosus: a fifty-two-week randomized, double-blind, placebo-controlled study. Arthritis Rheumatol 2017;69(5):1016–1027. doi:https://doi.org/10.1002/art.40049
26. GSK. Study of subcutaneous (SC) belimumab in pediatric participants with systemic lupus erythematosus (SLE). 2024. Accessed December 2024. https://cdn.clinicaltrials.gov/large-docs/32/NCT04179032/Prot_000.pdf
27. Centers for Disease Control and Prevention. Clinical growth charts. 2000. Accessed October 2024. https://www.cdc.gov/growthcharts/cdc-charts.htm
28. Dimelow R, Ji B, Struemper H. Pharmacokinetics of belimumab in children with systemic lupus erythematosus. Clin Pharmacol Drug Dev 2021;10(6):622–633. doi:https://doi.org/10.1002/cpdd.889
29. Struemper H, Thapar M, Roth D. Population pharmacokinetic and pharmacodynamic analysis of belimumab administered subcutaneously in healthy volunteers and patients with systemic lupus erythematosus. Clin Pharmacokinet 2018;57(6):717–728. doi:https://doi.org/10.1007/s40262-017-0586-5
30. Wang Y, Jadhav PR, Lala M, et al. Clarification on precision criteria to derive sample size when designing pediatric pharmacokinetic studies. J Clin Pharmacol 2012;52(10):1601–1606. doi:https://doi.org/10.1177/0091270011422812
31. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1997;40(9):1725. doi:https://doi.org/10.1002/art.1780400928
32. Medical Dictionary for Regulatory Activities. Accessed October 2024. https://www.meddra.org/
33. Doria A, Bass D, Schwarting A, et al. A 6-month open-label extension study of the safety and efficacy of subcutaneous belimumab in patients with systemic lupus erythematosus. Lupus 2018;27(9):1489–1498. doi:https://doi.org/10.1177/0961203318777634
34. Brunner HI, Abud-Mendoza C, Mori M, et al. Efficacy and safety of belimumab in paediatric and adult patients with systemic lupus erythematosus: an across-study comparison. RMD Open 2021;7(3):e001747. doi:https://doi.org/10.1136/rmdopen-2021-001747
35. Furie R, Petri M, Zamani O, et al; BLISS-76 Study Group. A phase III, randomized, placebo-controlled study of belimumab, a monoclonal antibody that inhibits B lymphocyte stimulator, in patients with systemic lupus erythematosus. Arthritis Rheum 2011;63(12):3918–3930. doi:https://doi.org/10.1002/art.30613
36. Arends EJ, Zlei M, Tipton CM, et al. Disruption of memory B-cell trafficking by belimumab in patients with systemic lupus erythematosus. Rheumatology (Oxford) 2024;63(9):2387–2398. doi:https://doi.org/10.1093/rheumatology/keae286

Arthritis Care & Research
Vol. 78, No. 6, June 2026, pp 771–780
DOI 10.1002/acr.25641

Preferences for Posttraumatic Osteoarthritis Prevention Strategies in Individuals With Anterior Cruciate Ligament Injury

Kevin Kennedy, PhD,[1] Lily Waddell, BA,[2] Adam Easterbrook, PhD,[3] Jeffrey N. Katz, MD, MSc,[4] Cale Jacobs, PhD,[5] Morgan Jones, MD, MPH,[5] Faith Selzer, PhD,[5] Elena Losina, PhD,[5] Liana Fraenkel, MD, MPH,[6] and Nick Bansback, MSc, PhD[1]

Objective. There is growing interest in evaluating new strategies to delay or prevent posttraumatic osteoarthritis (PTOA) in individuals who have sustained anterior cruciate ligament (ACL) injury. This study sought to determine characteristics of potential treatments that are acceptable to patients with ACL injury.

Methods. Participants with a history of ACL injury were recruited from Reddit, Facebook, and ResearchMatch.org. After consent and eligibility confirmation, participants completed a survey comprising questions on (1) demographics, (2) PTOA perceptions, (3) perceived PTOA risk, and (4) a discrete choice experiment (DCE) task. The DCE assessed treatment attributes including risk reduction, side effects, benefits, and out-of-pocket costs. In several scenarios, participants chose between two hypothetical treatments with various attributes or no treatment. The data were analyzed with multinomial logit, mixed logit, and latent class models.

Results. We enrolled 273 participants (median age 30.6 years [interquartile range 27–33]; 63% female; mean body mass index = 25.5 kg/m^2). Of these, 29% experienced daily knee pain, and 35% reported being very or extremely worried about knee OA. The two most influential attributes affecting treatment acceptability were monthly cost and potential mild side effects. Two preference phenotypes emerged: Class 1 members (n = 162, 59%) generally favored treatment, prioritizing effectiveness and injections but were deterred by high cost. Class 2 members (n = 111, 41%) were less inclined to use treatments with potential mild side effects and high cost.

Conclusion. These results can be used to develop tailored recruitment messaging for future trials. Messaging should emphasize how to manage side effects and out-of-pocket costs.

INTRODUCTION

Osteoarthritis (OA) is a disabling condition that affects >15% of adults in the United States.[1] Individuals who sustain anterior cruciate ligament (ACL) injury are at high risk to develop posttraumatic OA (PTOA), which accounts for nearly 12% of all cases of OA.[2–4] Published data have shown that approximately half of all individuals who undergo ACL reconstruction develop PTOA within 10 to 15 years.[5] Because onset is typically in middle adulthood, PTOA compromises the productivity and quality of life of working individuals. Individuals with PTOA are at a greater risk for total knee arthroplasty (TKA) at an earlier age relative to individuals with idiopathic OA.[6–9]

Because ACL injury often occurs in younger populations, the onset of PTOA can develop at earlier ages.[10] Those with ACL injury can show first structural changes in their 30s as compared with those with idiopathic OA, for whom the median age of diagnosis is 55 years.[5,10] Potential treatments may include rehabilitation, exercise programs, and oral or intra-articular medications.[11] Before designing trials to determine the safety and efficacy of such treatments, it is important to understand the treatment preferences of individuals at high risk for PTOA. When patients decide

Supported by the Arthritis Foundation.
[1]University of British Columbia, Vancouver, British Columbia, Canada; [2]Brigham and Women's Hospital, Boston, Massachusetts; [3]University of British Columbia and Centre for Advancing Health Outcomes, Vancouver, British Columbia, Canada; [4]Brigham and Women's Hospital and Harvard Medical School and T.H. Chan School of Public Health, Boston, Massachusetts; [5]Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts; [6]Yale University, New Haven, Connecticut, and Berkshire Medical Center, Pittsfield, 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/acr.25641).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25641.

Address correspondence via email to Elena Losina, PhD, at elosina@bwh.harvard.edu.

Submitted for publication September 10, 2024; accepted in revised form July 9, 2025.

SIGNIFICANCE & INNOVATIONS

- To our knowledge, this study is the first to apply a discrete choice experiment to understanding patient preferences for treatment strategies to prevent posttraumatic osteoarthritis (PTOA).
- In determining the acceptability of a potential PTOA treatment, we found the two most influential attributes were monthly cost and potential side effects.
- Two preference phenotypes resulted from our analysis. Class 1 members generally favored treatment, prioritizing effectiveness and injections and avoiding high cost, whereas Class 2 members were less inclined to adopt treatments with potential mild side effects and high cost.

whether to start a medication or treatment, they actively weigh the benefits and the risks to make the decision that is best for them. These risk-benefit decisions are collectively referred to as "trade-offs." Assessing patient preferences for treatment options to prevent PTOA can inform the trade-offs patients are willing to make among potential benefits, side effects, costs, and other treatment-related aspects (eg, the need to attend frequent medical visits). A study of older patients with OA without a history of traumatic injury found a significant proportion would accept small risks of moderate side effects to prevent worsening OA, but participants' tolerance of medication risks varied based on their knowledge of the illness and risk perception.[12] A critical aspect of the 21st Century Cures Act requires the US Federal Drug Administration (FDA) to consider patient perspectives and needs in the drug development and regulatory decision-making processes.[13,14] Individuals with significant knee injuries have variable expectations about their future and often have limited opportunities to discuss long-term knee health with their health care providers.[15] Studies like ours can help integrate patient voices into clinical trials, helping fulfill the new FDA criteria for Patient-Focused Drug Development.[13,14]

In this study, we evaluated preferences for potential prevention treatment regimens for individuals at high risk for PTOA. Discrete choice experiments (DCEs) are a reliable and widely applied survey methodology for eliciting individuals' preferences for health care.[16] The DCE method allows us to quantify the trade-offs between attributes and levels associated with preventive treatment, where we can examine in detail how individuals with ACL injury choose between treatment alternatives. DCE tasks afford a context that includes the PTOA risk reduction, costs, and possible side effects associated with hypothetical treatment scenarios. Understanding how patients weigh alternatives is crucial in anticipating the demand for treatments in clinical trials. We envision that the findings will inform future treatment development and help to optimize enrollment into clinical trials evaluating the efficacy of such treatments. The aim of the study is to inform future clinical research into the development of treatments for PTOA in conjunction with the preferences of individuals who have an ACL injury.

PATIENTS AND METHODS

Overview. This study used an internet survey in English to understand preventive treatment preferences for PTOA in individuals with previous ACL injury. The survey had three components, each designed to elicit how participants perceive PTOA risk following ACL injury and would choose among prospective treatment alternatives. First, we assessed participants' knowledge of general PTOA risk following ACL injury. Second, we determined the participants' perceived personal risk for developing PTOA. Third, we elicited preferences for trade-offs of the attributes of potential treatments that might prevent or delay the development of PTOA using a DCE.[16]

We aimed to enroll at least 200 participants to provide sufficient statistical power. The considerations for our power calculation are reported in the "DCE design" section. Participants were recruited from posts on the ACL forums of Reddit, ResearchMatch (a national health volunteer registry operated by Vanderbilt University Medical Center and supported by the National Institutes of Health), and Facebook (ACL injury support and resource groups) as shown in Supplements S1 and S2. Additional participants were recruited from the clinics of orthopedic sports medicine surgeons at an academic medical center. Before voluntarily proceeding to complete the web-based survey, participants reviewed the study consent form with information including the purpose of the study, the investigator, the time length of the survey, and plans for data storage. We collected data between July and September 2023. All questionnaire responses were electronically collected and stored securely on servers at the University of British Columbia. Email addresses were stored in a secure REDCap database at Mass General Brigham for remuneration purposes only. Participants received a $20 Amazon e-gift card for completing the survey. The study was approved by both the Mass General Brigham Institutional Review Board and University of the British Columbia behavioral ethics board. The questionnaire was distributed as an open survey.

Legitimate participants (excluding bots) who resided in North America, were aged between 18 and 45 years, and self-reported an ACL injury diagnosed using advanced imaging techniques (magnetic resonance imaging or computed tomography) in the past 5 years were included in the analyses. Only complete questionnaires were analyzed. We minimized the impact of bots by (1) including CAPTCHA requirements and questions such as "What is your favorite animal?" or "How did you injure your ACL? Tell us in a few sentences" and (2) identifying unlikely responses to weight and height for entries stratified by age and sex (see Supplement S3).

The upper age limit of 45 years was used as data suggest a nontrivial proportion of ACL injuries and ACL reconstructions

occur between the ages of 40 and 45 years.[17] The median age of symptomatic knee OA diagnosis is approximately 55 years, so the survey population was still markedly younger than the average individual with primary knee OA.[18]

Survey design. To guide survey development, we conducted semistructured interviews and focus groups with 25 participants who sustained a recent ACL injury, recruited from orthopedic practices at an academic medical center.[19,20] The interview data were analyzed and used to construct the survey and identify relevant treatment attributes for the DCE. The survey was reviewed with several participants, and, based on this feedback, the wording of some questions and attributes were revised before launching the survey for data collection. To maintain a logical sequence, the survey questions were not randomized or alternated. Adaptive questioning was not used. The survey asked 45 questions across 15 webpages with 1 to 12 questions per page (see Supplement S4). Participants were required to complete all items on a survey page before proceeding and were able to review and change their answers on previous pages.

Knowledge of knee arthritis. We included several true/false statement tasks to approximate participants' knowledge of the body parts affected by knee OA ("Knee arthritis affects all parts of the knee joint including the bone, cartilage, and muscles," correct answer: true), knee OA causes ("Knee arthritis is caused by old age," correct answer: false), knee OA symptoms ("Symptoms of knee arthritis include joint pain, stiffness, swelling, and difficulty with larger joint movements," correct answer: true), and knee OA treatment ("Knee arthritis can only be treated with surgery," correct answer: false). The proportion of correct statements was included as a composite measure of overall knee OA knowledge and included as a covariate in attempting to detect relationships with latent class groups.

Assessment of risk perceptions and preferences. The survey included a series of questions modified from a previous study that captured risk perceptions related to PTOA and its prevention.[21] Participants were asked to estimate their perceived risk of developing knee arthritis in the next 20 years and the chance they will need TKA in their lifetime, both measured on a scale from 0% to 100%. Furthermore, participants were asked if they considered their chance of developing knee OA to be higher or lower than most other individuals of their age and sex. Finally, participants were asked to assign a minimum level of lifetime knee-related PTOA risk for them to initiate the following treatments that may lower this risk: (a) take prescription medication, (b) participate in physical therapy, and (c) modify their lifestyle. Several related knee-specific questions were asked, including how often they experience pain in their ACL-injured knee, how often they are aware of their knee problem, and if they had modified their lifestyle to avoid activities that could potentially damage their knee. A complete knee-specific patient-reported outcomes measure was not included to reduce the time necessary to complete the survey.

Use of OA risk calculator. We collected age, sex at birth, ethnicity, height, weight, family history of arthritis, and exposure to work or movement-related risk factors associated with knee arthritis (eg, persistent kneeling, squatting, or lifting). These participant characteristics were used to estimate their individual risk of developing OA in the next 20 years with a previously validated OA risk calculator.[22]

Defining medication attributes. This work was performed to gauge interest in the use of oral medication for PTOA prevention. The items were informed by data collected from qualitative interviews and from recommendations by experts in a patient preference assessment to evaluate the interest in participating in the Preventing Injured Knees from OA: Severity Outcomes (PIKASO) study.[19,20]

To assess if participants accept oral medications in general, we asked participants how many over the counter or prescription medications they take daily. In addition, we collected participant attitudes on incentives and deterrents to taking medications to reduce the chance of OA on a five-point Likert scale from "strongly disagree" to "strongly agree." These included the importance of reducing the chance of developing PTOA, reducing the chance of requiring TKA, taking a medication daily, tablet size (in relation to ibuprofen), duration of treatment (less than or more than 6 months), and mild temporary side effects, such as diarrhea or nausea.

DCE design. We asked study participants to choose between alternative treatments with varying attributes (eg, patient costs) and levels (eg, $0 per month, $10 per month, or $100 per month) with eight discrete choice questions. We framed the DCE questions as a choice involving two hypothetical treatment options for preventing PTOA and an "opt out" option ("I prefer to choose neither option"). Alternatives varied attribute levels for each choice set. Attributes included (1) a description of treatment effectiveness at reducing the chance of developing PTOA in the next 20 years by one of three possible relative risk-reduction levels (50%, 25%, or 0%, all displayed as absolute risks based on individualized calculated baseline risk); (2) type of treatment (physical therapy alone or with a daily pill or one-time injection); (3) side effects (none, temporary mild nausea/diarrhea, or rare but very serious condition); (4) other potential benefits (none or delayed risk of diabetes/cancer); and (5) out-of-pocket cost ($0, $10 per month, or $100 per month). The complete set of attributes and levels included in the DCE are described in Supplement S5.

A d-efficient design was generated to maximize the statistical difference between the profiles in each choice set by iteratively comparing sets of alternatives with a modified Fedorov algorithm, using the idefix package in R.[23–25] The final design included 48 unique choice sets, each describing a pair of alternatives and an opt out option. The clear dominance condition is met when there is one clearly preferred treatment profile defined by greater benefits, fewer adverse events, and lower costs. Two additional choice tasks including the same set of attribute levels with clear dominance were inserted into the design as both the first question (warmup task) and the fifth question (consistency check). The design was blocked into eight unlabeled choice sets. Participants were assigned a random value between one and eight to allocate their choice tasks to one of the eight blocks and were asked to choose their preferred alternative for each question.

To determine the number of effective participants required to adequately measure preferences, we decided that an SE of 0.05 was sufficient to detect meaningful differences between attribute levels using a mixed logit (MXL) model. Based on a preliminary sample of 40 participants, we calculated a minimum sample size of 200 participants to estimate significant coefficients for level 2 parameters (from zero) from the observed pattern of variation across parameters included in an MXL model. An example choice question is shown in Figure 1. The study was conducted in accordance with good research practice for DCEs and follows the International Society for Pharmacoeconomics and Outcomes Research Good Practice Guidelines.[26]

Data analysis. We conducted the analysis in a series of steps to quantify preferences for DCE attribute levels. First, we estimated weights for the DCE attribute level using a dummy-coded multinomial logit (MNL) model to determine the overall sample preferences.[27] Second, we added a random parameter to the MNL model to estimate an MXL.[28] This model examined whether sufficient preference heterogeneity existed to justify further analysis of subgroup preferences. The random parameter included an error component that described variability for all the alternatives (independent and identically distributed [IID]) across all participants (not all observations).[29] In the event of statistical significance of the IID parameter, a third, scale-adjusted latent class (SALC) model was estimated. The SALC model allows for the accounting of correlated preferences and improves model fit, leading to more reliable inferences about preference heterogeneity. Latent class models with covariates can be used to examine preference heterogeneity and identify subgroups within the sample that differ in how they choose between treatment alternatives.[30] The central concept is that members of each distinct group are associated with an unobserved, or latent, construct that assigns different weights to observed attribute levels (or opting out).[29,30,31] This type of model affords insight into the characteristics associated with groups that have different segments by including covariates (eg, age or attitudes) associated with this set of preferences.

The number of latent classes selected was based on model parsimony and model fit (log likelihood tests and Akaike information criteria [AIC]). The statistical significance of an attribute-level coefficient ($P < 0.05$) indicated that participants assigned to that class considered the specific attribute level to affect their choice of treatment strategy. The coefficient sign reflects whether the attribute has a positive or negative effect on the perceived value of, and therefore preferences for, a treatment characteristic. To characterize class membership, we examined whether sociodemographic characteristics were associated with a specific preference-set class membership by entering them into the model as covariates one at a time. We retained all sociodemographic characteristics significantly associated with class membership at the $P < 0.05$ level that also resulted in equal or improved model fit using the AIC. We made predictions for treatment uptake by combining coefficients for specified levels and 95% confidence intervals (CIs) calculated using the SE of estimates. We estimated probabilities for all combinations of attributes relative to the null model to illustrate the differences between class preferences. Probability estimates reflect the likelihood an individual belonging to Class 1 or 2 would select each treatment profile. We conducted the statistical analyses and reporting with base R and the apollo package for the MNL and MXL models and LatentGold6 for the SALC model.[32,33]

RESULTS

Study sample. After excluding generated responses (bots), we deemed 273 responses to be legitimate, and their data comprised the analytic dataset (Figure 2). Of these 273 participants, 102 (37%) were aged 18 to 30 years and 173 (63%) were female. The average body mass index (BMI) for the sample was 25.55 (95% CI 24.91–26.19), and 86 (31.5%) participants reported having at least one immediate family member with a history of arthritis. Most participants (n = 197, 72%) underwent ACL reconstruction to treat their injury. Most participants (n = 224, 82%) reported infrequent (monthly) or no regular use of supplements. In terms of current knee pain, 179 participants (66%) reported at least weekly knee pain, 73 (27%) reported monthly knee pain, and 21 (8%) reported never having knee pain. Many participants were aware of their knee problem daily (n = 105, 39%), even if they were not currently experiencing pain (Table 1).

Knowledge and perceived risk of PTOA. Most participants (n = 245, 89%) responded correctly to all four questions about knee arthritis, demonstrating a good baseline knowledge. Participants considered their chance of developing knee OA to be higher (n = 156, 57%) or much higher (n = 32, 12%) than most

	Treatment A	Treatment B	I prefer to choose neither option
Effectiveness in reducing chance of developing arthritis and related pain in the next 20 years (and potential for future joint replacement)	18 out of 100 people will develop knee arthritis	15 out of 100 people will develop knee arthritis	22 out of 100 people will develop knee arthritis
Type of treatment	A pill every day for the next 12 months and physical therapy 2 times a week for 6 months	Physical therapy 2 times a week for 6 months	Physical therapy 2 times a week for 6 months
Side effects	Very rare chance (~6 in 100,000 people) of developing a very serious condition	Mild nausea/diarrhea that typically goes away after 2 weeks, but possibly persistent increase in bowel movements	No nausea/diarrhea
Other potential benefits	Potentially delay the onset of diabetes	Potentially lower the risk of some cancers in individuals with type 2 diabetes	None
Out-of-pocket cost	$10 per month	$100 per month	$0
Which do you prefer?	**O**	**O**	**O**

Figure 1. An example of a discrete choice experiment choice task.

others of their age or sex because of their knee injury. The median risk of developing OA after 20 years (measured from the OA risk calculator) was 10% (interquartile range [IQR] 8%–14%).[22] However, the median perceived risk was much higher at 65% (IQR 45%–80%). Most participants expressed some worry about developing knee OA in the future, either moderately worried (n = 86, 32%), very worried (n = 73, 27%), or extremely worried (n = 22, 8%).

Medication preferences. Most study participants were somewhat (n = 172, 63%) or very (n = 82, 30%) interested in medication use to reduce the risk of knee arthritis. For a hypothetical medication that would reduce the need to undergo knee replacement, 218 participants (80%) were at least somewhat interested in taking the medication. This was relative to the 195 participants (73%) that were at least somewhat interested in taking a medication that would reduce chances of developing arthritis (without mentioning reducing the need for TKA). Furthermore, 173 participants (63%) agreed that the presence of any (even mild) side effects was a barrier to taking an oral medication. The other barriers to taking an oral medication included the need to take a medication daily (n = 97, 36%), larger than normal pill size (n = 104, 38%), and requiring taking the medication for a long time (n = 110, 40%).

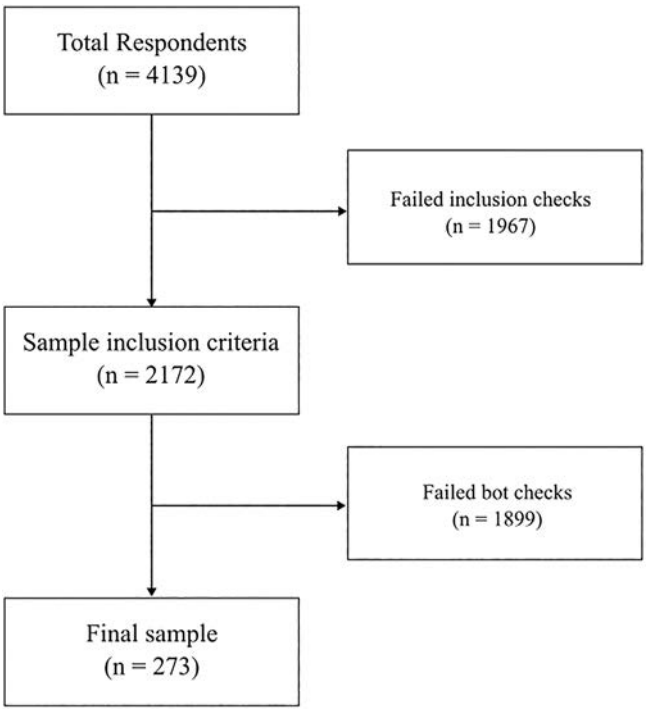


Figure 2. Sample inclusion flow diagram.

Preferences for treatment alternatives. Overall, 195 of 273 (71%) participants provided the same response for the warmup task and the consistency check question. In the initial MNL model, all levels followed the expected direction (eg, 50% effectiveness was preferred over 25%), and participants considered out-of-pocket cost ($100 per month) and mild nausea/diarrhea to be the most important attribute levels to consider when accepting a treatment for preventing PTOA. Notably, a 50% reduction in the chance of developing PTOA was considered less important than a chance of side effects, and a single injection was preferred over taking a daily pill (Table 2).

A two-class latent class model best fit the data, suggesting at least two preference phenotypes exist for this choice context and sample. Figure 3 illustrates differences between individual attribute levels. Overall, participants in Class 1, representing 59% of the sample (n = 162), were more likely to choose treatment over participants in Class 2. Class 1 participants prioritized the following: effectiveness of the medication, mode of administration (preferred injections over pills), avoidance of mild nausea/diarrhea, and concerns regarding $100 per month out-of-pocket costs. Participants in Class 2 (n = 111, 41%) were less likely to choose treatment overall, with strong preferences to avoid any side effects (mild nausea/diarrhea or small chance of serious side effects) and cost as low as $10 per month. These participants placed a greater value on treatments that could potentially lower cancer risk.

Characteristics of the participants were associated with class membership (Table 2), where participants who reported taking five or more daily medications were more likely to belong to Class 1 as were participants who were moderately to extremely worried about arthritis. A higher perceived lifetime risk for the development of OA was identified as a covariate for membership in Class 2, the group of participants less likely to accept treatment.

Table 1. Participant characteristics*

Characteristics	Total participants (N = 273)
Age group, n (%)	
18–29 y	102 (37.36)
30–45 y	171 (62.64)
Age, median (Q1–Q3), y	30.6 (27–33)
Location, n (%)	
Canada	5 (1.83)
United States	268 (98.17)
Sex at birth, n (%)	
Female	173 (63.37)
Male	100 (36.63)
Ethnicity, n (%)	
Asian	26 (9.52)
Black	35 (12.82)
Hispanic or Latino/a	24 (8.79)
Native American or other Pacific Islander	2 (0.73)
White	177 (64.84)
Other	9 (3.3)
Body composition, mean (95% CI)	
Height, cm	170.43 (169.34–171.52)
Weight, lbs	163.60 (159.42–167.78)
BMI, kg/m^2	25.55 (24.91–26.19)
Risk of developing OA, median (Q1–Q3)	10 (8–14)
Any parents or siblings with arthritis, n (%)	86 (31.5)
Underwent ACL reconstruction, n (%)	
Yes	197 (72.16)
No	76 (27.84)
Number of medications (not for knee), n (%)	
None	115 (42.12)
1–2	127 (46.52)
3–4	28 (10.26)
5+	3 (1.1)
Regular use of supplements, n (%)	
None	97 (35.5)
Daily	127 (46.5)
Weekly	33 (12.1)
Monthly	16 (5.9)
Pain in knee after ACL injury, n (%)	
Never	21 (7.7)
Monthly	73 (26.7)
Weekly	92 (33.7)
Daily	79 (28.9)
Always	8 (2.9)
Awareness of knee problem, n (%)	
Never	10 (3.7)
Monthly	57 (20.9)
Weekly	73 (26.7)
Daily	105 (38.5)
Constantly	28 (10.3)

* ACL, anterior cruciate ligament; BMI, body mass index; CI, confidence interval; OA, osteoarthritis; Q, quartile.

Predictions. From the results of the DCE, we calculated how a similar composition of patients with similar preferences

Table 2. Discrete choice experiment parameter estimates, overall and latent class with covariates*

Variable	Options	Models of preferences for OA prevention medication (dummy coded)				
		MNL coefficient (SE)	MXL coefficient (SE)	Class 1		Class 2 *P* value
Alternative-specific constant (choose no treatment)		−0.53 (0.13)[a]	−0.64 (0.15)[b]	−1.74	0.24	<0.001[a]
Effectiveness in reducing chance of developing arthritis in the next 20 years	0%	–	–	–	–	0.35
	25%	0.19 (0.09)[c]	0.25 (0.09)[c]	0.11	0.47	
	50%	0.35 (0.10)[a]	0.43 (0.09)[a]	0.32	0.64	
Type of treatment	None	–	–	–	–	<0.001[a]
	Pill	−0.26 (0.09)[b]	−0.25 (0.10)[b]	−0.08	−0.21	
	Injection	−0.01 (0.10)	0.03 (0.09)	0.17	−0.66	
Side effects	None	–	–	–	–	<0.001[a]
	Mild nausea/diarrhea	−0.72 (0.1)[a]	−0.84 (0.09)[a]	−0.52	−1.40	
	Small chance of serious side effect	−0.51 (0.09)[a]	−0.61 (0.09)[a]	−0.28	−1.38	
Other potential benefits	None	–	–	–	–	0.005[b]
	Potentially delay the onset of diabetes	0.02 (0.09)	0.11 (0.10)	−0.03	0.37	
	Potentially lower the risk of some cancers in T2D	0.14 (0.08)[c]	0.27 (0.11)[b]	−0.01	0.80	
Out-of-pocket cost	$0	–	–	–	–	<0.001[a]
	$10 per month	−0.04 (0.08)	−0.02 (0.09)	0.10	−0.50	
	$100 per month	−0.79 (0.12)[a]	−0.96 (0.08)[a]	−0.54	−1.91	
Random effect	IID	–	−0.99 (0.07)[a]	–	–	
Covariates (effects coded)						
Number of medications not for the knee	None	–	–	−0.76	0.76	0.072
	1–2	–	–	−0.39	0.39	
	3–4	–	–	−0.17	0.17	
	≥5	–	–	1.32	−1.32	
Worried about knee arthritis	Not at all	–	–	−0.49	0.49	<0.001[a]
	Somewhat	–	–	−0.50	0.50	
	Moderately	–	–	0.12	−0.12	
	Very much	–	–	0.40	−0.40	
	Extremely	–	–	0.46	−0.46	
Perceived lifetime risk	Low	–	–	0.22	−0.22	0.055
	Moderate	–	–	0.10	−0.10	
	High	–	–	−0.32	0.32	
	Log likelihood	−1,692	−1,580	−1,482		
	Adjusted R^2	0.054	0.115	0.19	0.34	

* The *P* values for the MNL and MXL model coefficients reflect the significance of attribute levels for the whole sample within each model. The *P* values for Class 1 and Class 2 indicate a significant difference among any attribute levels between the two classes. IID, independent and identically distributed; MNL, multinomial logit; MXL, mixed logit; OA, osteoarthritis; T2D, type 2 diabetes.
[a] $P < 0.001$.
[b] $P < 0.01$.
[c] $P < 0.05$.

would choose between treatment alternatives for a relevant base scenario. When presented with the base scenario of a pill treatment with 25% effectiveness that causes temporary mild diarrhea/nausea, no additional benefits, and costs $10/month, we would anticipate that 40% of participants assigned to Class 1 and 15% assigned to Class 2 would accept treatment. If, from this base scenario, the efficacy of the treatment was improved to 50% effectiveness, 5% more (40% to 45%) participants in Class 1 and 4% more (15% to 19%) assigned to Class 2 were estimated to accept treatment. If the out-of-pocket cost increased from $10 per month to $100 per month, 14% fewer participants assigned to Class 1 (40% to 26%) and 11% fewer (15% to 4%) assigned to Class 2 would accept treatment. If the method of medication administration was changed from a pill to an injection, 7% more (40% to 47%) participants assigned to Class 1 and 5% fewer (15% to 10%) assigned to Class 2 would accept treatment. Figure 4 illustrates a sample of 100 patients with ACL injury and the expected acceptance proportion across these scenarios. The complete set of scenarios with the most/least desired sets of attribute levels are illustrated in Supplement S6.

DISCUSSION

We surveyed individuals who experienced an ACL injury to examine the treatment preferences for lowering the risk of PTOA. Our findings suggest that there is considerable heterogeneity in

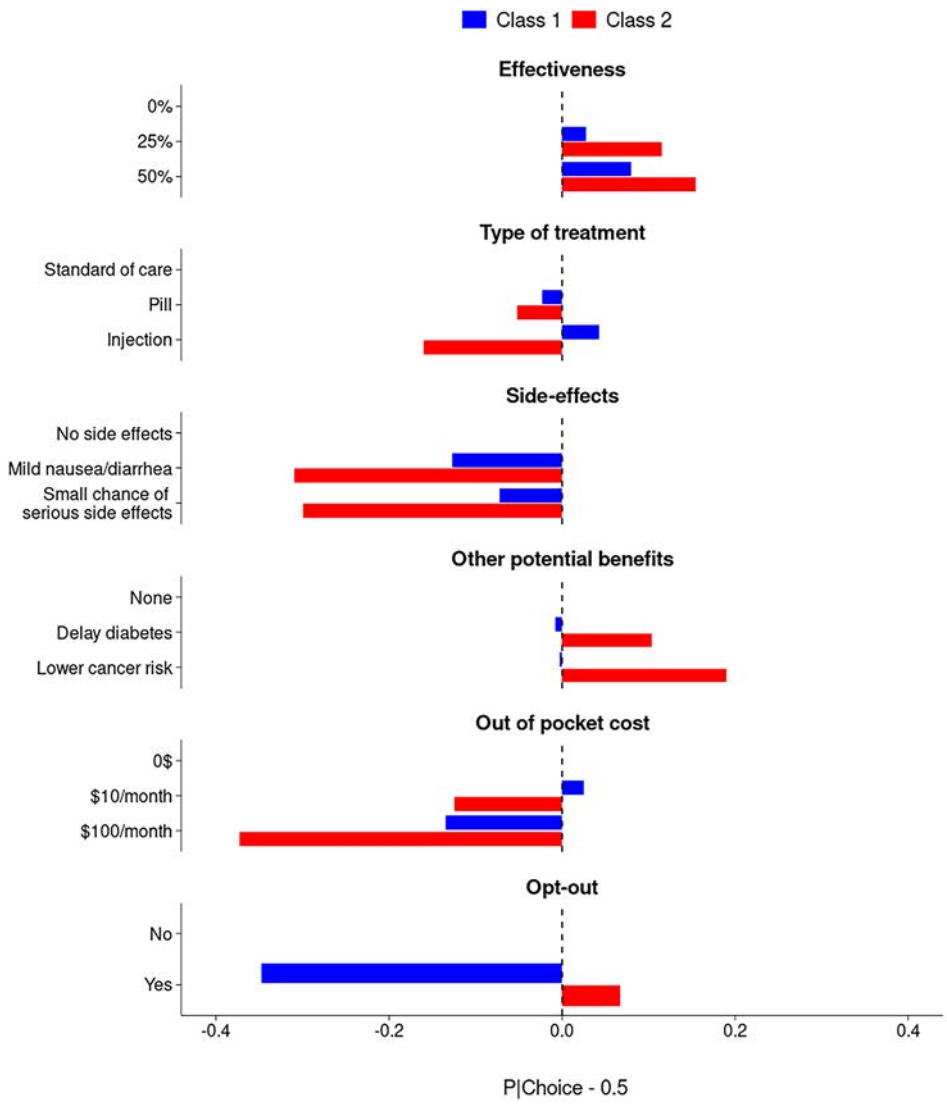


Figure 3. Discrete choice experiment attribute-level probability by latent class. The attribute level sizes indicate the relative importance and direction of preferences on a latent scale. Within class, the magnitude of values can be compared on a relative scale.

preferences, but we identified that there is a subgroup of individuals interested in treatments that may delay or lower their risk for PTOA. Their preferences are shaped by the perceived effectiveness of treatment, potential for side effects, mode of medication administration, and out-of-pocket costs. These factors can be important barriers to individuals using treatments. Methods used in this study could provide critical data to inform investigators designing OA prevention trials about key components of the study that would optimize recruitment by addressing factors important for individuals at risk of OA.

Our findings can be used by researchers designing trials for PTOA. For example, there is a clear preference for one-time injectable treatment for those opposed to daily pills. Regardless of the medication delivery method, the treatment needs to be inexpensive, as a $100 per month out-of-pocket cost is a clear and strong deterrent. Finally, although side effects are an inevitable feature of any medication, special attention should be given to how temporary mild side effects, such as nausea and diarrhea, are communicated. We found that these side effects are significant deterrents to participants choosing treatment alternatives within our scenarios. A review highlighting the use of verbal descriptors, such as "common" or "rare," to convey the frequency of side effects can lead to unrealistic patient expectations.[34] It suggested presenting numerical data on the side effects to set realistic expectations.[34]

Results of this study should be viewed in light of several limitations, which may affect the generalizability of the results to patients who reflect the characteristics of the participants included in this study. First, the recruitment of participants meeting the inclusion criteria was difficult to assess because the survey depended on the self-report of injury diagnosis and the type of imaging technology used to make that diagnosis. Ideally, these

Figure 4. Discrete choice experiment scenario prediction by latent class. The figure describes the predicted number of participants out of 100 who would choose treatment given the described characteristics (out-of-pocket cost, pill/injection, effectiveness, side effects, or other benefits). The columns describe the differences by different latent classes, and the rows show the changes in predicted uptake given changes in a single attribute level (all other attributes held the same).

participants would be identified directly from electronic medical records, which was not feasible to support the analysis of DCE trade-offs with a sufficient sample size. The sampling methodology collected responses from multiple sources, including in-clinic recruitment. Second, in the study advertisement, a monetary incentive was listed in the post on public forums, leading to excessive illegitimate responses from automated machines. Although we attended to the quality of the responses, there may have been participants who were excluded from the analysis that provided an account of their experience and preferences.

Third, the design of the attributes and levels in the discrete choice tasks presented challenging trade-offs. The risk reduction associated with treatment effectiveness was dependent on the estimated risk for each participant. This approach to risk presentation allows for more realistic trade-offs in evaluating treatment effectiveness. However, it may have resulted in internally inconsistent responses within choice sets where participants were presented with relatively small changes in their absolute risk, particularly for the MNL model. Specifically, some participants saw proportionally similar changes but with larger differences in magnitude, which may have led to varying interpretations. Furthermore, the language used to describe side effects and treatment options was simplified to facilitate trade-offs between choice alternatives. Although this was intended to make the tasks more manageable, it may have been overly simplistic, potentially leading participants to select dominant attribute levels. Fourth, because the survey relied on self-report data, a definitive, current diagnosis of OA was not verified in this study. Regarding the potential impact on our overall modeling approach, we considered this limitation to have a minor effect. We observed consistent incremental preferences across levels of proportionately incremental risk. These issues were addressed using the MXL model, which accounts for individual-level variation within participants (IID).

Our study was not sufficiently powered to measure all possible relationships of the covariates and their relation to latent classes. This is a challenging task because there is inherent uncertainty in the number of latent classes, the preference functions of those groups, and the associated characteristics (covariates) of membership. Future studies aimed at understanding patient treatment preferences should consider evaluating scenarios of latent class fit for the number and potential effects for associated covariates in a priori power calculations.

In conclusion, this study demonstrates how patient preferences inform the acceptability of treatment alternatives. This can guide treatment development, inform clinical trial design, and/or understand patient population preferences for available treatment alternatives. Further studies should consider other factors

involved in participating in clinical trials beyond the treatment itself, such as incentives, randomization, and modes of recruitment.

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 Kennedy 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. Theis KA, Steinweg A, Helmick CG, et al. Which one? What kind? How many? Types, causes, and prevalence of disability among U.S. adults. Disabil Health J 2019;12(3):411–421.
2. Osteoarthritis Research Society International. Osteoarthritis: a serious disease. Published December 1, 2016. Accessed January 6 2026. https://oarsi.org/sites/oarsi/files/docs/2016/oarsi_white_paper_oa_serious_disease_121416_1.pdf
3. Thomas AC, Hubbard-Turner T, Wikstrom EA, et al. Epidemiology of posttraumatic osteoarthritis. J Athl Train 2017;52(6):491–496.
4. Brown TD, Johnston RC, Saltzman CL, et al. Posttraumatic osteoarthritis: a first estimate of incidence, prevalence, and burden of disease. J Orthop Trauma 2006;20(10):739–744.
5. Luc B, Gribble PA, Pietrosimone BG. Osteoarthritis prevalence following anterior cruciate ligament reconstruction: a systematic review and numbers-needed-to-treat analysis. J Athl Train 2014;49(6):806–819.
6. Buechel FF Sr. Knee arthroplasty in post-traumatic arthritis. J Arthroplasty 2002;17(4) (suppl 1):63–68.
7. Lonner JH, Pedlow FX, Siliski JM. Total knee arthroplasty for posttraumatic arthrosis. J Arthroplasty 1999;14(8):969–975.
8. Wu LD, Xiong Y, Yan SG, et al. Total knee replacement for posttraumatic degenerative arthritis of the knee. Chin J Traumatol 2005;8(4):195–199.
9. Weiss NG, Parvizi J, Hanssen AD, et al. Total knee arthroplasty in post-traumatic arthrosis of the knee. J Arthroplasty 2003;18(3) (suppl 1):23–26.
10. Roos H, Adalberth T, Dahlberg L, et al. Osteoarthritis of the knee after injury to the anterior cruciate ligament or meniscus: the influence of time and age. Osteoarthritis Cartilage 1995;3(4):261–267.
11. Watt FE, Corp N, Kingsbury SR, et al. Towards prevention of posttraumatic osteoarthritis: report from an international expert working group on considerations for the design and conduct of interventional studies following acute knee injury. Osteoarthr. Cartil. 2019;27(1):23–33
12. Fraenkel L, Suter L, Cunningham CE, et al. Understanding preferences for disease-modifying drugs in osteoarthritis. Arthritis Care Res (Hoboken) 2014;66(8):1186–1192.
13. 21st Century Cures Act, Pub L No. 114-255, 130 Stat 1033 (2016).
14. US Food and Drug Administration. CDER patient-focused drug development. Accessed January 6 2026. www.fda.gov/drugs/development-approval-process-drugs/cder-patient-focused-drug-development
15. Holm PM, Simonÿ C, Brydegaard NK, et al. An early care void: the injury experience and perceptions of treatment among knee-injured individuals and healthcare professionals - a qualitative interview study. Phys Ther Sport 2023;64:32–40.
16. Soekhai V, Whichello C, Levitan B, et al. Methods for exploring and eliciting patient preferences in the medical product lifecycle: a literature review. Drug Discov Today 2019;24(7):1324–1331.
17. Collins JE, Katz JN, Donnell-Fink LA, et al. Cumulative incidence of ACL reconstruction after ACL injury in adults: role of age, sex, and race. Am J Sports Med 2013;41(3):544–549.
18. Losina E, Weinstein AM, Reichmann WM, et al. Lifetime risk and age at diagnosis of symptomatic knee osteoarthritis in the US. Arthritis Care Res (Hoboken) 2013;65(5):703–711.
19. Waddell LM, Mitchener DP, Frier KC, et al. Patient perceptions of medication therapy for prevention of posttraumatic osteoarthritis following anterior cruciate ligament injury: a qualitative content analysis. Arthritis Care Res (Hoboken) 2025;77(7):881–891.
20. Jacobs CA, Jones MH, Collins JE, et al; for PIKASO Team. The PIKASO trial (Preventing Injured Knees from Osteoarthritis: Severity Outcomes): rationale and design features for a randomized controlled trial. Osteoarthr Cartil Open 2024;7(1):100563.
21. Michl GL, Katz JN, Losina E. Risk and risk perception of knee osteoarthritis in the US: a population-based study. Osteoarthr. Cartil. 2016;24(4):593–596.
22. Losina E, Klara K, Michl GL, et al. Development and feasibility of a personalized, interactive risk calculator for knee osteoarthritis. BMC Musculoskelet Disord 2015;16(1):312.
23. Vanniyasingam T, Cunningham CE, Foster G, et al. Simulation study to determine the impact of different design features on design efficiency in discrete choice experiments. BMJ Open 2016;6(7):e011985.
24. Traets F, Sanchez DG, Vandebroek M. Generating optimal designs for discrete choice experiments in R: the idefix package. J Stat Softw 2020;96(3):1–41.
25. R Core Team. R: a language and environment for statistical computing. MSOR Conn 2014;1.
26. Bridges JF, Hauber AB, Marshall D, et al. Conjoint analysis applications in health--a checklist: a report of the ISPOR Good Research Practices for Conjoint Analysis Task Force. Value Health 2011;14(4):403–413.
27. Greene WH, Hensher DA. A latent class model for discrete choice analysis: contrasts with mixed logit. Transp Res, Part B: Methodol 2003;37(8):681–698.
28. Hensher DA, Greene WH. The mixed logit model: the state of practice. Transportation 2003;30(2):133–176.
29. Hess S, Rose JM, Hensher DA. Asymmetric preference formation in willingness to pay estimates in discrete choice models. Transp Res, Part E Logist Trans Rev 2008;44(5):847–863.
30. Dayton CM, Macready GB. Use of categorical and continuous covariates in latent class analysis. In: Hagenaars JA, McCutcheon AL, eds. Applied Latent Class Analysis. Cambridge University Press; 2002:213–233.
31. Swait J. A structural equation model of latent segmentation and product choice for cross-sectional revealed preference choice data. J. Retail. Consum. Serv. 1994;1(2):77–89.
32. Hess S, Palma D. Apollo: A flexible, powerful and customisable freeware package for choice model estimation and application. J. Choice Model. 2019;32:100170. https://doi.org/10.1016/j.jocm.2019.100170.
33. Vermunt JK, Magidson J. LG-syntax user's guide: manual for latent gold syntax module version 6.0. Arlington, MA: Statistical Innovations 2021.
34. Smith LE, Webster RK, Rubin GJ. A systematic review of factors associated with side-effect expectations from medical interventions. Health Expect 2020;23(4):731–758.

Arthritis Care & Research
Vol. 78, No. 6, June 2026, pp 781–787
DOI 10.1002/acr.25695

Response to Allopurinol and Febuxostat According to the Fractional Excretion of Urate in Men With Gout

Pascal Richette,[1,2] Anouk Walter-Petrich,[3,4] Quang Dinh Nguyen,[5] Matthieu Resche-Rigon,[3,4] Quang Huy Dinh Nguyen,[5] Minh Duc Do,[6] Thao Phuong Mai,[6] Tuan Duc Nguyen,[6] Fernando Perez-Ruiz,[7] Emmanuel Letavernier,[8,9] Tristan Pascart,[10] Augustin Latourte,[1,2] Hang Korng Ea,[1,2] and Thomas Bardin[1,2]

Objective. Body mass index (BMI), glomerular filtration rate (GFR), and pretreatment urate levels have been reported to influence the urate-lowering response to allopurinol. We investigated whether the fractional excretion of uric acid (FEUA) also modulates this response and relates to oxypurinol concentrations. We further evaluated its potential influence on febuxostat, not as a direct comparison, but to determine whether the effect of FEUA was specific to allopurinol.

Methods. The data are from n = 1,547 and n = 296 patients starting allopurinol and febuxostat, respectively. The relationship between FEUA (≤5.5% or >5.5%) and the dose response to allopurinol or febuxostat was assessed by linear mixed-effects regression models on serum urate levels and adjusted for BMI, estimated GFR (eGFR), and treatment doses. Concentrations of oxypurinol were measured in a subgroup of patients (n = 181). A multiple linear regression model was used to assess the association between FEUA and oxypurinol concentrations, adjusted for BMI, eGFR, allopurinol dosage, and serum urate levels.

Results. The median FEUA in the whole population was 4.0% (quartile 1–3: 3%–5.1%). The changes in serum urate levels for each 150-mg increase in allopurinol in patients with FEUA ≤5.5% or >5.5% were −72.37 (confidence interval [CI] −74.81 to −69.94) µM and −65.96 (CI −71.29 to −60.62) µM, respectively (P = 0.032). We found higher oxypurinol concentrations in patients with the lowest FEUA (P = 0.032). However, we did not observe any interaction between the febuxostat response and FEUA (P = 0.13).

Conclusion. Allopurinol is more effective in patients with low FEUA, probably because of the reduced renal excretion of oxypurinol. These data highlight the similarity between the renal handling of oxypurinol and urate.

INTRODUCTION

Allopurinol, an analog of hypoxanthine, is the recommended first-line medication for most people with gout.[1] After ingestion, approximately 80% to 90% of allopurinol is rapidly metabolized to oxypurinol, which is almost completely eliminated by the kidneys. Both allopurinol and oxypurinol inhibit xanthine oxidoreductase, the enzyme responsible for converting hypoxanthine to xanthine and xanthine to urate, primarily in the liver.[2]

There is considerable interindividual variability in response to allopurinol. Several clinical factors have been reported to be associated with a better response to allopurinol, including low body

Supported by the French Society of Rheumatology, ArtViggo, and Bridge Point (Francois Davy).

[1]Pascal Richette, MD, PhD, Augustin Latourte, MD, PhD, Hang Korng Ea, MD, PhD, Thomas Bardin, MD: Hôpital Lariboisière, Université Paris Cité, Assistance Publique Hôpitaux de Paris, Paris, France; [2]Pascal Richette, MD, PhD, Augustin Latourte, MD, PhD, Hang Korng Ea, MD, PhD, Thomas Bardin, MD: Hôpital Lariboisière, Université Paris Cité, INSERM U1132, Paris, France; [3]Anouk Walter-Petrich, PhD, Matthieu Resche-Rigon, MD, PhD: Hôpital Saint-Louis, Université Paris Cité, Assistance Publique Hôpitaux de Paris, Paris, France; [4]Anouk Walter-Petrich, PhD, Matthieu Resche-Rigon, MD, PhD: ECSTRRA team, UMR U1342, INSERM, Université Paris Cité, Paris, France; [5]Quang Dinh Nguyen, BSc, Quang Huy Dinh Nguyen, BSc: Vien Gut Medical Clinic, Ho Chi Minh City, Vietnam; [6]Minh Duc Do, MD, PhD, Thao Phuong Mai, MD, Tuan Duc Nguyen, PhD: University of Medicine and Pharmacy at Ho Chi Minh City, Ho Chi Minh City, Vietnam; [7]Fernando Perez-Ruiz, MD, PhD: Medicine and Nursery School, University of the Basque Country, Basque Country, Spain; [8]Emmanuel Letavernier, MD, PhD: Hôpital Tenon, Sorbonne Université, Assistance Publique Hôpitaux de Paris, Paris, France; [9]Emmanuel Letavernier, MD, PhD: INSERM, UMR S 1155, Paris, France; [10]Tristan Pascart, MD, PhD: Hôpital Saint-Philibert, Université Lille Catholique, Lomme, France.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25695).

[Correction added on 27 February 2026, after first online publication: The degree and affiliation of Dr. Tuan Duc Nguyen have been changed in this version.]

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25695.

Address correspondence via email to Pascal Richette, MD, PhD, at pascal.richette@aphp.fr.

Submitted for publication April 25, 2025; accepted in revised form October 16, 2025.

SIGNIFICANCE & INNOVATIONS

- Body mass index and renal function are two factors known to influence the response to allopurinol.
- We show that the fractional excretion of urate acid is also a variable that modulates the response to allopurinol.

mass index (BMI), poor renal function as measured by the estimated glomerular filtration rate (eGFR), and high pretreatment urate levels.[3] Heritability might also influence the response to allopurinol: the presence of the lysine (K) allele of ABCG2, a urate transporter expressed in the kidney but also in the intestine and the liver, has been reported to be associated with a reduced effect of allopurinol in some studies,[4,5] but not all.[6,7]

At the kidney level, two main factors influence the urate levels. The eGFR determines the filtered urate load, whereas the fractional excretion of uric acid (FEUA) reflects tubular secretion and reabsorption. Although largely independent when renal function is preserved, FEUA may increase as eGFR declines in advanced chronic kidney disease as a compensatory mechanism for urate excretion by residual nephrons or the involvement of intestinal urate excretion.[8–11]

Because renal underexcretion of urate, as defined by an FEUA ≤5.5%,[8,12] is the dominant cause of hyperuricemia in 80% to 90% of patients with gout,[1] studies have investigated the influence of FEUA on the response to uricosuric agents, particularly benzbromarone. Overall, these studies showed that patients with a low FEUA, that is, less than 5.5%, respond better to benzbromarone than patients with a higher FEUA do.[13]

Very few studies have investigated the effect of FEUA on the response to allopurinol.[14] Studies have shown that uricosurics increase the clearance of oxypurinol[15–17] because it is likely a substrate for URAT1, as is urate.[18]

We hypothesized that a low FEUA would enhance the urate-lowering response to allopurinol, independently of glomerular filtration, by reducing the renal clearance of oxypurinol. In contrast, because febuxostat is primarily metabolized in the liver, we hypothesized that its efficacy would not be influenced by FEUA.

Accordingly, the objective of this study was to assess whether FEUA modulates the hypouricemic effect of allopurinol or febuxostat. We also examined the relationship between FEUA and oxypurinol concentrations.

PATIENTS AND METHODS

Study population. Patients were recruited from the Vien Gut Medical Center, an outpatient clinic for people with gout for whom clinical, biologic, and imaging data are collected at each visit, according to a predefined protocol. For this study, we analyzed data from newly diagnosed patients between June 2017 and January 2023 for whom we had at least one month of follow-up of urate-lowering therapy (ULT). Gout was diagnosed according to the American College of Rheumatology/EULAR criteria.[19] We excluded women from the study population because the number of women among newly diagnosed patients was very low (n = 30) and because sex is associated with FEUA in different ethnic groups.[20]

This study was approved by the ethics committee of the University of Medicine and Pharmacy, Ho Chi Minh City, Vietnam. Written informed consent was obtained from all participants. The data sets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Treatment with xanthine oxidase inhibitors. All patients were started on a xanthine oxidase inhibitor with a treat-to-target strategy as follows: for those with an eGFR ≥60 mL/min/1.73 m^2, allopurinol was started at 150 mg/d (half a 300-mg tablet) for one month, and the dose was increased by 150 mg increment per month until the serum urate (SUA) target was <360 μM (60 mg/L) for two consecutive months. For patients with tophi, urate arthropathies, or more than six flares per year, the SUA target was <300 μM (50 mg/L). In patients with an eGFR <60 mL/min/1.73 m^2, febuxostat was started at 40 mg/d, with monthly increments of 40 mg until the urate target was achieved, up to 120 mg/d. Prophylactic colchicine (0.5–1 mg/d) was prescribed during the first months of the study. The median length of follow-up was 4.1 months (95% confidence interval [CI] 3.9–4.4).

Demographics and laboratory tests. Data on gout features and comorbidities, SUA levels, serum creatinine levels, and eGFR were collected before the initiation of ULT. SUA levels were monitored monthly until the urate target was obtained. The eGFR was estimated by the Modification of Diet in Renal Disease equation. The FEUA was calculated as (UUA × Pcr)/(PUA × Ucr) × 100, expressed as a percentage, where U and P represent the urinary and plasma concentrations, respectively; PUA represents the SUA levels; and cr represents the creatinine level. It is the ratio of urate clearance to creatinine clearance. The FEUA was determined by using spot-urine samples, as it has been demonstrated to be a valid method,[21] not different from the FEUA estimated by using 24-hour urine collection, which is very difficult to obtain for many patients.

Measurement of plasma oxypurinol and allopurinol concentrations. Among the study population, oxypurinol and allopurinol assays were available for 187 patients. A 50-μL solution of 40-μg/mL acyclovir was added to 500 μL of plasma from the patient sample and vortexed for 10 seconds. Then, 250 μL of a solution of 10% perchloric acid was added, and the mixture was vortexed for one minute. The sample was centrifuged at 17,000 revolutions per minute at 20°C for 10 minutes. Finally, the sample was filtered through a 0.22-μm filter and injected into the chromatographic system. The concentrations of allopurinol and oxypurinol in the plasma samples were identified via LC software from Shimadzu (Kyoto, Japan). Concentrations of

oxypurinol and allopurinol were determined one month and two months after the initiation of allopurinol. Blood samples were taken on average four hours after the morning dose of allopurinol. Among the 187 patients, n = 6 were excluded from the analysis because the plasma oxypurinol concentration was <3 mg/L (20 μM), a low concentration suggesting that these patients were not adherent to treatment, as previously reported.[3]

Statistical analysis. Categorical variables are described as frequencies (percentages); continuous variables are described as medians and interquartile ranges. Comparisons of continuous data were performed using the nonparametric Wilcoxon rank sum test. For comparisons of binary data, Fisher exact test was used.

Associations between FEUA and clinical or biochemical variables were first assessed using univariate logistic regression models. Variables significantly associated with FEUA ($P < 0.05$) in the univariate analysis were then entered into a multivariate logistic regression model to identify independent predictors of FEUA ≤5.5%. The interaction between FEUA and the dose response to either allopurinol or febuxostat (ie, lowering SUA levels) was assessed using linear mixed-effects regression models for urate levels, including a random intercept for patients to account for repeated measurements. The models were adjusted for BMI, eGFR, and treatment doses as fixed effects.

Subgroup analyses were performed with treatment-by-subset interactions tested using the Gail and Simon statistics, a method that analyzes qualitative interactions between treatment effects and patient subgroups.[22] An FEUA cutoff of 5.5% was used, as was done in several studies conducted in Asia.[8,12,23] To further assess whether FEUA provided information independent of eGFR, we also tested for collinearity between FEUA and eGFR using correlation analyses and variance inflation factors.[24]

A linear mixed-effects model for oxypurinol concentrations was used to assess the association between FEUA and oxypurinol concentrations, with a random intercept for patients to account for repeated measures, and adjusted for BMI, eGFR, allopurinol dosage, and SUA levels as fixed effects.

Because treatment allocation was based on renal function (allopurinol in patients with eGFR ≥60 mL/min/1.73 m^2 and febuxostat in those with eGFR <60 mL/min/1.73 m^2), the study was not designed to compare outcomes between the two drugs. All the tests were two sided, with $P \leq 0.05$ considered statistically significant. Statistical analyses were performed via R version 4.2.1 (The R Foundation, https://www.R-project.org/).

RESULTS

Demographic characteristics. The data are from 1,843 male patients who were newly diagnosed with gout and started ULT. Among these patients, 1,547 patients received allopurinol, and 296 received febuxostat. The demographic characteristics of the study population are shown in Table 1. The median (quartile 1–3 [Q1–Q3]) age of patients treated with allopurinol was 45 (38–53) years, with a median BMI of 24.7 (22.8–27) kg/m^2. In these patients, the median SUA level was 543 (464–613) μM, and the median eGFR was 95 (83–106) mL/min/1.73 m^2. Patients treated with febuxostat were older, had lower eGFR, were more likely to have diabetes or hypertension, and had higher SUA levels (all P values < 0.001).

FEUA in the study population. Before the initiation of ULT, the median (Q1–Q3) FEUA in the whole population was 4.0% (3%–5.1%; Figure 1A). It was slightly greater in patients receiving febuxostat who had lower renal function than in those receiving allopurinol: 4.4% (3.2%–6%) and 4.0% (3%–5%), respectively ($P < 0.0001$; Figure 1B).

According to the univariate analysis, an FEUA ≤5.5% was associated with higher eGFR, BMI, and triglyceride levels, and it was associated with younger age (all P values < 0.001). There was no association between FEUA and low-density lipoprotein cholesterol levels (Supplementary Table 1). Age (odds ratio [OR] 0.98 [95% CI

Table 1. Baseline characteristics of patients treated with either allopurinol or febuxostat*

	Allopurinol (n = 1,547)	Febuxostat (n = 296)	*P* value
Age, y	45 [38–53]	57 [48–65]	<0.0001
BMI, kg/m^2	24.7 [22.8–27]	24.0 [21.8–26]	<0.0001
Comorbidities			
Chronic kidney disease	5 (0.3)	157 (53.2)	<0.0001
Type 2 diabetes	114 (7.4)	46 (15.5)	<0.0001
Hypertension	476 (30.8)	153 (51.7)	<0.0001
Cardiovascular diseases	1,022 (66.1)	234 (79.1)	<0.0001
Serum urate level, μM	543 [464–613]	561 [476.8–639.2]	0.007
eGFR, mL/min/1.73 m^2	95 [83–106]	59 [49–76.5]	<0.0001
LDL cholesterol, mmol/L	3.0 [2.4–3.7]	2.8 [2.2–3.6]	0.016
Triglycerides, g/L	2.5 [1.7–3.7]	2.3 [1.5–3.5]	0.017
FEUA, %	4.0 [3–5]	4.4 [3.2–6]	<0.0001

* Data are presented as medians [Q1–Q3] or numbers (percentages). CKD was defined as an eGFR less than 60 mL/min/1.73 m^2. Cardiovascular diseases include coronary artery diseases, heart failure, arrhythmia, and valvular heart disease. BMI, body mass index; CKD, chronic kidney disease; eGFR, estimated glomerular filtration rate; FEUA, fractional excretion of uric acid; LDL, low-density lipoprotein; Q, quartile.

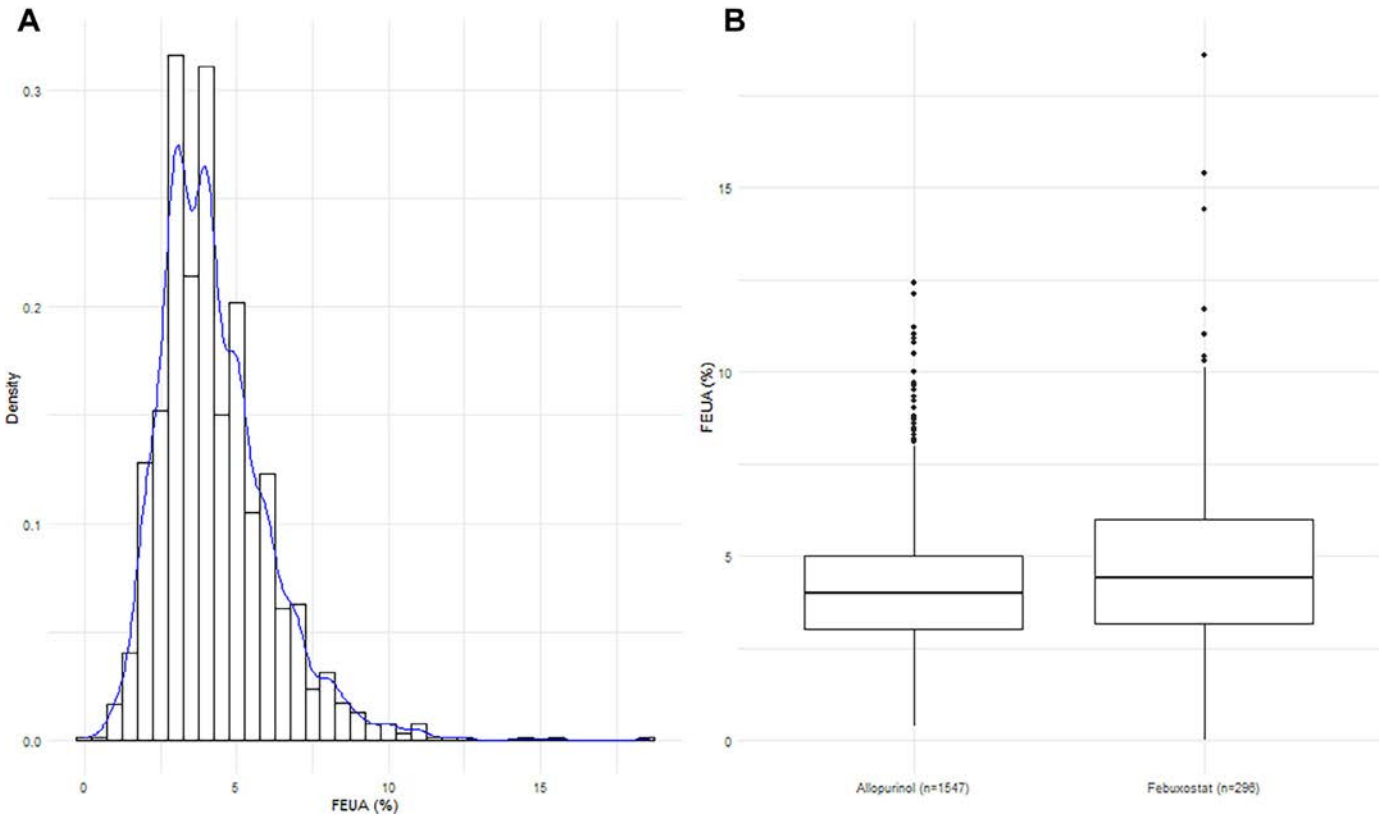


Figure 1. FEUA in the study population. (A) Distribution of the FEUA in 1,843 male patients with gout. (B) FEUA in patients who received allopurinol or febuxostat. FEUA, fractional excretion of uric acid.

0.96–0.99], P = 0.0002) and eGFR (OR 1.01 [95% CI 1–1.02], P = 0.0008) were significantly associated with FEUA ≤5.5% in the multivariate model (Table 2), with a trend toward an association with triglyceride levels (OR 1.06 [95% CI 1.0–1.13], P = 0.064).

Response to allopurinol or febuxostat according to the FEUA. We then assessed the interaction between the FEUA and the hypouricemic effect of allopurinol. We found a significant interaction between FEUA and each 150-mg increment of allopurinol on reducing SUA levels (P = 0.032, adjusted for allopurinol, BMI, and eGFR; Table 3). The hypouricemic effect of allopurinol was more pronounced in patients with a low FEUA: the decreases in SUA levels for each 150-mg increment of allopurinol were −72.37 (CI −74.81 to −69.94) µM and −65.96 (CI −71.29 to −60.62) µM in patients with an FEUA ≤5.5% and an FEUA >5.5%, respectively (Figure 2A and Supplementary Table 2).

In contrast, we found no interaction between FEUA and febuxostat on reducing SUA levels (P = 0.13; Table 4). The association between the decrease in SUA levels for each 40-mg increment of febuxostat did not significantly differ across the two groups of patients defined by an FEUA ≤5.5% or >5.5% (Figure 2B and Supplementary Table 3). Variance inflation factors for all variables, including FEUA and eGFR, were below 2, indicating no collinearity in our models (Supplementary Figure 1).

Table 2. Multivariate analysis of factors associated with an FEUA ≤5.5%*

	OR (95% CI)	P value
Age, y	0.98 (0.96–0.99)	0.0002
BMI, kg/m^2	1.02 (0.99–1.06)	0.19
eGFR, mL/min/1.73 m^2	1.01 (1–1.02)	0.0008
Triglycerides, g/L	1.06 (1–1.13)	0.064

* BMI, body mass index; CI, confidence interval; eGFR, estimated glomerular filtration rate; FEUA, fractional excretion of uric acid; OR, odds ratio.

Table 3. Adjusted analysis of the effect of allopurinol on serum urate levels with a linear regression model (n = 1,537 patients)*

Variables	Beta	95% CI	P value
FEUA > 5.5%	−39.93	−53.27 to −26.59	<0.0001
Allopurinol (each 150 mg)	−72.37	−74.81 to −69.94	<0.0001
BMI, kg/m^2	3.75	2.56 to 4.95	<0.0001
eGFR, mL/mn/1.73 m^2	−0.26	−0.5 to −0.02	0.035
FEUA > 5.5% × allopurinol interaction	6.46	0.33 to 12.59	0.032[a]

* BMI, body mass index; CI, confidence interval; eGFR, estimated glomerular filtration rate; FEUA, fractional excretion of uric acid.
[a] Gail and Simon interaction test.

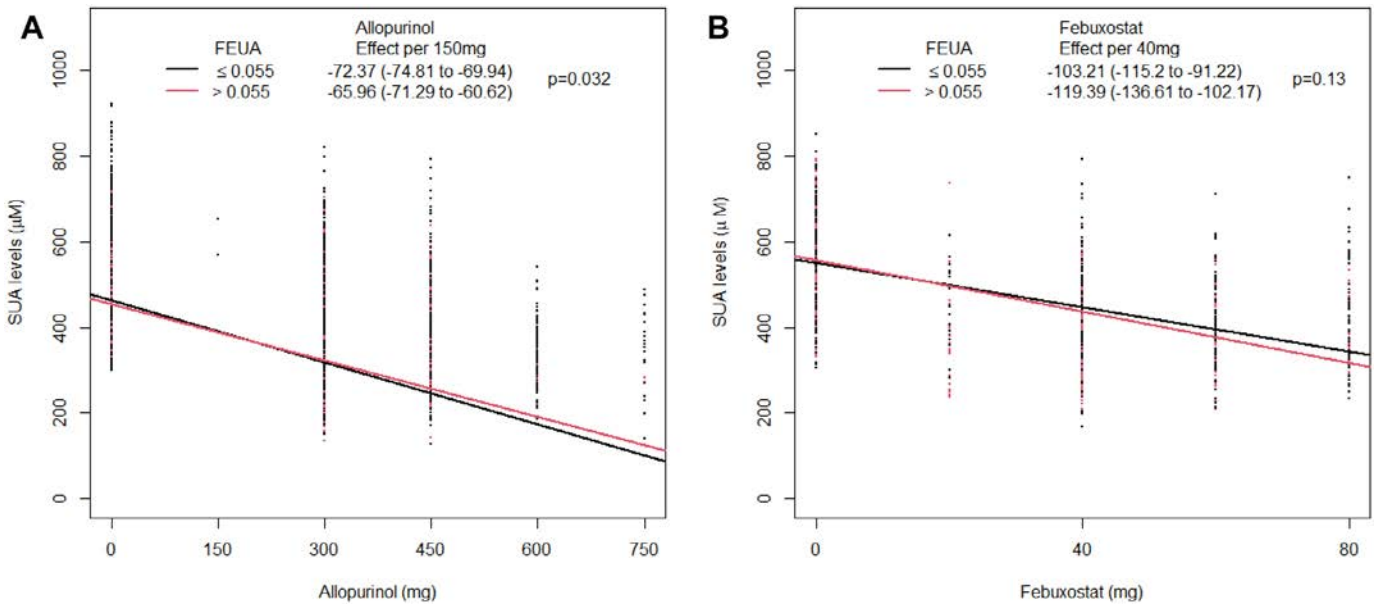


Figure 2. FEUA and response to xanthine oxidase inhibitors. (A) Response to allopurinol according to the FEUA. (B) Response to febuxostat according to the FEUA. FEUA, fractional excretion of uric acid; SUA, serum urate. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25695/abstract.

Relationship between oxypurinol concentrations and the FEUA. To explain the relationship we observed between an FEUA ≤5.5% and a better response to allopurinol, we evaluated the concentrations of oxypurinol, an active metabolite of allopurinol (Supplementary Table 4), according to the FEUA, in a subgroup of patients (n = 181). These patients had a median BMI of 24.7 (Q1–Q3: 22.5–27.2) kg/m^2 and a median eGFR of 93 (Q1–Q3: 79–106) mL/min/m^2. Only two patients took diuretics. We found a significant association between lower FEUA and higher oxypurinol concentrations (P = 0.032, adjustment for BMI, eGFR, allopurinol dosage, and SUA levels). There was a strong association between the concentrations of allopurinol and oxypurinol (P < 0.0001; Supplementary Table 5).

Table 4. Adjusted analysis of the febuxostat effect on serum urate levels via a linear regression model (n = 291 patients)*

Variables	Beta	95% CI	*P* value
FEUA > 5.5%	−39.76	−67.04 to −12.48	0.004
Febuxostat (each 40 mg)	−103.21	−115.2 to −91.22	<0.0001
BMI, kg/m^2	1.34	−1.48 to 4.16	0.35
eGFR, mL/mn/1.73 m^2	−0.74	−1.2 to −0.29	0.001
FEUA > 5.5% × febuxostat interaction	−15.26	−36.56 to 6.04	0.13[a]

* BMI, body mass index; CI, confidence interval; eGFR, estimated glomerular filtration rate; FEUA, fractional excretion of uric acid.
[a] Gail and Simon interaction test.

DISCUSSION

This study identified FEUA as a new covariate that modulated the hypouricemic effect of allopurinol. Although statistically significant, the difference in the hypouricemic effect of allopurinol according to FEUA was modest and of uncertain clinical relevance. Nevertheless, this finding reinforces the concept that the renal mechanisms regulating urate and oxypurinol handling are quite similar. A low FEUA was associated with a better response to allopurinol, independent of the GFR and BMI, which was likely due to a decrease in oxypurinol renal clearance. Although FEUA and eGFR were modestly associated, there was no collinearity supporting the independent contribution of FEUA beyond glomerular filtration to the response to allopurinol.

The kidney is responsible for the majority of urate excretion in humans.[1] Renal excretion of urate is dependent on both the filtered urate load and the fraction of the filtered load that is eliminated in the urine, that is, the FEUA,[9,25] which is under the control of multiple urate transporters in the proximal tubule.[26]

We observed that the influence of FEUA on the allopurinol response is independent of the glomerular function. Although eGFR determines the filtered load of oxypurinol, FEUA captures the balance between secretion and reabsorption in the tubule, which critically shapes oxypurinol clearance. Thus, the improved hypouricemic response in patients with low FEUA, albeit modest, likely reflects increased tubular reabsorption of oxypurinol rather than decreased glomerular filtration.

Several factors have been reported to lower the FEUA and are therefore risk factors for hyperuricemia and gout. Heredity,

which plays a key role in the genesis of the disease,[27] is undoubtedly one of the most important. Indeed, several SUA-associated single nucleotide polymorphisms have been reported to be associated with a lower FEUA,[28] notably at the *SLC2A9* locus, with differences depending on ethnicity groups.[20] In contrast, the Q141K variant of *ABCG2* has not been found to be associated with FEUA variations[9] but might modify the response to allopurinol, although the data are conflicting. Other clinical variables reported to influence FEUA are male sex and BMI.[20] We did find an association between BMI and FEUA ≤5.5% in the univariate analysis, which was no longer significant in the multivariate analysis, suggesting an indirect effect of BMI on FEUA. Notably, a study conducted in people with overweight and obesity revealed that this association was observed only when urate levels increased following a fructose load.[29]

The doses of allopurinol needed to lower the SUA levels below 360 μM vary from one individual to another and depend on clinical and biologic factors. According to a recent study, these factors include ethnicity, BMI, pretreatment SUA levels, and renal function.[6] In contrast to what has been reported in previous studies,[3,30] neither the *ABCG2* genotype (rs2231142) nor the presence of diuretics were predictors of allopurinol efficacy in more recent papers.[6,7]

Multivariate analysis revealed that the hypouricemic effect of allopurinol was greater in patients with a low FEUA, that is, less than ≤5.5%, independent of BMI and kidney function. These data suggest that the mechanisms involved in the tubular secretion and excretion of urate could also modify the pharmacokinetics of allopurinol, independent of the GFR.

In contrast, we found no interaction between febuxostat and FEUA, a result that was expected because febuxostat is metabolized mainly in the liver and is virtually not eliminated by the kidneys. In line with our findings, Qi et al reported that the response to febuxostat, as assessed by the percentage of patients achieving SUA levels of less than 360 μM, was identical in patients with an FEUA ≥5.5% or <5.5%.[31]

Because the renal elimination of oxypurinol is the most important aspect of the pharmacokinetics of allopurinol,[15] we explored the association between FEUA and oxypurinol concentrations. We found that oxypurinol concentrations were higher in patients with a low FEUA, independent of BMI, eGFR, allopurinol dosage, and SUA levels. This is probably because the oxypurinol filtered at the glomerulus is then largely reabsorbed by the proximal tubules.

The mechanisms underlying the renal elimination of oxypurinol in patients with gout have not been extensively studied. Unlike allopurinol, which has a very short half-life of one hour, the half-life of oxypurinol is approximately 23 hours.[15] The finding that the renal clearance rate of oxypurinol was lower than the GFR but two to three times greater than that of urate led to the hypothesis that oxypurinol could be reabsorbed by renal tubules, such as urate.[32] This hypothesis was further supported by the observation of an increase in renal clearance of oxypurinol by uricosurics in patients with gout[15,16,33] but also in Dalmatian dogs.[32] In this breed of dog, urate excretion equals or exceeds the GFR because of a mutation in the *SLC2A9* gene, which codes for the urate transporter GLUT9,[34] explaining why renal urate reabsorption is lost almost entirely.[34]

It was subsequently shown that oxypurinol is a substrate of URAT1, highlighting the similarity in the renal handling of urate and oxypurinol. Thus, URAT1 might play a key role in the renal reabsorption of oxypurinol to the difference of OAT4, which seems to not be involved in this process.[18] Other studies have suggested that GLUT9 may also transport oxypurinol.[33,35]

Our study has limitations that must be considered. We recognize that the difference in the magnitude of the hypouricemic effect of allopurinol according to the FEUA is not large, so the clinical relevance at the individual level of this result is unknown. Our study population did not include women, which is explained by the very low prevalence of women treated at the Vien Gut Medical Center. Therefore, it is not possible to generalize our results to all patients with gout. Furthermore, the phenotype of the patients in our study are different from those included in similar studies conducted notably in Oceania and the US[3,6,7,20]: overweight and obesity are rare in our study population. In addition, individuals treated with allopurinol in our study had a preserved renal function (eGFR ≥ 60 mL/min/1.73 m^2), which may restrict the generalizability of our findings to gout populations with a higher prevalence of chronic kidney disease, in which the relationships between FEUA, oxypurinol clearance, and allopurinol response may differ. It is therefore difficult to compare our results with those already published in this area. Lastly, the oxypurinol threshold of 3 mg/L that we used to define nonadherent patients is questionable, as this threshold varies according to patient characteristics that we did not take into account.[36] In conclusion, our study revealed that, in addition to the GFR, the FEUA is a renal parameter that influences the response to allopurinol but not to febuxostat.

ACKNOWLEDGMENTS

We thank the patients who participated in this study, as well as all the medical and nonmedical staff at Vien Gut. DeepL was used to check spelling and grammar.

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 Richette 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. Dalbeth N, Gosling AL, Gaffo A, et al. Gout. Lancet 2021;397(10287): 1843–1855.
2. Sekine M, Okamoto K, Pai EF, et al. Allopurinol and oxypurinol differ in their strength and mechanisms of inhibition of xanthine oxidoreductase. J Biol Chem 2023;299(9):105189.
3. Stamp LK, Chapman PT, Barclay M, et al. Relationships between allopurinol dose, oxypurinol concentration and urate-lowering response-in search of a minimum effective oxypurinol concentration. Clin Transl Sci 2020;13(1):110–115.
4. Stamp LK, Wallace M, Roberts RL, et al. ABCG2 rs2231142 (Q141K) and oxypurinol concentrations in people with gout receiving allopurinol. Drug Metab Pharmacokinet 2018;33(6):241–242.
5. Wallace MC, Roberts RL, Nanavati P, et al. Association between ABCG2 rs2231142 and poor response to allopurinol: replication and meta-analysis. Rheumatology (Oxford) 2018;57(4):656–660.
6. Wright DFB, Hishe HZ, Stocker SL, et al. The development and evaluation of dose-prediction tools for allopurinol therapy (Easy-Allo tools). Br J Clin Pharmacol 2024;90(5):1268–1279.
7. Hishe HZ, Stocker SL, Stamp LK, et al. The impact of genetic variability in urate transporters on oxypurinol pharmacokinetics. Clin Transl Sci 2023;16(3):422–428.
8. Ichida K, Matsuo H, Takada T, et al. Decreased extra-renal urate excretion is a common cause of hyperuricemia. Nat Commun 2012; 3(1):764.
9. Hoque KM, Dixon EE, Lewis RM, et al. The ABCG2 Q141K hyperuricemia and gout associated variant illuminates the physiology of human urate excretion. Nat Commun 2020;11(1):2767.
10. Guan H, Zheng Y, Zhou X, et al. Efficacy of different urinary uric acid indicators in patients with chronic kidney disease. BMC Nephrol 2020;21(1):290.
11. Hsiao PJ, Liao CY, Kao YH, et al. Comparison of fractional excretion of electrolytes in patients at different stages of chronic kidney disease: a cross-sectional study. Medicine (Baltimore) 2020;99(2):e18709.
12. Yan F, Xue X, Lu J, et al. Superiority of low-dose benzbromarone to low-dose febuxostat in a prospective, randomized comparative effectiveness trial in gout patients with renal uric acid underexcretion. Arthritis Rheumatol 2022;74(12):2015–2023.
13. Xue X, Yuan X, Han L, et al. Effect of clinical typing on serum urate targets of benzbromarone in Chinese gout patients: a prospective cohort study. Front Med (Lausanne) 2022;8:806710.
14. Perez-Ruiz F, Alonso-Ruiz A, Calabozo M, et al. Efficacy of allopurinol and benzbromarone for the control of hyperuricaemia. A pathogenic approach to the treatment of primary chronic gout. Ann Rheum Dis 1998;57(9):545–549.
15. Day RO, Graham GG, Hicks M, et al. Clinical pharmacokinetics and pharmacodynamics of allopurinol and oxypurinol. Clin Pharmacokinet 2007;46(8):623–644.
16. Stocker SL, Williams KM, McLachlan AJ, et al. Pharmacokinetic and pharmacodynamic interaction between allopurinol and probenecid in healthy subjects. Clin Pharmacokinet 2008;47(2):111–118.
17. Perez-Ruiz F, Sundy JS, Miner JN, et al; RDEA594-203 Study Group. Lesinurad in combination with allopurinol: results of a phase 2, randomised, double-blind study in patients with gout with an inadequate response to allopurinol. Ann Rheum Dis 2016;75(6):1074–1080.
18. Iwanaga T, Kobayashi D, Hirayama M, et al. Involvement of uric acid transporter in increased renal clearance of the xanthine oxidase inhibitor oxypurinol induced by a uricosuric agent, benzbromarone. Drug Metab Dispos 2005;33(12):1791–1795.
19. Neogi T, Jansen TL, Dalbeth N, et al. 2015 Gout classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheumatol 2015;67(10): 2557–2568.
20. Narang RK, Vincent Z, Phipps-Green A, et al. Population-specific factors associated with fractional excretion of uric acid. Arthritis Res Ther 2019;21(1):234.
21. Kannangara DR, Ramasamy SN, Indraratna PL, et al. Fractional clearance of urate: validation of measurement in spot-urine samples in healthy subjects and gouty patients. Arthritis Res Ther 2012;14(4): R189.
22. Gail M, Simon R. Testing for qualitative interactions between treatment effects and patient subsets. Biometrics 1985;41(2):361–372.
23. Xue X, Sun M, Yan F, et al. Superiority of low dose benzbromarone add-on to low-dose febuxostat compared with febuxostat monotherapy in gout with combined-type hyperuricemia. Arthritis Care Res (Hoboken) 2024;76(5):703–711.
24. Fox J, Monette G. Generalized collinearity diagnostics. J Am Stat Assoc 1992;87(417):178–183.
25. Perez-Ruiz F, Calabozo M, Erauskin GG, et al. Renal underexcretion of uric acid is present in patients with apparent high urinary uric acid output. Arthritis Rheum 2002;47(6):610–613.
26. Takada T, Miyata H, Toyoda Y, et al. Regulation of urate homeostasis by membrane transporters. Gout Urate Cryst Depos Dis 2024;2(2): 206–219.
27. Major TJ, Takei R, Matsuo H, et al. A genome-wide association analysis reveals new pathogenic pathways in gout. Nat Genet 2024;56(11): 2392–2406.
28. Köttgen A, Albrecht E, Teumer A, et al; LifeLines Cohort Study; CARDIoGRAM Consortium; DIAGRAM Consortium; ICBP Consortium; MAGIC Consortium. Genome-wide association analyses identify 18 new loci associated with serum urate concentrations. Nat Genet 2013;45(2):145–154.
29. Dalbeth N, Phipps-Green A, House ME, et al. Body mass index modulates the relationship of sugar-sweetened beverage intake with serum urate concentrations and gout. Arthritis Res Ther 2015; 17(1):263.
30. Graham GG, Kannangara DR, Stocker SL, et al. Understanding the dose-response relationship of allopurinol: predicting the optimal dosage. Br J Clin Pharmacol 2013;76(6):932–938.
31. Qi H, Sun M, Terkeltaub R, et al. Response to febuxostat according to clinical subtypes of hyperuricemia: a prospective cohort study in primary gout. Arthritis Res Ther 2023;25(1):241.
32. Elion GB, Yü TF, Gutman AB, et al. Renal clearance of oxipurinol, the chief metabolite of allopurinol. Am J Med 1968;45(1):69–77.
33. Stocker SL, McLachlan AJ, Savic RM, et al. The pharmacokinetics of oxypurinol in people with gout. Br J Clin Pharmacol 2012;74(3): 477–489.
34. Bannasch D, Safra N, Young A, et al. Mutations in the SLC2A9 gene cause hyperuricosuria and hyperuricemia in the dog. PLoS Genet 2008;4(11):e1000246.
35. Anzai N, Ichida K, Jutabha P, et al. Plasma urate level is directly regulated by a voltage-driven urate efflux transporter URATv1 (SLC2A9) in humans. J Biol Chem 2008;283(40):26834–26838.
36. Smith-Diaz N, Stocker SL, Stamp LK, et al. An allopurinol adherence tool using plasma oxypurinol concentrations. Br J Clin Pharmacol 2023;89(7):1956–1964.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 78, No. 6, June 2026, pp 788–797
DOI 10.1002/acr.25698

REVIEW ARTICLE

Advances in Synovial Fluid Analysis for the Diagnosis of Crystal Arthropathies

Tom Niessink,[1,2] Tim L. Jansen,[2] Matthijs Janssen,[2] and Cees Otto[1]

Synovial fluid analysis remains to be a critical aspect of the diagnosis of crystal arthropathies. The gold standard, compensated polarized light microscopy (CPLM), has significant shortcomings due to poor reproducibility and improper training of specialists. Especially in cases of artifacts, low crystal counts, and combinations of crystals, errors are frequently made. There are several new techniques in development to improve the accuracy of crystal identifications. With modified CPLM, microscopes, often aided by machine-learning algorithms, enhance the image contrast, which significantly improves sensitivity. Examples are lens-free polarized light microscopy and polychromatic polarized light microscopy. These techniques have relatively low costs and are easy to use. Another approach is using Raman spectroscopy alone or as a second (verification) step for the objective identification of crystals. This improves objectivity and specificity of diagnosis. Examples include a point-of-care Raman spectroscope and integrated Raman-polarized light microscopy. Although these are more accurate than the microscopy-based platforms, they are often more expensive and require some additional training. Both Fourier transform infrared spectroscopy (FTIR) and scanning electron microscopy (SEM) have also been used for crystal identification in research. FTIR contributes specificity in a manner similar to Raman methods. SEM adds a higher resolution than optical methods and provides a clear view of morphology. In this narrative review, we provide an overview of the available literature comparing innovational techniques in synovial fluid analysis.

Introduction

Crystalline arthropathies are a class of generally autoinflammatory diseases caused by deposited crystals in the synovial space. The most important crystalline arthropathies include gout, calcium pyrophosphate deposition disease (CPDD), and basic calcium phosphate (BCP)–associated arthritis.[1,2] The presentation of these crystals, alone or in combinations, can trigger the inflammasome via a second signal (for example, interleukin-1 β) and offset a strong inflammatory response.[3] The crystal arthropathies can be very painful and can limit patients in daily activities. When these conditions are left untreated, patients can experience permanent joint damage and disability.[4]

To enable administering low-cost and high-quality health care, a detailed and thorough determination of all crystals present in the synovial fluid is required, such that effective therapy can be started with certainty. The only definitive method of diagnosing crystalline arthropathies is by the direct identification of the present potentially causative crystals within the synovial fluid drawn from affected patients.[5] Current methods to perform synovial fluid analyses, predominantly compensated polarized light microscopy (CPLM), are known to be highly dependent on the quality of the microscope used, the number of crystals present, and skills of the trained analyst, and therefore, one may raise doubts on the correctness of the clinical diagnosis.[6–8] There is a demand for more objective technologies. In this review, different types of crystals that commonly occur in synovial fluid will be discussed. This review intends to present recent advances in technology for the diagnosis of crystal arthropathies and discuss the advantages and limitations of each diagnostic method. This is especially important

Supported by the PPP Allowance made available by Health-Holland, Top Sector Life Sciences & Health, to Stichting ReumaNederland to stimulate public-private partnerships.

[1]University of Twente, Enschede, the Netherlands; [2]VieCuri Medical Centre, Venlo, 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/acr.25698).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25698.

Address correspondence via email to Tom Niessink, PhD, at t.niessink@utwente.nl.

Submitted for publication June 2, 2025; accepted in revised form October 22, 2025.

Table 1. Harvey Ball assessment of the compared innovations*

Modality	Sensitivity	Specificity	Ease of use	Readiness[a]	Cost
CPLM	◑	◕	◯	◯	◔
CPLM + PC	◔	◕	◯	◯	◔
SCPLM	◔	◑	◔	◑	◔
DL-HPM	◔	◑	◔	◑	◯
Polychromatic PLM	◔	◑	◯	◑	◔
POC Raman	◑	◯	◑	◔	◑
SRS microscopy	◔	◔	◕	◑	◕
iRPolM	◔	◯	◔	◔	◕
(s)FTIR	◔	◯	◑	◑	●
SEM/EDX	◯	◑	◑	◑	●
In vivo Raman	◕	◑	◔	◕	◑
Sweat SERS	◑	◕	◑	◕	◕

* ● = worst, ◕ = moderate, ◑ = median, ◔ = good, ◯ = best. CPLM, compensated polarized light microscopy; DL-HPM, deep-learning holographic polarization microscopy; iRPolM, integrated Raman-polarized light microscopy; PC, phase contrast; PLM, polarized light microscopy; POC, point-of-care; SCPLM, single shot computed polarized light microscopy; SEM/EDX, scanning electron microscopy with energy dispersive x-ray spectroscopy; SERS, surface-enhanced Raman spectroscopy; (s)FTIR, (synchrotron) Fourier transform infrared spectroscopy; SRS, stimulated Raman scattering.
[a] With readiness, we imply technology readiness level or time to clinic. All evaluations are performed with respect to CPLM as the gold standard.

for calcium crystals, as these are difficult to diagnose.[9] An overview of these methods is given in Table 1, together with a scoring on criteria important for implementation (sensitivity, specificity, ease of use, readiness, and costs).

Classes of synovial fluid crystals and crystal mimickers

There is a range of different crystals that can be found in synovial fluids, most of which are directly related to the arthropathy involved. Synovial fluids can also simultaneously contain two or more types of crystals, a condition known as mixed crystal disease.[10] Classes of crystals and other biologic deposits are given in Table 2. Classes of foreign bodies with known prevalence in synovial fluid are given in Table 3.

For the diagnosis of crystalline arthropathies, especially in suspected gout and CPDD, both EULAR and the American College of Rheumatology (ACR) recommend microscopical analysis of synovial fluid aspirates.[5,11] Identification of monosodium urate monohydrate (MSU) and calcium pyrophosphate (CPP) crystals in synovial fluid in the setting of arthritis will help clinicians and will often lead to a direct and definitive diagnosis of the underlying condition.

CPLM

The gold standard in synovial fluid identification is CPLM. CPLM is a well-established technique and was first applied in rheumatology in the 1960s.[12,13] CPLM enables crystals based on their morphology and their polarizing (birefringent) properties. MSU crystals, for example, show a strong, negative birefringence and are typically needle shaped, whereas CPP crystals show a weak positive birefringence and more rhomboid or rod-like shapes (Supplementary Figure 1A and B).[14]

Although rheumatologists are increasingly using CPLM, it has limited reliability as a diagnostic tool. First reports on poor reproducibility, sensitivity, and specificity stem from 1986 to 1998.[15–18] Some of the surveyed centers had a mere 50% accuracy rate in the diagnosis. More recent studies showed the accuracy of detection is heavily dependent on the observer's training and structured training and further education are not uniform across countries.[6,7] To find a correct diagnosis, an analyst should not only know how to identify crystals but also know how to handle a specialized microscope. CPLM enables the identification of the sign of birefringence and morphology of the object but does not enable precise characterization of crystal composition. Most often, an identification of MSU is more accurate than CPP. Berendsen et al (2017)[6] showed a correct identification of pathognomonic MSU crystals by 81% and CPP crystals by 68% of participants. Errors in an individual diagnosis can have a large impact on the patient's treatment success, prognosis, and long-term outcome.

MSU might also be mistaken for other synovial fluid crystals with similar morphology and/or birefringence. Cholesterol for example (Supplementary Figure 1D), which is often presented as large geometric plates with notched corners with varying negative birefringence, may present dendrites with similarities to MSU and can sometimes be found in patients with gout.[19]

Hydroxyapatite crystals are small (sub micrometer) and lack birefringence, which is problematic for CPLM detection.[20,21] For detection of these crystals, an additional alizarin red staining is required, but the effectiveness of this method has been questioned.[18,22] Deposited glucocorticoids are often present as amorphous masses and are therefore difficult to recognize

Table 2. An overview of known crystals and biologic deposits which can be identified in synovial fluid samples*

Crystal type	Description
Monosodium urate	Presence is evidence for the diagnosis of gout in affected patients. Patients with gout have hyperuricemia to such an extent that uric acid precipitates in their synovial fluids as MSU. When these MSU crystals are recognized by resident synovial macrophages, they will activate the NLRP-3 inflammasome and release pro-inflammatory cytokines like interleukin-1, which promotes a large influx of neutrophils in the synovial space.[3,66] Symptoms include a rapid onset of severe pain, swelling, redness, heat, and loss of function. Gout mostly affects the first metatarsophalangeal or ankle joint but can also affect other joints, such as the knee joints. Renal or ureteral urate stones are a common comorbidity (odds ratio 1.4) in patients with gout.[67]
Calcium pyrophosphate	Presence is evidence for CPDD in specific clinical scenarios. These crystals occur when pyrophosphate anions form with calcium cations.[12] The exact pathologic pathway behind this process is not fully understood, but prevalence shows a strong correlation with arthritis of larger joints, mostly in elderly patients, although hereditary forms also affect patients at a younger age in Gitelman syndrome.[12–14] The severity of symptoms is similar to gout, and the disease was therefore formerly known as pseudogout. Acute CPP arthritis mostly occurs in the knee and the wrist but other joints can also be affected.
Hydroxyapatite	BCP crystals relate to HADD, including the Milwaukee shoulder. HADD is known to be a very destructive joint disease.[20,68] The inflammatory response in HADD is similar to gout and CPDD.[68] HADD can occur in every joint but predominantly affects shoulder joints, followed by the knee, hip, elbow, wrist, and ankle joints.[20] BCP is also known to be highly prevalent in osteoarthritis and is suspected to be associated with the pathogenesis of this disease.[69]
Calcium oxalate	Associates with oxalate arthritis. This arthritis is often preceded by hyperoxaluria, which is mostly an inherited metabolic disorder but can also occur in renal insufficient patients with a high oxalate or vitamin C intake.[18,19] Oxalate crystals occur in two forms, m-CO or d-CO. Hyperoxaluria is a serious condition and may lead to kidney failure when left untreated.
Cholesterol monohydrate plates	Can be found in the synovial fluid of patients with rheumatoid arthritis, gout, and other rheumatic diseases.[19,70–72] They are rare, and a clear relationship between cholesterol deposits and inflammation has not been found.[73] When present in a bursal sack or joint, the deposits cannot be cleared easily, so a chronic swelling may persist.
Maltes cross birefringent spherules	Maltese cross birefringence can be caused by either lipid droplets or polysaccharides (starch) or can be a rare presentation of monosodium urate.[14,44] The clinical relevance of these objects remains elusive. Maltese cross birefringence is a result of radial (spherical) symmetry of the organization of anisotropic molecules within the particle. It has been quantitively described and is not a specific material property.
Calcites	The presence of calcium carbonate crystals in the joint space is a recent finding,[46] and the clinical relevance of these crystals remains to be studied. In vitro assays have demonstrated their pro-inflammatory properties, and the crystals are highly prevalent in patients with end-stage osteoarthritis.[46]

* Many of these are known to be associated with inflammatory joint disease. BCP, basic calcium phosphate; CPDD, calcium pyrophosphate deposition disease; CPP, calcium pyrophosphate; d-CO, dihydrate calcium oxalate; HADD, hydroxyapatite deposition diseases; m-CO, monohydrate calcium oxalate; MSU, monosodium urate monohydrate.

(Supplementary Figure 1C). Dust particles, silicates, foreign bodies, and other artifacts are also a cause of misinterpretation in CPLM.

The coexistence of multiple types of crystals can further complicate the identification; for example, CPP crystals can be found together with BCP crystals, like hydroxyapatite, in patients with osteoarthritis.[23] Sometimes CPP and MSU crystals can be found in the same joint, leading to the diagnosis of both gout and CPDD.[24–26] This is a difficult diagnosis to make, however, as finding a few CPP crystals lying among thousands of MSU crystals can easily be compared to a needle-in-a-haystack situation. The presence of a combination of CPP and MSU crystals is also being referred to as mixed crystal arthritis.[10]

Thus, the use of CPLM in specific cases may be problematic, especially in more difficult cases of crystalline arthritis. Therefore, rheumatologists worldwide are depending on a subjective, imperfect, and sometimes unreliable technique and make high-impact clinical decisions based on questionable data, especially when applied to patients with a low a priori risk of gout or acute CPP arthritis.

Alternatives for CPLM-based crystal identification

Due to the problems associated with CPLM, several methods have been developed or are under development to improve the reliability of microscopic synovial fluid analysis.[27] These methods either rely on different contrast mechanisms (alizarin red staining and phase contrast microscopy) or are designed to be cheaper alternatives, such as lens-free polarized light microscopy, or use a combination of high-tech optical elements and computational enhancement to improve on CPLM image contrast, such as single shot computational polarized light microscopy (SCPLM), deep-learning holographic polarization microscopy, and polychromatic polarization microscopy. We will discuss these methods in the following sections.

Alizarin red staining. As BCP crystals have very low birefringence (+0.007, compared to –0.3 of MSU), identification of these crystals by CPLM is not straightforward with the equipment that is commonly clinically available. Alizarin red is a soluble sodium salt of alizarin sulfonic acid, which binds to calcium to

Table 3. An overview of foreign body material with known occurrence in the joint and synovial fluid

Foreign body	Description
Injected glucocorticoids	Glucocorticoid injections are anti-inflammatory therapeutics used for different types of arthritis, including crystalline arthritis. Glucocorticoids themselves are also crystals, which may be pro-inflammatory on their own and cause postinjection flares.[74,75] Due to their varying morphology, misclassification of these particles is a risk.
Micrometals and oxide minerals	It is well known that friction between orthopedic joint implants can lead to wear particle generation.[76,77] Common metal components of orthopedic implants include titanium, zirconium, and cobalt-chromium alloys.[78] Oxide minerals, such as titanium dioxide, can be found in and around the prosthetic joint[47] but have also been found in healthy joints,[48] where their presence is potentially related to ingestion of these particles though food. Titanium is for instance also a food additive and may enter the body by absorption through the skin or gut. These minerals are possibly pro-inflammatory[79] and are therefore potentially clinically relevant.
Microplastics	These have been identified in both prosthetic joints[47] and native joints.[80] Microplastics are pro-inflammatory, and their presence has been associated with an increased risk of cardiovascular disease.[81] Microplastics might enter the body via ingestion and can be found even in otherwise healthy individuals.[82] There is ample scientific evidence between the presence of microplastics and inflammatory joint disease, although some case reports on periprosthetic inflammation associated with foreign body infiltration exist too.[83]

form an orange to red pigment. The method is inexpensive and can be performed in any modern pathology laboratory. Although frequently used in research,[22,28] the method is not used often in clinical practice.[29] Alizarin red binds to calcium from any form of calcium-containing crystal and is therefore not specific for BCP. Furthermore, alizarin red assays are known to be not sensitive for small microcrystals, which makes it unsuitable for identification of often tiny BCP crystals.

Phase contrast microscopy. Where polarized light microscopy provides image contrast based on birefringence, phase contrast relies on the phase shift that occurs when light is scattered by (biologic) specimens such as a crystal. As light can be described as a wave, the phase of a light ray varies between 0 and 2π between a crest of the wave and the next crest. When a light ray scatters, the phase jumps with π/2. After passage of the scattered light through a phase plate, where it acquires another π/2 phase shift, it is recombined with the nonscattered wave. The interference between the scattered and nonscattered light ray can be constructive when the rays are in phase with each other or destructive if the rays are shifted half a wavelength (or π) in phase. Similar to a compensated polarized light microscope, a phase contrast microscope contains a series of specialized optical elements to allow for phase contrast microscopy, such as an annular diaphragm and a phase plate. The phase plate translates the gained phase differences into intensity differences and therefore into visible image contrast. The idea to use phase contrast microscopy for synovial fluid analysis is not novel, as it was already introduced by Kohn et al in 1968.[12] Phase contrast microscopy is an inexpensive and widely available technique and can be combined with polarized light microscopy in a single microscope platform. The visibility for small CPP crystals is significantly better with phase contrast microscopy compared to polarized light microscopy (Supplementary Figure 2). Unlike CPLM, phase contrast microscopy never became popularized in clinical use, and statistics on sensitivity and specificity are therefore unknown. Phase contrast does not provide a specificity to distinguish between MSU and CPP, and crystal identification has to be done on the basis of morphology or with an (integrated) polarized light microscope. It is to be expected that the sensitivity of a phase contrast microscope, when integrated with polarized light microscopy, is increased for CPP, with similar specificity. Operating of a phase contrast microscope is not more complex than a CPLM, and the method is readily available for clinicians; associated purchase costs are comparable to CPLM.

SCPLM. In SCPLM, an ordinary microscope platform is altered with a left hand circular polarizer and a specialized camera with four orthogonal polarizing filters per pixel.[30] In a single shot, the camera takes four images (one for every polarization) and then computes a pseudocolor CPLM image. These pseudocolor images offer enhanced contrast compared to CPLM, and the microscope can operate at several magnifications. As the method is based on an ordinary microscope platform and the additionally required optical components are relatively affordable, it could be possible to upgrade existing microscopes in rheumatology departments to SCPLM systems. Expert raters tested SCPLM and demonstrated higher crystal detection rates for both CPP and MSU crystals.[31] Statistics on diagnostic sensitivity and specificity are as of yet unknown. The developers of this device expect the sensitivity to be higher than CPLM, but as birefringence is the discriminating factor among crystal types, specificity should be similar. Operation of this device should be comparable to CPLM, although additional software is required. The first studies with patient data have been published,[30] but additional research is necessary before market entry. The costs should be similar to CPLM.

Lens-free polarized light microscopy and deep-learning based holographic polarization microscopy. In an attempt to create a more affordable and practical platform for synovial fluid analysis, the lens-free polarized light microscope is only several centimeters large.[32] The components of a lens-free polarized light microscope are mostly similar to that of a CPLM, including polarizing filters and a compensator. Instead of lenses, a large complementary metal oxide semiconductor (CMOS) opto-electronic image sensor is used to capture light directly from the synovial fluid. The method is designed to mimic CPLM as close as possible, using Jones calculus as a physical basis for its image-processing algorithm. For every sample, two holographic (depth-sensitive) images are captured with different analyzer angles, and a pseudocolor CPLM image is calculated. The method has been demonstrated to accurately depict MSU crystals and glucocorticoid deposits, although no proper diagnostic accuracy study was performed. The main advantage of this method is a very large field of view of 20 mm^2, which is directly equal to the size of the CMOS chip and two orders of magnitude larger compared to CPLM, and the small size of the device. The resolution of the device is stated to be similar to that of ordinary CPLM.

In a second iteration of the lens-free device, deep-learning algorithms were added to improve image contrast. A SCPLM microscope was used to generate training data for the deep-learning algorithm, and the output of the deep-learning based holographic polarization microscope indeed closely resembles the enhanced contrast offered by SCPLM. An additional advantage of adding deep learning is that only one image with one analyzer angle is used for imaging. At this point, no clinical performance studies have been performed, and the sensitivity and specificity of this device is unknown. The developers expect that performance matches that of SCPLM, thus an increased sensitivity (compared to CPLM) and similar specificity. Like SCPLM, there are some additional steps in software required to get to a diagnosis. There is no price range known thus far, but the device is designed to be low cost.

Polychromatic polarized light microscopy. Shribak[33] developed a device called the "polychromatic polarized light microscope." This device uses a bandpass filter and two special filters called "spectral polarization state generators" to enhance the visibility of small birefringent materials in microscopy.[33] These filters contain an ordinary linear polarizer and a z-cut quartz crystal that are oriented orthogonally to each other, similar to CPLM. This approach allows researchers to study weakly birefringent biologic materials, including cells, organelles, and organic matrixes. In a small study, they also demonstrated that this can be used to identify MSU and CPP. They report a perfect match in crystal classification between CPLM and their technique in 55 joint fluid aspirates (28 gout, 20 CPDD, 1 mixed crystal disease, and 6 negative),[34] which indicates at least similar sensitivity and specificity as CPLM, although increased sensitivity is to be expected according to the author. The operation of the polychromatic polarized light microscope should be very similar to CPLM. Although some clinical data are published, additional validation is a necessity. The pricing should be similar to CPLM.

Raman spectroscopic tools for synovial fluid analysis

Raman spectroscopy is a spectroscopic technique based on the interaction between light and electrons in a material (Supplementary Figure 3). The atoms in a molecular bond make oscillatory motions, which modulate the electronic response. The modulation takes place with unique frequencies, which enables objective material identification. When irradiated with monochromatic (laser) light, the energy associated with the modulation of the bonds will be small but leads to a presentation of higher and lower frequencies in the scattered light. This can be measured as a change in color with respect to the original monochromatic light source. Analysis of the color spectrum of the scattered light provides a fingerprint of the bonds present in the sample and can be used to identify the material.

Although Raman spectroscopy is a widely used analytical technique for research, clinical use of Raman spectroscopy is limited. The ability to detect MSU and CPP crystals in synovial fluid was already demonstrated in 1991.[35] We discuss several applications here.

A proof of principle of Raman microscopy for crystal identification. McGill et al used a BGSC RM III Raman microscope with a 514-nm, 10-mW laser to identify MSU crystals in a patient with gout and CPP crystals in a patient with CPDD.[35] Acquisition of spectra with their microscope, which was a state-of-the-art device for that time, ranged from 1 to 30 minutes. They show a strong correlation between patient material and spectra from synthetic crystals.

A point-of-care Raman spectroscope. The first publication on the possibility of a point-of-care Raman device for gout diagnosis was in 2009 by Cheng et al.[36] They used centrifugation to collect crystal pellets in a conical tube, which was then scanned with a Labram HR800 Raman microscope, using a 660-nm laser source. The complete Raman analysis was automated, and a series of fixed points on the conical tube was measured, which took 15 minutes. They found consistent results with CPLM in 32 of 35 patient samples.

This was followed by Li et al, who produced a shoebox-sized Raman device.[37,38] It can filter the sample from an aspirated joint with an integrated microsieve with 1-micron pores, which is subsequently scanned in the device. Before scanning, a 30-minute chemical digestion step was performed. Thirty fixed scanning points, with a 30-μm step size, were analyzed per sample, and

crystals were identified based on their expression of either a 630 cm^{-1} (MSU) or a 1,050 cm^{-1} peak (CPP).[39] They analyzed 174 samples using both CPLM and their device. There was a consistency between the two in 89.7% of all results. In eight samples, CPLM could identify MSU crystals where Raman could not, and in eight samples, Raman could identify CPP where CPLM could not. In two samples, Raman identified the presence of both CPP and MSU where CPLM did not.

The produced point-of-care device was inexpensive and showed a good correlation with CPLM (Cohen's $\kappa = 0.84$ for gout and $\kappa = 0.61$ for CPDD). Their technique does not allow visual examination of the crystals however, and they recommend the use alongside CPLM rather than being a substitution. As the method performs area scans at fixed locations instead of locating individual crystals, it is to be expected that sensitivity is low in samples with low crystal counts. As a Raman spectrum provides a full chemical fingerprint of the analyte, specificity should be excellent. Although the Raman measurement is automated, there is a digestion step required. Given all available evidence, this method is relatively close to clinical implementation, although no commercial activity can be found in more recent years. Pricing has been stated to be around 10,000 United States dollars (USD).[37]

Stimulated Raman scattering microscopy. Zhang et al used stimulated Raman scattering (SRS) for the detection of CPP and MSU in rat models, synovial fluid aspirates, and surgical tissue.[40] In SRS, the presence of a specific vibrational band is measured with two laser beams, which are shifted in frequency from each other by the frequency of interest in the material.[41] The so-called pump laser will donate intensity to the Raman excitation laser when the band of interest is present. This increase in laser intensity can be measured and therefore be used for identification. In this experiment, a 1,040-nm Raman excitation laser was used, together with a pump laser which was set to 976 nm for MSU and 937 nm for CPP. They were able to detect MSU in synovial fluid samples from two patients with gout and ex vivo surgical resections (metatarsophalangeal 1 [MTP1], ankle, and elbow) of tophi or synovium from four patients with gout. Furthermore, they showed an ability to detect and identify MSU and CPP in synovial tissue of rats (ex vivo), which were injected with these crystals three days before sacrifice.

The spectroscopic contrast offered by SRS increases specificity for crystals that are difficult to detect with CPLM, such as CPP. SRS microscopy allows for spectroscopy-based, visual observation of the crystals. However, the method of Zhang et al only allows analysis of a portion of the Raman spectrum, inhibiting the identification of crystals other than MSU and CPP. The specificity is therefore lower compared to other spectroscopic techniques. SRS is challenging to do and is probably not suitable for standard clinical applications, due to the complexity of the associated technology, and SRS devices are costly (<400,000 USD).

A hybrid Raman-polarized light microscopy platform. One of the shortcomings of Raman spectroscopic analysis for synovial fluid samples is the inability to localize birefringent objects, such as most crystals in the sample. Crystals are small (<10 μm) objects and can be difficult to find, especially when low in number. The use of polarized light microscopy provides crystal localization with high contrast but low chemical specificity. By combining polarization microscopy with Raman microscopy, a hybrid platform is created: the integrated Raman-polarized light microscope (iRPolM).

Diagnostic accuracy studies have been performed with this device for gout and CPDD. MSU crystals were identified in 77.6% of patients with gout who were positive for the 2015 ACR/EULAR gout classification criteria,[42,43] and CPP crystals were identified in 86.0% of patients with CPDD who were positive for the 2023 ACR/EULAR CPDD classification criteria.[44,45] The primary advantage of this modality is the ability to find the crystals and to obtain a Raman spectrum to characterize the crystal with high chemical specificity. It has shown itself as a proven technology not only for the identification of MSU or CPP crystals but also has been shown to be important in the study of osteoarthritis, as it demonstrated the ability to characterize seven crystal types in end-stage osteoarthritis[46] and was able to identify foreign body material, both in patients with suspected orthopedic implant–related complications[47] and in native joints.[48] iRPolM demonstrates increased sensitivity compared to CPLM, given the spectroscopic contrast, which makes identification of CPP crystals easier. As Raman spectra are chemical fingerprints of the studied crystals, specificity is excellent. Identification of crystals with iRPolM requires operators to locate crystals with polarized light microscopy, which can then be scanned with the integrated Raman spectroscope. Several studies on diagnostic performance have been published, and further development toward clinical implementation is being pursued. At this point, no price is available, but it is expected to be around 150,000 to 250,000 USD.

Automated classification of Raman spectra. Raman spectra are highly suitable for machine-learning approaches due to their unique, specific characteristics. Popular chemometric techniques for the classification of Raman spectra include linear classifiers, such as linear discriminant analysis or logistic regression,[49] but also more sophisticated methods, including one-dimensional convolutional neural networks.[50] Raw Raman spectroscopic data often has to be preprocessed before classification is possible. Common preprocessing steps include baseline removal, normalization, and dimensional reduction. A support vector machine-based model demonstrated to be highly effective in classifying spectra from patients with gout and CPDD. This model was trained on data from 246 patients and validated on 200 new, unseen synovial fluids. The model had a sensitivity of 81.4% for gout and 90.5% for CPDD, compared to the respective ACR/EULAR classification criteria sets.[51]

Research tools: infrared spectroscopy and scanning electron microscopy/energy dispersive x-ray spectroscopy

Infrared (IR) spectroscopy and scanning electron microscopy (SEM), the latter often combined with energy dispersive x-ray spectroscopy (EDX), have been used extensively as a research tool in crystal analysis. These methods are generally considered unfit for clinical applications, however, due to high associated costs, laborious sample preprocessing, and extensive training required. We therefore briefly discuss these methods in this section.

Fourier transform infrared spectroscopy (FTIR). IR spectroscopy, like Raman spectroscopy, is a form of vibrational spectroscopy and uses IR light to measure the molecular vibrations of an object. In FTIR, the sample is illuminated with multiple wavelengths of IR light at once, and a Fourier transform of variable pathlength information renders the absorption spectrum. FTIR often requires dehydration of the samples to avoid the intense absorption of IR light by water and is therefore less suitable for synovial fluids, although it has been shown that crystals can be identified in tissue.[52,53] In the study of Rosenthal et al (2008),[54] a synchrotron is used to generate this collection of wavelengths. This is a costly and highly specialized device but leads to a sensitive system. In their study, Rosenthal et al demonstrate their ability to identify CPP and BCP crystals in synovial fluid. Like Raman spectroscopy, IR spectra are chemical fingerprints of crystals and specificity is excellent.[55]

SEM/EDX. The resolution offered by SEM is much higher than any available by optical microscopic technique, due to the much shorter wavelengths of the electrons compared to photons. SEM performs a raster scan with an electron beam and provides image contrast due to the scattering of these electrons by the sample. These devices can be complemented with an integrated EDX, which measures the emission of atom-specific x-ray spectra stimulated by the electron beam. Due to the high resolution of SEM, it is well-suited for the study of tiny crystals such as BCP.[56] The sensitivity of SEM is very high, and with EDX, it is also specific in crystal identification. Sample preprocessing before SEM analysis is extensive, however, and includes sample dehydration. These devices are complicated in use and therefore not suitable for normal clinical practice. Pricing is around 400,000 USD.

Noninvasive approaches for diagnosing gout

As a joint puncture is a painful procedure for patients and the puncture site is a potential entry point for infections, noninvasive approaches would be beneficial for rheumatology care. Crystals can be seen with imaging techniques such as x-ray, echography, and, more recently, dual-energy computed tomography, but these techniques have been thoroughly discussed elsewhere.[57] The main advantage of dual-energy computed tomography is increased chemical specificity for specific types of crystals (BCP, CPP, and MSU[58]), although the sensitivity is limited. Here we will focus on two approaches using visible light to diagnose gout in vivo.

In vivo Raman spectroscope. Abhishek et al used a Snowy Range Sierra Raman spectrometer device as a noninvasive Raman probe.[59] This device operates a 785-nm, 100-mW laser and has a spectral resolution of 4 cm^{-1}. They placed their device directly on the skin of the affected first MTP joint. The in vivo measurement principle of Abhishek et al eliminates the need for synovial fluid aspiration, as it can directly scan the affected joint. The operator places the device about 2 cm from the affected joint, oriented such that the laser is in line with the synovial space. In their pilot-sized study, MSU crystals were found in 7 of 10 patients with gout and 1 of 10 patients with osteoarthritis. They did not include a comparison with CPLM.

In vivo Raman spectroscopy has been used as a diagnostic tool before in other fields of interest with promising results.[60–62] While being noninvasive, in principle, the penetration of light into the skin is limited. Also, measuring through the skin will generate a significant amount of background signal, which can be problematic. In the Abhishek et al study, 16 of 18 known MSU peaks were identified. However, 12 of these peaks overlapped with peaks present in the signal from a healthy patient, and only 4 peaks were unique. In their paper, a background removal algorithm was recommended to tackle some of these problems, but none was applied. There has been no update on this work since the first paper. The sensitivity of this technique is low, as it can only identify MSU in patients with sufficient deposits in their joint. Additionally, the presence of strong background signals make the identification of Raman spectra more difficult. The used device is easy to operate, although good alignment with the joint of interest is required. Although some interesting results were published, no activity toward clinical implementation has been pursued. No pricing is available, but the device should be relatively low cost.

Noninvasive Raman spectroscopic urate monitoring in sweat. The level of uric acid in human sweat is directly related to the level of uric acid in the serum.[63] As sweat is easily accessible without penetrating the skin, it provides a window for in vivo metabolite monitoring. Most techniques for analyzing sweat use electrochemistry, relying on enzymes and antibodies to retrieve valuable information. In their 2024 paper, Chen et al use surface-enhanced Raman spectroscopy (SERS) to measure uric acid levels in sweat.[64] They designed a specialized bandage-sized wearable to collect sweat, which can be analyzed with their spectroscope. Combined with a deep-learning

artificial intelligence model, they state an accuracy of 97% for diagnosing gout in simulated sweat samples.

As this device is measuring uric acid levels in serum, not monosodium urate crystals in synovial fluid, it is not suitable as a diagnostic for gout. Asymptomatic hyperuricemia is very common, and hyperuricemia in serum is therefore not a specific biomarker for gout. This device can be interesting as a tool in clinical research in gout, however, as frequent, noninvasive serum uric acid level monitoring provides deep insight into uric acid metabolism. No actual clinical data have been published as of yet, and at this point, no pricing is available, although the required SERS device is costly (>150,000 USD).

Physical constraints in transdermal measurements and how to tackle these. Due to the high attenuation of visible light in tissue,[65] we are only able to penetrate a few millimeters deep into the human skin. Attenuation of light in tissue is a combination of absorption and scattering processes and is highly dependent on the wavelength (color). Abhishek et al used a 785-nm laser, which is considered near-IR, and demonstrated promising results with their device in an MTP1 joint.[59] However, it is to be expected that these results cannot be reproduced with larger joints such as an ankle, knee, or shoulder.

There is much research on novel photonic principles to penetrate deeper in human tissue. Molecules such as melanin, fat, collagen, and hemoglobin are high absorbing in the visible light region, which leads to a direct physical constraint on how we can perform transdermal measurements. If near-IR or IR light is used, we see significantly lower attenuation, which could provide a window of opportunity.

Conclusions

In this narrative literature review, we discuss a wide range of innovations designed to improve the diagnosis of gout, CPDD, other crystallopathies, and crystal mimickers. These innovations are designed to either enhance the currently available CPLM technique, which leads to improved sensitivity, or change the crystal characterization method to a spectroscopic approach, which leads to improved specificity. Most techniques still require arthrocentesis, although two promising noninvasive in vivo approaches have been published. None of the described innovations have made it into the clinic, and all will require approval from the Food and Drug Administration (USA), In Vitro Diagnostic Regulation (EU), or a comparable agency before clinical acceptation. Clinical performance data are available for some techniques, but none have been tested by independent agencies. As all described devices are in a prototype stage (most technology readiness level 3–4, some 6–7), it will take several years before any of these in vitro diagnostic tools enter the market.

Traditional treatments of crystal arthropathies like gout and CPDD are often based on treating the inflammation rather than the cause of the disease. The scientific consensus is shifting, however, with the appearance of drugs targeting the formation of specific crystals and inflammasome inhibitors. Gaining more insight into the different crystals present in the synovial fluid of these patients will allow physicians to target specific crystals and make therapies more personalized. Crystal identification can then be used to find biomarkers for both diagnosis and disease monitoring.

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 Niessink 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. Dalbeth N, Gosling AL, Gaffo A, et al. Gout. Lancet 2021;397(10287):1843–1855.
2. Latourte A, Ea H-K, Richette P. Calcium pyrophosphate and basic calcium phosphate crystal arthritis: 2023 in review. Gout Urate Cryst Depos Dis 2024;2(2):101–107.
3. Cronstein BN, Terkeltaub R. The inflammatory process of gout and its treatment. Arthritis Res Ther 2006;8(suppl 1):S3.
4. ElTaraboulsi R. Crystalline Arthropathy. Clinical Foundations of Musculoskeletal Medicine. Springer; 2021:143–149.
5. Richette P, Doherty M, Pascual E, et al. 2018 updated European League Against Rheumatism evidence-based recommendations for the diagnosis of gout. Ann Rheum Dis 2020;79(1):31–38.
6. Berendsen D, Neogi T, Taylor WJ, et al. Crystal identification of synovial fluid aspiration by polarized light microscopy. An online test suggesting that our traditional rheumatologic competence needs renewed attention and training. Clin Rheumatol 2017;36(3):641–647.
7. Lumbreras B, Pascual E, Frasquet J, et al. Analysis for crystals in synovial fluid: training of the analysts results in high consistency. Ann Rheum Dis 2005;64(4):612–615.
8. Bernal JA, Andrés M, López-Salguero S, et al. Agreement among multiple observers on crystal identification by synovial fluid microscopy. Arthritis Care Res (Hoboken) 2023;75(3):682–688.
9. Buchholz A, Stücker S, Koßlowski F, et al. High-resolution imaging methods for identification of calcium crystal types in osteoarthritis. Gout Urate Cryst Depos Dis 2023;1(2):62–82.
10. Pascual E, Sivera F, Andres M. Mixed crystal disease: a tale of 2 crystals. J Rheumatol 2020;47(8):1158–1159.
11. Zhang W, Doherty M, Bardin T, et al. European League Against Rheumatism recommendations for calcium pyrophosphate deposition. Part I: terminology and diagnosis. Ann Rheum Dis 2011;70(4):563–570.
12. Kohn NN, Hughes RE, McCarty DJ Jr, et al. The significance of calcium phosphate crystals in the synovial fluid of arthritic patients: the "pseudogout syndrome". II. Identification of crystals. Ann Intern Med 1962;56(5_Part_1):738–745.

13. Phelps P, Steele AD, McCarty DJ Jr. Compensated polarized light microscopy. Identification of crystals in synovial fluids from gout and pseudogout. JAMA 1968;203(7):508–512.
14. Niessink T, Jansen TL, Kleinherenbrink M, et al. Synovial fluid analysis with compensated polarization light microscopy: a physics approach to quantitative understanding of birefringence, polarization, and Maltese crosses. Clin Rheumatol 2024;43(12):3597–3602.
15. Schumacher HR Jr, Sieck MS, Rothfuss S, et al. Reproducibility of synovial fluid analyses. A study among four laboratories. Arthritis Rheum 1986;29(6):770–774.
16. von Essen R, Hölttä AMH, Pikkarainen R. Quality control of synovial fluid crystal identification. Ann Rheum Dis 1998;57(2):107–109.
17. McGill NW, York HF. Reproducibility of synovial fluid examination for crystals. Aust N Z J Med 1991;21(5):710–713.
18. Gordon C, Swan A, Dieppe P. Detection of crystals in synovial fluids by light microscopy: sensitivity and reliability. Ann Rheum Dis 1989; 48(9):737–742.
19. Fam AG. Cholesterol crystals in gouty bursitis. J Rheumatol 2001; 28(8):1928–1929.
20. Hongsmatip P, Cheng KY, Kim C, et al. Calcium hydroxyapatite deposition disease: imaging features and presentations mimicking other pathologies. Eur J Radiol 2019;120:108653.
21. Yavorskyy A, Hernandez-Santana A, McCarthy G, et al. Detection of calcium phosphate crystals in the joint fluid of patients with osteoarthritis - analytical approaches and challenges. Analyst 2008;133(3): 302–318.
22. Shoji K. Alizarin red S staining of calcium compound crystals in synovial fluid. Article in Japanese. Nihon Seikeigeka Gakkai Zasshi 1993; 67(4):201–210.
23. Derfus BA, Kurian JB, Butler JJ, et al. The high prevalence of pathologic calcium crystals in pre-operative knees. J Rheumatol 2002; 29(3):570–574.
24. Wang YH, Ho S. A rare case of gout and pseudogout occurring in the same joint. Malays Orthop J 2019;13(1):54–56.
25. Qureshi AM, Tariq S, Javed N, et al. One joint aspirate: three diagnoses. Cureus 2021;13(9):e17714.
26. Yoo Y, Seo Y-J, Huh M, et al. Gout and coexisting pseudogout in the knee joint. Knee Surg Sports Traumatol Arthrosc 2011;19(4): 553–555.
27. FitzGerald JD. Novel Techniques for Synovial Fluid Crystal Analysis. In: Mandell BF, ed. Synovial Fluid Analysis and The Evaluation of Patients With Arthritis. Springer International Publishing; 2022: 133–142.
28. Bernar A, Gebetsberger JV, Bauer M, et al. Optimization of the alizarin red s assay by enhancing mineralization of osteoblasts. Int J Mol Sci 2022;24(1):723.
29. Rosenthal AK. Analytic Methods to Detect Articular Basic Calcium Phosphate Crystals. In: Mandell BF, ed. Synovial Fluid Analysis and The Evaluation of Patients With Arthritis. Springer International Publishing; 2022:125–132.
30. Bai B, Wang H, Liu T, et al. Pathological crystal imaging with single-shot computational polarized light microscopy. J Biophotonics 2020; 13(1):e201960036.
31. FitzGerald JD, Barrios C, Liu T, et al. A novel polarized light microscope for the examination of birefringent crystals in synovial fluid. Gout Urate Cryst Depos Dis 2024;2(4):315–324.
32. Zhang Y, Lee SY, Zhang Y, et al. Wide-field imaging of birefringent synovial fluid crystals using lens-free polarized microscopy for gout diagnosis. Sci Rep 2016;6(1):28793.
33. Shribak M. Polychromatic polarization microscope: bringing colors to a colorless world. Sci Rep 2015;5(1):17340.
34. Jum'ah H, Shribak M, Keikhosravi A, et al. Detection of crystals in joint fluid aspirates with polychromatic polarisation microscopy. Ann Rheum Dis 2023;82(11):1501–1502.
35. McGill N, Dieppe PA, Bowden M, et al. Identification of pathological mineral deposits by Raman microscopy. Lancet 1991;337(8733): 77–78.
36. Cheng X, Haggins DG, York RH, et al. Analysis of crystals leading to joint arthropathies by Raman spectroscopy: comparison with compensated polarized imaging. Appl Spectrosc 2009;63(4):381–386.
37. Li B, Singer NG, Yeni YN, et al. A point-of-care Raman spectroscopy-based device for the diagnosis of gout and pseudogout: comparison with the clinical standard microscopy. Arthritis Rheumatol 2016; 68(7):1751–1757.
38. Li B, Yang S, Akkus O. An automated Raman device for gout diagnosis. Abstract presented at: CLEO 2013; June 9–14, 2013; San Jose, CA.
39. Akkus O, Slipchenko M, Akkus A, et al. inventors; US20180363024A1 Methods and devices for diagnosis of particles in biological fluids. US. 2014.
40. Zhang B, Xu H, Chen J, et al. Highly specific and label-free histological identification of microcrystals in fresh human gout tissues with stimulated Raman scattering. Theranostics 2021;11(7):3074–3088.
41. Agrawal GP. Stimulated Raman scattering. In: Nonlinear Fiber Optics. 6th ed. Academic Press; 2019:297–354.
42. Niessink T, Giesen T, Efdé M, et al. Test characteristics of Raman spectroscopy integrated with polarized light microscopy for the diagnosis of acute gouty arthritis. Joint Bone Spine 2023;90(6):105611.
43. Neogi T, Jansen TLTA, Dalbeth N, et al. 2015 gout classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheumatol 2015;67(10):2557–2568.
44. Niessink T, Janssen M, Giesen T, et al. Diagnostic accuracy of Raman spectroscopy integrated with polarized light microscopy for calcium pyrophosphate-associated arthritis. Arthritis Care Res (Hoboken) 2024;76(9):1333–1341.
45. Abhishek A, Tedeschi SK, Pascart T, et al. The 2023 ACR/EULAR classification criteria for calcium pyrophosphate deposition disease. Arthritis Rheumatol 2023;75(10):1703–1713.
46. Niessink T, Stassen RHMJ, Kischkel B, et al. Discovery of calcite as a new pro-inflammatory calcium-containing crystal in human osteoarthritic synovial fluid. Osteoarthritis Cartilage 2024;32(10):1261–1272.
47. Niessink T, Schoenmakers JWA, Janssen M, et al. Raman spectroscopy reveals microparticles in synovial fluids of patients with suspected implant-related complications. J Raman Spectroscopy 2024; 56(2):127–134.
48. Niessink T, Janssen M, Jansen TL, et al. The prevalence of titanium dioxide particles in synovial fluid samples drops after European Union ban. Gout Urate Cryst Depos Dis 2024;2(1):45–51.
49. Brereton RG. One-class classifiers. J Chemometr 2011;25(5): 225–246.
50. Sang X, Zhou R, Li Y, et al. One-dimensional deep convolutional neural network for mineral classification from Raman spectroscopy. Neural Process Lett 2022;54(1):677–690.
51. Niessink T, Jansen TL, Coumans FAW, et al. An objective diagnosis of gout and calcium pyrophosphate deposition disease with machine learning of Raman spectra acquired in a point-of-care setting. Rheumatology (Oxford) 2025;64(4):1791–1798.
52. Ryu K, Iriuchishima T, Oshida M, et al. The prevalence of and factors related to calcium pyrophosphate dihydrate crystal deposition in the knee joint. Osteoarthritis Cartilage 2014;22(7):975–979.
53. Nguyen C, Bazin D, Daudon M, et al. Revisiting spatial distribution and biochemical composition of calcium-containing crystals in human osteoarthritic articular cartilage. Arthritis Res Ther 2013;15(5):R103.

54. Rosenthal AK, Mattson E, Gohr CM, et al. Characterization of articular calcium-containing crystals by synchrotron FTIR. Osteoarthritis Cartilage 2008;16(11):1395–1402.
55. Ea H-K, Gauffenic A, Nguyen QD, et al. Calcium pyrophosphate dihydrate crystal deposition in gouty tophi. Arthritis Rheumatol 2021; 73(2):324–329.
56. Frallonardo P, Ramonda R, Peruzzo L, et al. Basic calcium phosphate and pyrophosphate crystals in early and late osteoarthritis: relationship with clinical indices and inflammation. Clin Rheumatol 2018; 37(10):2847–2853.
57. Reijnierse M, Schwabl C, Klauser A. Imaging of crystal disorders: calcium pyrophosphate dihydrate crystal deposition disease, calcium hydroxyapatite crystal deposition disease and gout pathophysiology, imaging, and diagnosis. Radiol Clin North Am 2022;60(4):641–656.
58. Pascart T, Falgayrac G, Norberciak L, et al. Dual-energy computed-tomography-based discrimination between basic calcium phosphate and calcium pyrophosphate crystal deposition *in vivo*. Ther Adv Musculoskelet Dis 2020;12: 1759720x20936060.
59. Abhishek A, Curran DJ, Bilwani F, et al. In vivo detection of monosodium urate crystal deposits by Raman spectroscopy-a pilot study. Rheumatology (Oxford) 2016;55(2):379–380.
60. Baker MJ, Byrne HJ, Chalmers J, et al. Clinical applications of infrared and Raman spectroscopy: state of play and future challenges. Analyst 2018;143(8):1735–1757.
61. Matousek P, Draper ERC, Goodship AE, et al. Noninvasive Raman spectroscopy of human tissue in vivo. Appl Spectrosc 2006;60(7): 758–763.
62. Parlatan U, Inanc MT, Ozgor BY, et al. Raman spectroscopy as a non-invasive diagnostic technique for endometriosis. Sci Rep 2019;9(1): 19795.
63. Huang CT, Chen ML, Huang LL, et al. Uric acid and urea in human sweat. Chin J Physiol 2002;45(3):109–115.
64. Chen Z, Wang W, Tian H, et al. Wearable intelligent sweat platform for SERS-AI diagnosis of gout. Lab Chip 2024;24(7):1996–2004.
65. Tao R, Gröhl J, Hacker L, et al. Tutorial on methods for estimation of optical absorption and scattering properties of tissue. J Biomed Opt 2024;29(6):060801.
66. Zamudio-Cuevas Y, Fernández-Torres J, Martínez-Nava GA, et al. Phagocytosis of monosodium urate crystals by human synoviocytes induces inflammation. Exp Biol Med (Maywood) 2019;244(5): 344–351.
67. Jin S, Geng H, Lu Y, et al. Association between gout and kidney stone: results from mendelian randomization and the NHANES study. Front Genet 2024;15:1417663.
68. Doumas C, Vazirani RM, Clifford PD, et al. Acute calcific periarthritis of the hand and wrist: a series and review of the literature. Emerg Radiol 2007;14(4):199–203.
69. Rosenthal AK. Crystals, inflammation, and osteoarthritis. Curr Opin Rheumatol 2011;23(2):170–173.
70. Dey B, Khonglah Y, Raphael V, et al. Hypercholesterolemic arthritis: a case report with emphasis on synovial fluid cytology. J Family Med Prim Care 2020;9(8):4428–4430.
71. Ettlinger RE, Hunder GG. Synovial effusions containing cholesterol crystals report of 12 patients and review. Mayo Clin Proc 1979;54(6): 366–374.
72. Lazarevic MB, Skosey JL, Vitic J, et al. Cholesterol crystals in synovial and bursal fluid. Semin Arthritis Rheum 1993;23(2):99–103.
73. Gur H, Ehrenfeld M. The benign long-term effect of cholesterol crystal synovial cysts. Clin Rheumatol 1994;13(3):537–542.
74. Young P, Homlar KC. Extreme postinjection flare in response to intra-articular triamcinolone acetonide (Kenalog). Am J Orthop 2016;45(3): E108–E111.
75. Courtney P, Doherty M. Joint aspiration and injection and synovial fluid analysis. Best Pract Res Clin Rheumatol 2013;27(2):137–169.
76. Stratton-Powell AA, Williams S, Tipper JL, et al. Mixed material wear particle isolation from periprosthetic tissue surrounding total joint replacements. J Biomed Mater Res B Appl Biomater 2022;110(10): 2276–2289.
77. Xie Y, Peng Y, Fu G, et al. Nano wear particles and the periprosthetic microenvironment in aseptic loosening induced osteolysis following joint arthroplasty. Front Cell Infect Microbiol 2023;13:1275086.
78. Billi F, Benya P, Kavanaugh A, et al. The John Charnley Award: an accurate and extremely sensitive method to separate, display, and characterize wear debris: part 2: metal and ceramic particles. Clin Orthop Relat Res 2012;470(2):339–350.
79. Yao JJ, Lewallen EA, Trousdale WH, et al. Local cellular responses to titanium dioxide from orthopedic implants. Biores Open Access 2017; 6(1):94–103.
80. Li Z, Zheng Y, Maimaiti Z, et al. Identification and analysis of microplastics in human lower limb joints. J Hazard Mater 2024;461: 132640.
81. Marfella R, Prattichizzo F, Sardu C, et al. Microplastics and nanoplastics in atheromas and cardiovascular events. N Engl J Med 2024; 390(10):900–910.
82. Leslie HA, van Velzen MJM, Brandsma SH, et al. Discovery and quantification of plastic particle pollution in human blood. Environ Int 2022; 163:107199.
83. van der Weele L, Niessink T, Maas M, et al. Clinical image: synovitis of the metacarpophalangeal joints. Ann Rheum Dis 2023;82(7): 992–993.

Arthritis Care & Research
Vol. 78, No. 6, June 2026, pp 798–809
DOI 10.1002/acr.25697

AMERICAN COLLEGE of RHEUMATOLOGY
Empowering Rheumatology Professionals

REVIEW ARTICLE

Performance Characteristics of Anti–Collagen II Antibodies in Relapsing Polychondritis and Related Diseases: Prospective Analysis, Systematic Review, and Meta-Analysis

Karyssa Stonick,[1] Marcela A. Ferrada,[2] Alice Fike,[1] Kaitlin A. Quinn,[1] Benjamin A. Turturice,[1] Casey Stein,[1] and Peter C. Grayson[1]

Objective. Relapsing polychondritis (RP) is a rare disease defined by recurrent cartilaginous inflammation. Anti–collagen II (Col2) antibodies have been proposed as a diagnostic biomarker for RP, but their performance characteristics are not well defined.

Methods. In an observational cohort, anti-Col2 antibody levels were measured in patients with RP compared to other inflammatory diseases that mimic RP. In parallel, a systematic review and meta-analysis was performed to assess anti-Col2 antibody prevalence across a broad spectrum of diseases. Individual study risk ratios (RRs) and pooled disease category RRs, study bias, and interstudy heterogeneity were assessed.

Results. In the observational cohort, anti-Col2 antibody prevalence did not differ between RP and comparators. The performance characteristics of anti-Col2 to diagnose RP were poor (sensitivity = 18%; specificity = 72%). Anti-Col2 antibody titers did not correlate with disease activity in RP (r = 0.08, P = 0.44). In the systematic review, 71 of 2,443 reviewed articles were included. Anti-Col2 antibodies were not associated with RP across five pooled studies (RR: 2.09; 95% confidence interval [CI]: 0.05–81.80; P = 0.69). Anti-Col2 antibodies were significantly associated with a composite group of inflammatory diseases with cartilaginous involvement (RR: 2.99; 95% CI: 1.29–6.91; P = 0.01). Studies using healthy controls reported increased effect sizes compared to studies that used disease controls (β-estimate = 1.14, I^2 = 17.08%; P = 0.0004).

Conclusion. Anti-Col2 antibodies are neither sensitive nor specific for RP, are detected in the minority of patients with RP, and are detected at similar prevalences across a spectrum of inflammatory diseases with cartilage inflammation. Use of these antibodies to diagnose or monitor RP is not advisable.

INTRODUCTION

Relapsing polychondritis (RP) is a rare, systemic inflammatory disease characterized by recurrent episodes of cartilage inflammation that can lead to organ and life-threatening complications. RP is primarily diagnosed based on clinical recognition of an appropriate pattern of symptoms and organ system involvement. The clinical criteria to diagnose RP include auricular, nasal, and airway/laryngotracheal chondritis, seronegative nonerosive arthritis, ocular inflammation, hearing loss, vestibular damage, response to glucocorticoids or dapsone, and histologic confirmation.[1–3] Given the rarity and heterogeneous nature of RP, developing adequate expertise to diagnose the disease is challenging, which can lead to diagnostic delays.[1–3] Identification of biomarkers that can accurately and objectively confirm disease when there is clinical suspicion remains a major unmet need in RP.

Anti–collagen II (Col2) antibodies have been associated with RP, and testing for these antibodies is clinically available.[4–7] A few small studies have reported an association between anti-

The findings and conclusions presented in this paper are those of the authors and do not necessarily reflect the views of the NIH or the US Department of Health and Human Services. The contributions of the NIH authors are considered Works of the US Government.

Supported by the Intramural Research Program of the National Institute of Arthritis and Musculoskeletal and Skin Diseases, NIH (ZIA AR041204).

[1]National Institutes of Health, Bethesda, Maryland; [2]University of Maryland, Baltimore.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25697).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25697.

Address correspondence via email to Peter C. Grayson, MD, MSc, at peter.grayson@nih.gov.

Submitted for publication July 2, 2025; accepted in revised form October 3, 2025.

SIGNIFICANCE & INNOVATIONS

- The performance characteristics of anti–collagen II autoantibodies to diagnose or monitor relapsing polychondritis (RP) are poor.
- Presence of anti–collagen II antibodies likely indicates a history of cartilage inflammation rather a disease-specific pathologic antibody.
- Anti–collagen II antibodies have limited utility as a clinical biomarker and should not be routinely ordered in cases of suspected RP.

Col2 antibody titers and RP.[4–6] Animal models of disease support a potentially pathogenic role of anti-Col2 antibodies in RP and related diseases. Immunization of mice with anti-Col2 antibodies induces collagen-induced arthritis and ear chondritis, which recapitulates clinical features of both rheumatoid arthritis (RA) and RP.[8,9] Anti-Col2 antibodies have been proposed as a potential diagnostic biomarker for RP. Whether these antibodies are reliable clinical biomarker to diagnosis and monitor RP has not been well characterized, and it remains unclear if anti-Col2 antibodies are directly involved in disease pathogenesis in humans or simply represent a marker of cartilage disease.

The objective of this study was to determine the performance characteristics of anti-Col2 antibodies to differentiate RP from other diseases. Performance of anti-Col2 antibodies was assessed within an observational cohort study of patients with RP and comparator diseases that can present with similar cartilaginous and extracartilaginous features as RP. A systematic review and meta-analysis was also performed to synthesize existing evidence on the prevalence and diagnostic utility of anti-Col2 antibodies in RP and other conditions.

METHODS

Cohort study. *Study population.* Patients with RP and inflammatory diseases with clinical overlap were enrolled into an ongoing, prospective observational cohort study at the National Institute of Arthritis, Musculoskeletal, and Skin Diseases of the National Institutes of Health. Each patient provided written informed consent, and the protocol was approved by local ethics review (14-AR-0200). All patients with RP fulfilled diagnostic criteria for this disease.[1–3] A subset of patients with RP who have mouth and genital ulcers with inflammatory chondritis (MAGIC) syndrome were studied independently. A comparator group of patients with other forms of vasculitis was selected from within the same study protocol, including patients with vacuoles, E1 enzyme, X-linked, autoinflammatory, somatic (VEXAS) syndrome, large vessel vasculitis (LVV), or small and medium vessel vasculitis (SMVV). Patients with LVV included those with giant cell arteritis or Takayasu arteritis. Patients with SMVV included those with granulomatosis with polyangiitis (GPA), Behçet disease, microscopic polyangiitis, eosinophilic GPA, isolated cutaneous vasculitis, cryoglobulinemic vasculitis, or IgA-associated vasculitis. The composite disease comparator group was chosen because these conditions have clinical features that can substantially overlap with RP, and thus are a likely population in which a diagnostic test to rule out RP would be clinically indicated.

Clinical assessment. Study investigators clinically evaluated each patient at every study visit per protocol. Clinical features of disease were systematically recorded. Chondritis was defined as auricular chondritis, nasal chondritis, airway chondritis, or inflammatory arthritis. Cartilage damage was defined as subglottic stenosis, tracheomalacia, bronchomalacia, saddle nose deformity, or external ear deformity, as previously defined.[10] Physician Global Assessment (PGA) scores were assessed on a scale of 0: remission to 10: severe disease at each visit. Demographic information including age, sex, race, ethnicity, CD19 B cell counts, and treatments was recorded. Patients were considered B cell depleted if they had an absolute CD19 cell count of <70/uL by flow cytometry.

Anti-Col2 antibody testing. Anti-Col2 antibody testing (Mayo Clinic, test identifier: FFTYC) was performed on the serum of consecutive patients evaluated between 2019 and 2023. If testing from multiple study visits was available for an individual patient, results from the initial test was used. The Mayo Clinic test uses an IgG antibody against the antigen of denatured Col2 derived from chick sternal cartilage. Reference ranges were applied per predefined test standards from the Mayo Clinic: an anti-Col2 antibody titer level of 20 to 25 equivalent units (EU)/mL was considered borderline elevated, and a titer level of >25 EU/mL was considered unequivocally positive. Analyses were performed at both cutoff levels.

Statistical analysis. Performance characteristics of anti-Col2 antibodies (sensitivity, specificity, receiver operating characteristic curves) were assessed in patients with RP and MAGIC syndrome in reference to the comparator disease group (R-Studio pROC).[11] The comparator group was further stratified into diseases of cartilaginous inflammation (VEXAS, GPA) and extracartilaginous symptom overlap (all other comparators). Continuous variables were compared between groups using Kruskal-Wallis tests. Correlations were calculated using Spearman rank coefficient. Categorical variables were compared using Fischer's exact tests or chi-squared test, as appropriate. The frequency of specific clinical symptoms was compared between those who were anti-Col2 antibody positive vs antibody negative for patients with RP alone and for patients with diseases of cartilage inflammation (RP, MAGIC, VEXAS, GPA). Analyses were performed using GraphPad Prism 10, JMP 16, and R-Studio 4.4.0.

Systematic review and meta-analysis. *Literature search.* To evaluate the diagnostic performance of anti-Col2 antibodies across a broader spectrum of diseases, a systematic review and synthesis of evidence was performed in accordance with the recommendations of the Cochrane Collaboration and the Preferred

Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA). The study was registered in PROSPERO (April 1, 2025; Centre for Reviews and Dissemination no. 1023286). A comprehensive PubMed search was performed using the terms ((anti-collagen II) OR (collagen II antibodies)) AND (human)), and snowballing was performed to identify any additional articles to include. One study investigator (KS) performed the article search. Rayyan AI was used to increase the efficiency of the review by highlighting words that the authors designated as keywords for inclusion (i.e., human, collagen II, etc.) and exclusion (i.e., animal, mouse, rat, etc.) and by tracking which articles were marked for inclusion/exclusion.[12] Articles were excluded first by title, then abstract, and finally by full article review. Studies were excluded if they (1) only measured anti-Col2 antibody titers in animal models, (2) did not include the data from the control population and/or only used the control population to set the reference range within their study and did not show the rate positivity of anti-Col2 antibodies within the control population, (3) did not report the rate of positivity or number of participants positive within the study/disease population, (4) did not measure anti-Col2 antibodies in serum (ie, only measured it in synovial fluid), (5) were not printed in English, or (6) were unable to be accessed. Studies were included in the meta-analysis if they were published in English, were accessible, and measured and reported anti-Col2 antibody prevalence rates in the serum from at least one human disease/participant population and the serum from at least one human control population. Positivity definitions from each study were used for that respective study to define the rate of positivity in the disease/participant population and control population. Individual study definitions for anti-Col2 positivity were frequently based on the titer levels of the control population (ie, ±2 or 3 SDs from the average titer level of the control population).

Study data extraction. For each study, one study investigator (KS) extracted and recorded the sample size, proportion of participants who were anti-Col2 antibody positive, proportion of controls who were anti-Col2 antibody positive, whether healthy and/or disease controls were used, year of publication, country of publication, proportion of female participants (if reported), Col2 antigen species (if reported), Col2 antigen anatomic location (if reported), whether native or denatured anti-Col2 antibodies were used (if reported), average age of the participants (if reported), and the quality of the study. The quality of each study was assessed using the Newcastle-Ottawa Quality Assessment Scale for Case Control Studies. For comparability, studies should account for treatment effect for therapies that could impact antibody production. Additionally, composition of the control group (healthy vs appropriate disease comparator conditions) was considered when assessing study quality.

If more than one disease/population was investigated within an article, the disease that was emphasized in the discussion and conclusion was designated as the primary disease for that study. To reduce multiple comparisons, only the primary disease and control populations were included in the meta-analysis. If multiple antibody comparisons were included, the one most similar to the antibody used by the Mayo Clinic test was included (IgG, denatured, chick, sternal cartilage). If multiple control populations were included, the comparison to the disease control population rather than the healthy control population was included.

Study categorization. Studies were grouped into four categories based on classification of the primary disease of interest as predisease, noninflammatory diseases, inflammatory diseases without cartilage involvement, and inflammatory diseases with cartilage involvement (see Supplementary Table 1 for definitions). Two authors (CS, BAT) independently categorized all the articles included in the study according to these definitions. Discrepancies were resolved with review by a third author (KAQ) and discussion.

Statistical analysis. Risk ratios (RRs) and 95% confidence intervals (CIs) were calculated for each study, and pooled RRs were estimated using the restricted maximum likelihood method. Studies were weighted by cohort size relative to the total number of participants in all the studies within that category combined (equation: (no. of participants within the disease cohort of an article) / (sum of the number of participants with the disease cohort of all the studies) = weight of the study). This method was chosen rather than inverse variance to minimize weighting of studies where the total sample size was disproportionally driven by controls rather than cases. Study bias was calculated using Egger's test. Heterogeneity between studies was evaluated using meta-regression analyses. Forest plots, funnels plots, and meta-regression plots were generated to visualize the data. For meta-regression analyses, the following variables were considered: publication year, sample size of the disease cohort, total study size, age and sex of patients in the disease cohort, disease category (predisease, noninflammatory, inflammatory with or without cartilage involvement), country of publication (Americas, Australia and Asia, Europe), antigen species (human, nonhuman, both), antigen location (axial, appendicular), and control population (healthy vs disease). Studies with missing data or that did not fit into the defined subcategories above were excluded. Subcategories with a minimum of two studies were included in the final regression analysis. Variables with a P value < 0.20 and an R^2 > 0% on univariable regression were included in multivariable models (R-Studio metafor, ggplot2).[13,14] R-Studio (meta, metafor) was used to run meta-analysis and meta-regressions.[11,14,15]

RESULTS

Cohort study. *1. Patient characteristics.* A total of 171 patients were included in the present study: 84 with RP, 11 with MAGIC, and 76 with comparator diseases (45 with VEXAS, 13 with LVV, and 18 with SMVV). Baseline demographics including sex, race, and factors relevant to treatment and disease, reported by disease category, are displayed in Table 1. Except for the patients with VEXAS, which consisted entirely of men because

Table 1. Study population characteristics for the observational cohort (n = 177)*

Characteristics	Relapsing polychondritis (n = 84)	MAGIC (n = 11)	VEXAS (n = 45)	LVV (n = 13)	SMVV (n = 18)
Age, median (IQR), y	45.5 (33.0–57.25)	41.0 (26.5–48.5)	67.0 (63.0–72.0)	37.0 (23.0–69.0)	62.5 (39.5–69.25)
Female sex, n (%)	60 (71)	10 (91)	0 (0)	12 (92)	13 (72)
Race, n (%)	74 (88)	10 (91)	42 (93)	11 (85)	15 (83)
CD19 absolute count <70/uL, n (%)	2 (2)	3 (27)	33 (73)	3 (23)	8 (44)
B cell–depleting therapy, n (%)	3 (4)	0 (0)	4 (9)	0 (0)	4 (22)
IVIG, number (%)	7 (8)	1 (9)	0 (0)	0 (0)	1 (6)
Prednisone dose, median (IQR)	3.75 (0.00–12.12)	0.00 (0.00–15.00)	20.00 (10.00–25.00)	0.00 (0.00–4.00)	0.00 (0.00–5.00)
PGA scores, median (IQR)	2.00 (1.00–3.00)	3.00 (2.00–4.00)	2.00 (2.00–3.00)	0.00 (0.00–0.00)	0.00 (0.00–0.75)

* IQR, interquartile range; IVIG, intravenous immunoglobulin; LVV, large vessel vasculitis; MAGIC, mouth and genital ulcers and inflamed cartilage; PGA, physician global assessment; SMVV, small/medium vessel vasculitis; VEXAS, vacuoles, E1 enzyme, X-linked, autoinflammatory, somatic syndrome.

it is an X-linked disease, the cohort was predominantly female (Table 1). A proportion of patients within each disease category were B cell depleted at the time of antibody testing (Table 1)—this was sometimes, but not always, due to B cell–depleting therapy. Median PGA scores significantly differed by disease ($P < 0.0001$; Table 1), which likely reflects lack of effective clinical treatments for RP, MAGIC, and VEXAS relative to other forms of vasculitis.

2. Performance characteristics of anti-Col2 antibodies and levels across diseases. The performance characteristics of anti-Col2 antibodies to diagnose RP were poor at a reference cutoff of both >20 EU/mL (sensitivity = 31%; specificity = 57%) and of >25 EU/mL (sensitivity = 18%; specificity = 72%). The prevalence of anti-Col2 antibodies was greater in the comparator diseases (28%) than RP (18%), with an area under the curve <50% observed for RP relative to comparators (Figure 1A; Supplementary Figure 1). There were no significant differences in the proportion of patients with detectable anti-Col2 antibodies by disease (RP = 15 of 84, 18%; MAGIC = 5 of 11, 46%; LVV = 3 of 13, 23%; SVV = 6 of 18, 33%; VEXAS = 10 of 45, 22%; $P = 0.21$). The performance characteristics of anti-Col2 antibodies for each disease was poor (Figure 1A). No significant differences in median anti-Col2 antibody titer levels were observed between diseases (RP = 17.1 EU/mL; MAGIC = 21.6 EU/mL; VEXAS = 18.2 EU/mL; LVV = 18.3 EU/mL; SMVV = 20.8 EU/mL; $P = 0.20$, Figure 1B).

3. Correlation of anti-Col2 antibody levels and disease activity. Anti-Col2 antibody titers did not correlate with disease activity for any of the diseases (RP $r = 0.08$, $P = 0.44$; MAGIC $r = 0.01$, $P = 0.99$; VEXAS $r = -0.05$, $P = 0.76$; LVV $r = 0.03$, $P = 0.94$; SMVV $r = 0.10$, $P = 0.70$). Results did not change when excluding CD19 B cell–depleted patients (RP $r = 0.08$, $P = 0.46$; MAGIC $r = 0.29$, $P = 0.45$; VEXAS $r = -0.21$, $P = 0.46$; LVV $r = 0.04$, $P = 0.91$; SMVV $r = 0.05$, $P = 0.88$). Anti-Col2 titers did not correlate with PGA in patients with RP, even when excluding those who were B cell depleted (Figure 1C and 1D). Comparing anti-Col2 positivity in patients with RP who were in remission or had low disease activity (defined as a PGA score of 0 or 1) to those with active disease (defined as PGA score ≥2), no statistically significant association was observed (21% vs 15%, $P = 0.73$). No associations were observed between anti-Col2 antibody titer levels and specific comparator diseases (Supplementary Table 2).

4. Anti-Col2 antibody association with specific clinical manifestations of RP. At a reference cutoff of >25 EU/mL, anti-Col2 antibody positivity did not correlate with symptoms of chondritis or cartilage damage in patients with RP (Supplementary Figure 2A). Across all studied clinical associations, anti-Col2 antibodies were only detected in a greater proportion of patients with RP who had ear chondritis (93 vs 87%) or inflammatory arthritis (93 vs 86%), but these results were not statistically significant. In patients with diseases that have known cartilage involvement (RP, VEXAS, MAGIC, and GPA), there was no association between anti-Col2 antibody positivity and symptoms of chondritis or cartilage damage (Supplementary Figure 2B).

Systematic review and meta-analysis of anti-Col2 antibodies in RP and related diseases. *1. Study characteristics.* Out of 2,443 articles screened, 71 met the inclusion criteria. Most exclusions were due to animal only studies, missing control data, or inaccessible text. The PRISMA flow diagram is show in Figure 2.

2. Anti-Col2 antibodies in RP. Five studies, including the present observational cohort, assessed anti-Col2 antibody positivity prevalence in RP compared to controls, and these

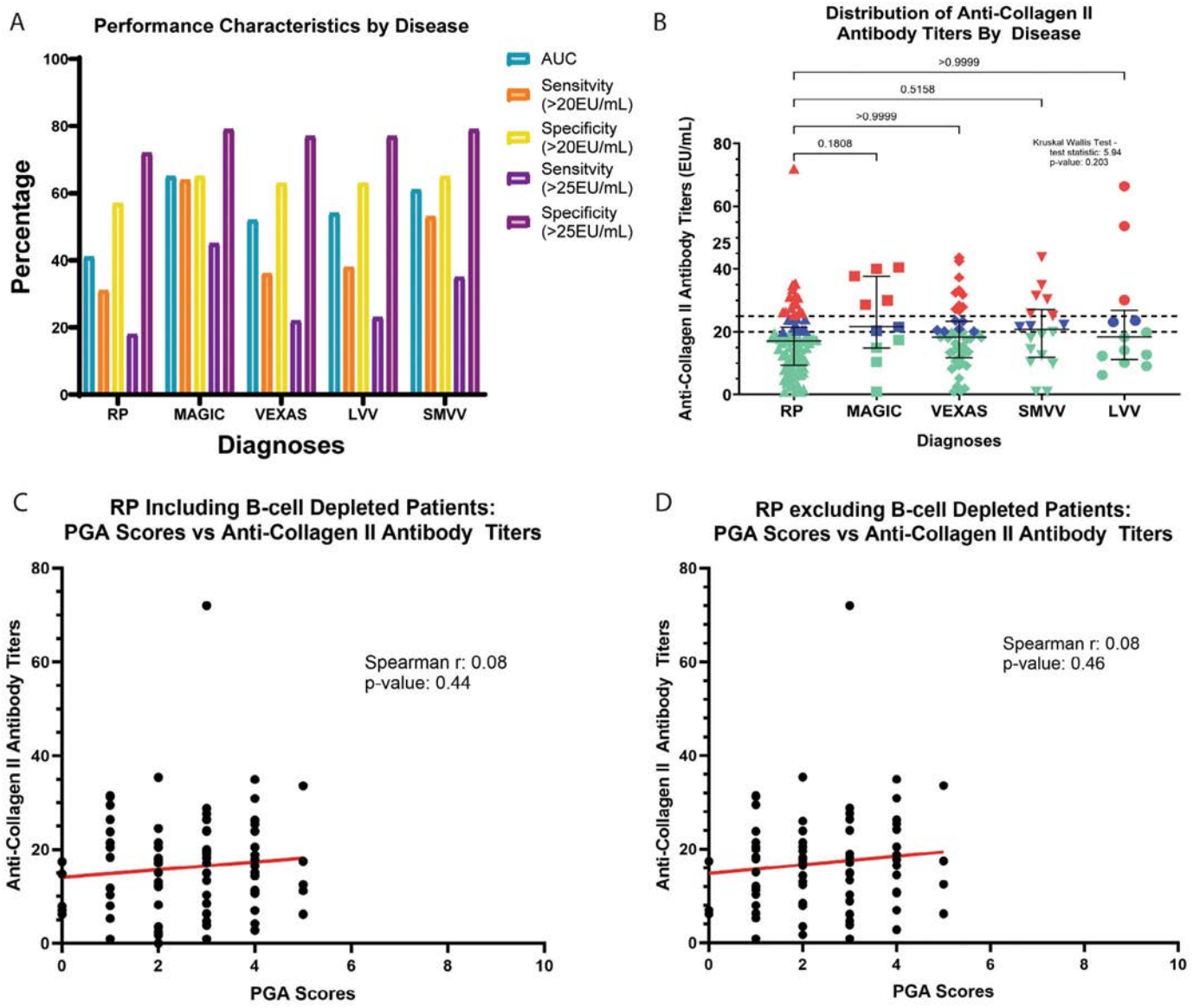


Figure 1. Anti-Col2 antibodies and clinical associations. (A) Prevalence of anti-Col2 antibodies at two different titers did not differ between RP, MAGIC syndrome, VEXAS, LVV, and SMVV. (B) Anti-Col2 antibodies did not differ by disease where green indicates normal, blue is borderline elevated, and red is elevated. (C) Anti-Col2 antibody titers were not associated with PGA scores of disease activity in patients with RP (D) or when excluding patients with RP who were B cell depleted. AUC, area under the curve; Col2, collagen II; EU, equivalent unit; LVV, large vessel vasculitis; MAGIC, mouth and genital ulcers with inflammatory chondritis; PGA, physician global assessment; RP, relapsing polychondritis; SMVV, small/medium vessel vasculitis; VEXAS, E1 enzyme, X-linked, autoinflammatory, somatic.

studies were conducted between 1978 and 2025.[4,5,16,17] The pooled RR for these studies was 2.09 (95% CI: 0.05–81.80; P = 0.69) (Figure 3). The quality of articles ranged from 4 to 7 of 9 with a mean of 5.2 ± 1.3 (95% CI: 3.6–6.8). When studies were excluded based on a low-quality score (≤4 of 9), the pooled RR decreased to 1.66 (95% CI: 0.07–41.30; P = 0.76) (Supplementary Figure 3). There was significant heterogeneity between the studies (I^2 = 92.7%; P < 0.001).

3. Anti-Col2 antibodies in other diseases. There were 44 studies published between 1978 and 2025 comparing the prevalence of anti-Col2 antibodies in inflammatory diseases with cartilaginous involvement to controls, including the present observational cohort study.[4,5,16–57] There was a significant association between anti-Col2 antibodies and inflammatory diseases with cartilaginous involvement compared to controls (RR: 2.99; 95% CI: 1.29–6.91; P = 0.01) (Figure 4A). Because studies were weighted based on disease cohort size, statistical significance was largely driven by three larger studies—Too et al, which was weighted at 13.5%; Manivel et al, which was weighted at 16.2%; and Zeng et al, which was weighted at 23.2% (Supplementary Figure 4).[41,53,55] There was a significant amount of heterogeneity

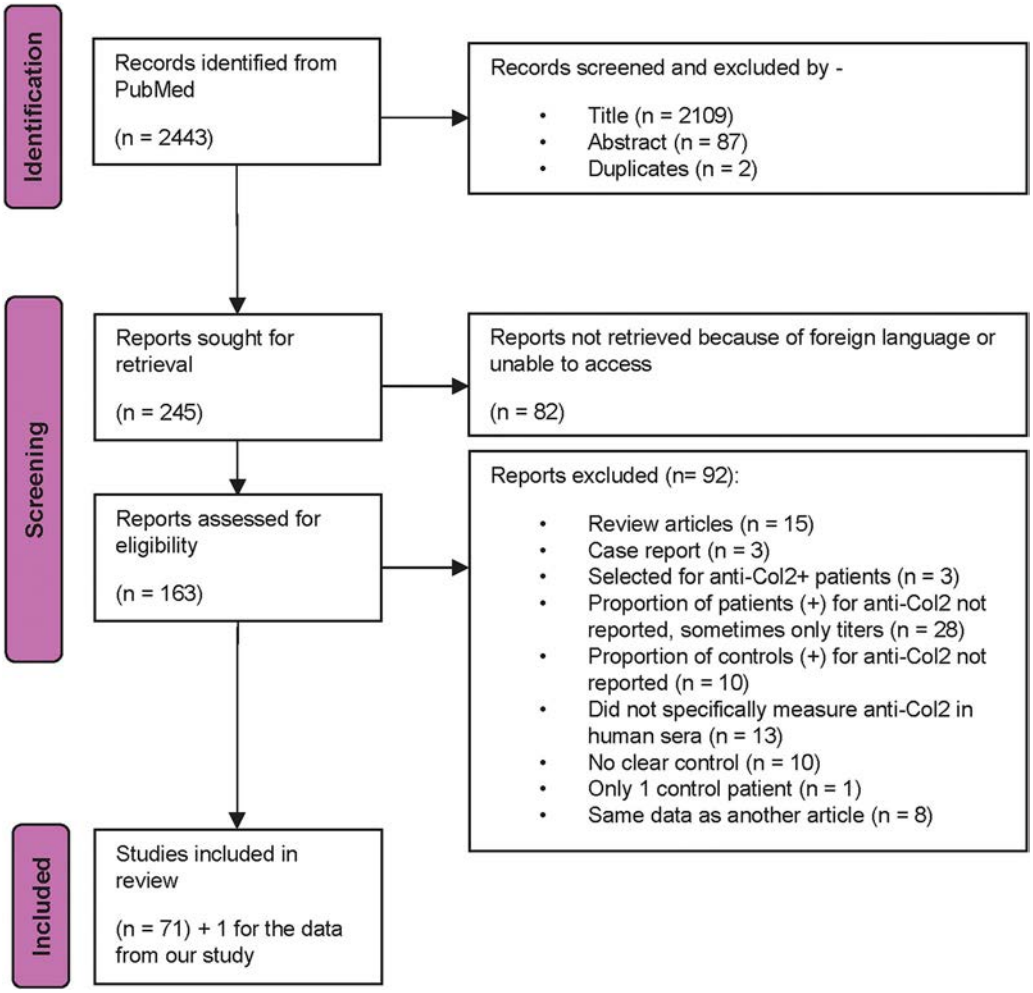


Figure 2. Systematic review study selection flow diagram. This flow diagram describes the process for identifying 71 articles included in this review out of 2,443 potential articles identified by the search strategy. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25697/abstract.

between studies in this group (I^2 = 92.6%; $P < 0.001$). The Egger's test for study bias did not find significant study bias ($P = 0.057$). The line generated from Egger's test on the funnel plot had a negative slope, indicating a nonstatistically significant trend toward a bias in which smaller studies showed larger effect sizes (Supplementary Figure 5A).

There were seven studies published between 1988 and 2016 comparing the prevalence of anti-Col2 antibodies in inflammatory diseases without cartilaginous involvement to controls.[58–63] Risk of anti-Col2 antibodies in inflammatory diseases without cartilaginous involvement was not significantly different compared to controls (RR: 1.74, 95% CI: 0.84–3.61; $P = 0.11$) (Figure 4B, Supplementary Figure 6). There was a significant amount of heterogeneity between the studies within this group (I^2 = 79.5%; $P < 0.001$). The Egger's test for study bias did not find significant study bias ($P = 0.13$) (Supplementary Figure 5B). The line generated from Egger's test on the funnel plot had a positive slope indicating bias toward larger studies showing larger effect sizes.

There were 17 studies published between 1979 and 2018 examining the prevalence of anti-Col2 antibodies in noninflammatory diseases,[64–79] and three studies published between 1988 and 2015 focused on anti-Col2 antibodies in a predisease state before onset of RA.[80–82] Compared to controls, there was no significant association between the noninflammatory group (RR: 1.10, 95% CI: 0.44–2.75; $P = 0.84$) or the predisease group (RR: 1.15, 95% CI: 0.25–5.40; $P = 0.86$) (Figure 4C and 4D; Supplementary Figures 7 and 8). There was a significant amount of heterogeneity between the studies within these groups (I^2 = 82.9%; $P < 0.001$ and I^2 = 94.3%; $P < 0.001$ respectively). The Egger's test for the noninflammatory disease category did not find significant study bias ($P = 0.40$). The line generated from Egger's test on the funnel plot had a negative slope indicating bias toward smaller studies showing larger effect sizes (Supplementary Figure 5C). The Egger's test for the predisease category did not find significant study bias ($P = 0.83$). The line generated from Egger's test on the funnel plot had positive slope indicating bias

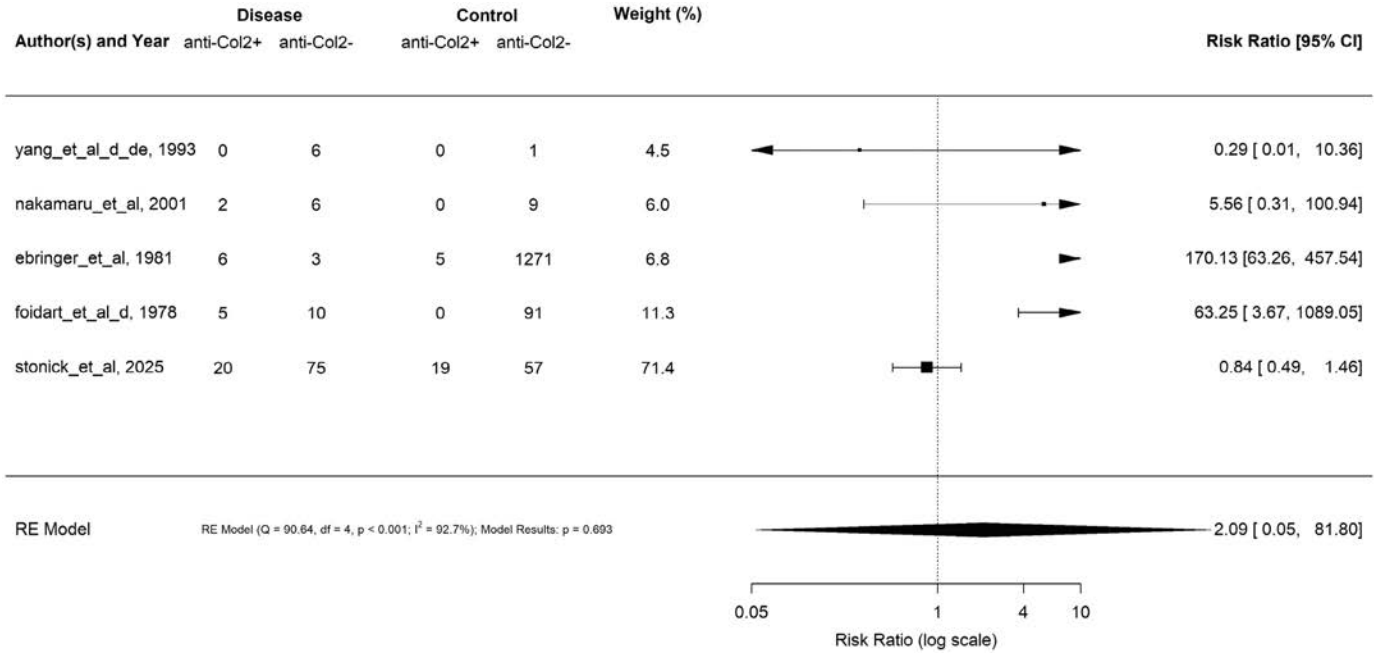


Figure 3. Forest plots of the association of anti-Col2 antibodies and relapsing polychondritis. Anti-Col2 antibodies were not associated with relapsing polychondritis across five pooled studies. The number of disease cases of relapsing polychondritis and controls is reported for each study along with individual and pooled risk ratios. CI, confidence interval; Col2, collagen II; RE, random effects.

toward larger studies showing larger effect sizes (Supplementary Figure 5D).

4. Meta-regressions. On univariable regression, only one variable was significantly associated with interstudy heterogeneity with an R^2 value >0%. Use of healthy controls rather than disease comparators was significantly associated with an increased effect size (β estimate = 1.14, $P = 0.0004$; Figure 5). However, this variable only explained 17.08% of the heterogeneity.

DISCUSSION

Anti-Col2 antibodies have poor performance characteristics to diagnose or monitor disease activity in RP. Within the cohort analysis portion of this study, the presence of anti-Col2 antibodies in patients with RP was not associated with diagnosis, disease activity, or any specific clinical characteristics of disease. Depending on the titer level used to define anti-Col2 antibody positivity, only 18% to 32% of patients with RP had detectable antibodies. There was a numerically greater proportion of patients with RP who had detectable anti-Col2 antibodies in association with ear chondritis and inflammatory arthritis—which are phenotypic features that mice sensitized with anti-Col2 antibodies can develop.[8,9] However, these associations were not statistically significant and were not replicated in a broader set of diseases defined by cartilage involvement. More broadly, the systematic review and meta-analysis portion of the study confirmed that beyond a lack of sensitivity, anti-Col2 antibodies are also not specific for RP and can be detected at similar frequencies across a wide range of related and unrelated conditions. Therefore, anti-Col2 antibodies are not useful in clinical practice to diagnose a patient with suspected RP or to monitor disease activity in a patient with confirmed RP.

Anti-Col2 antibodies have been studied across a wide spectrum of diseases. Although there was significant variability in individual study results within the meta-analysis, a significant pattern of association was observed. Compared to noninflammatory diseases, there was an association between the presence of anti-Col2 antibodies and inflammation, with a three-fold increased risk to detect anti-Col2 antibodies in a group of diseases defined by inflammation with cartilage involvement that was statistically significant and a near significant two-fold increased risk in a group of inflammatory diseases without cartilage involvement. For the few studies in which anti-Col2 antibodies were measured in the predisease state, there was no clinical association. Taken together, these findings suggest that anti-Col2 antibodies may form in response to systemic inflammation that involves cartilage or collagen rich structures but are less likely to play a primary role in disease pathogenesis for a specific condition.

Significant heterogeneity was observed between the studies included in this meta-analysis. Meta-regression did not identify factors that accounted for a substantial proportion of heterogeneity between studies. The only variable that was significantly associated with heterogeneity was the composition of the control group. Studies that used disease controls rather than healthy controls were significantly less likely to identify an association between anti-Col2 antibodies and the disease of interest,

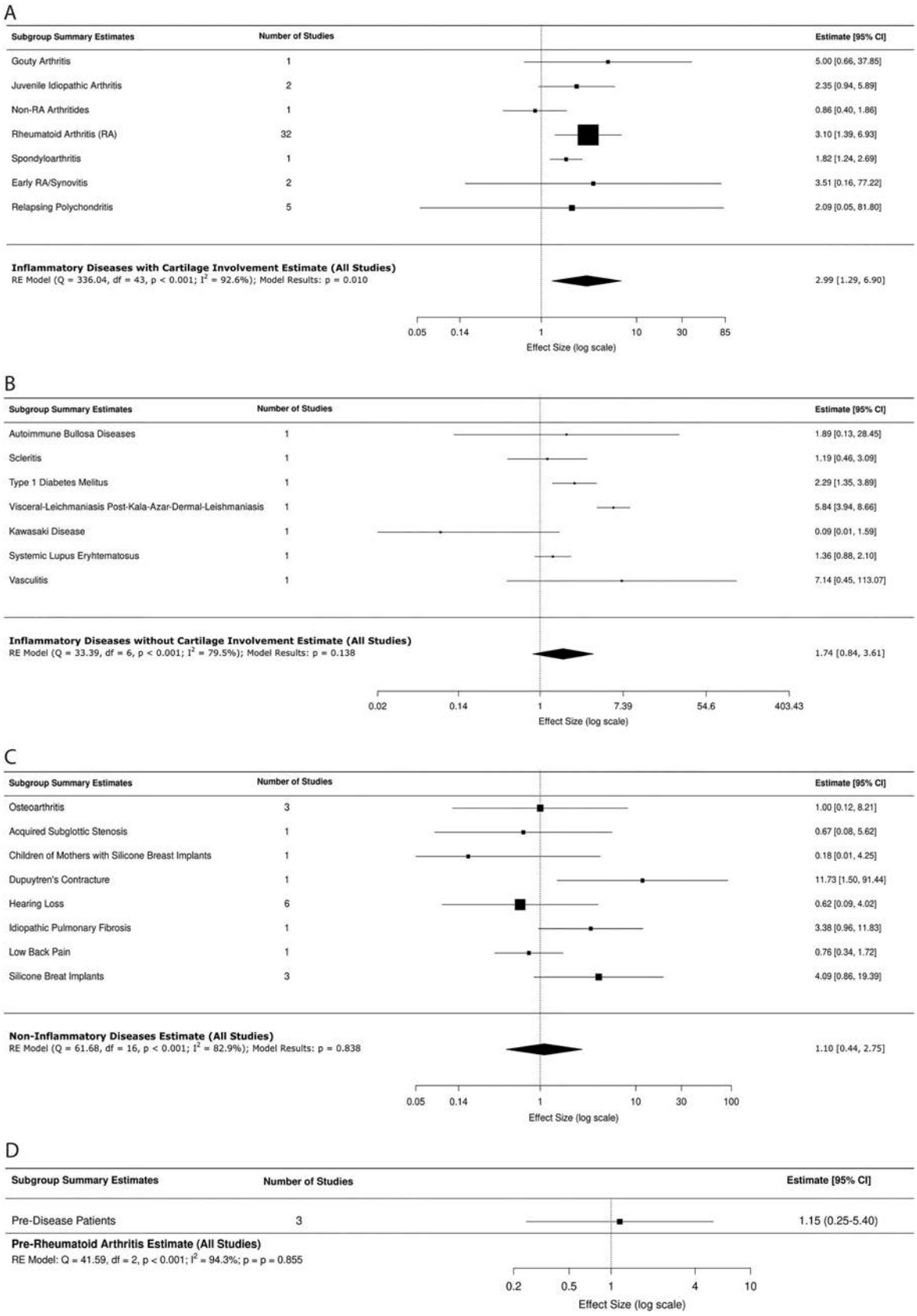


Figure 4. Forest plots of the association of anti-Col2 antibodies and a broad spectrum of clinical diseases. (A) Pooled risk ratios show a significant association between anti-Col2 antibodies and inflammatory diseases with cartilaginous involvement. Significant associations were not detected between (B) inflammatory diseases without cartilaginous involvement, (C) noninflammatory diseases, (D) or in studies in which antibody testing was performed before the onset of rheumatoid arthritis and anti-Col2 anitbody positivity. CI, confidence interval; Col2, collagen II; RA, rheumatoid arthritis; RE, random effects.

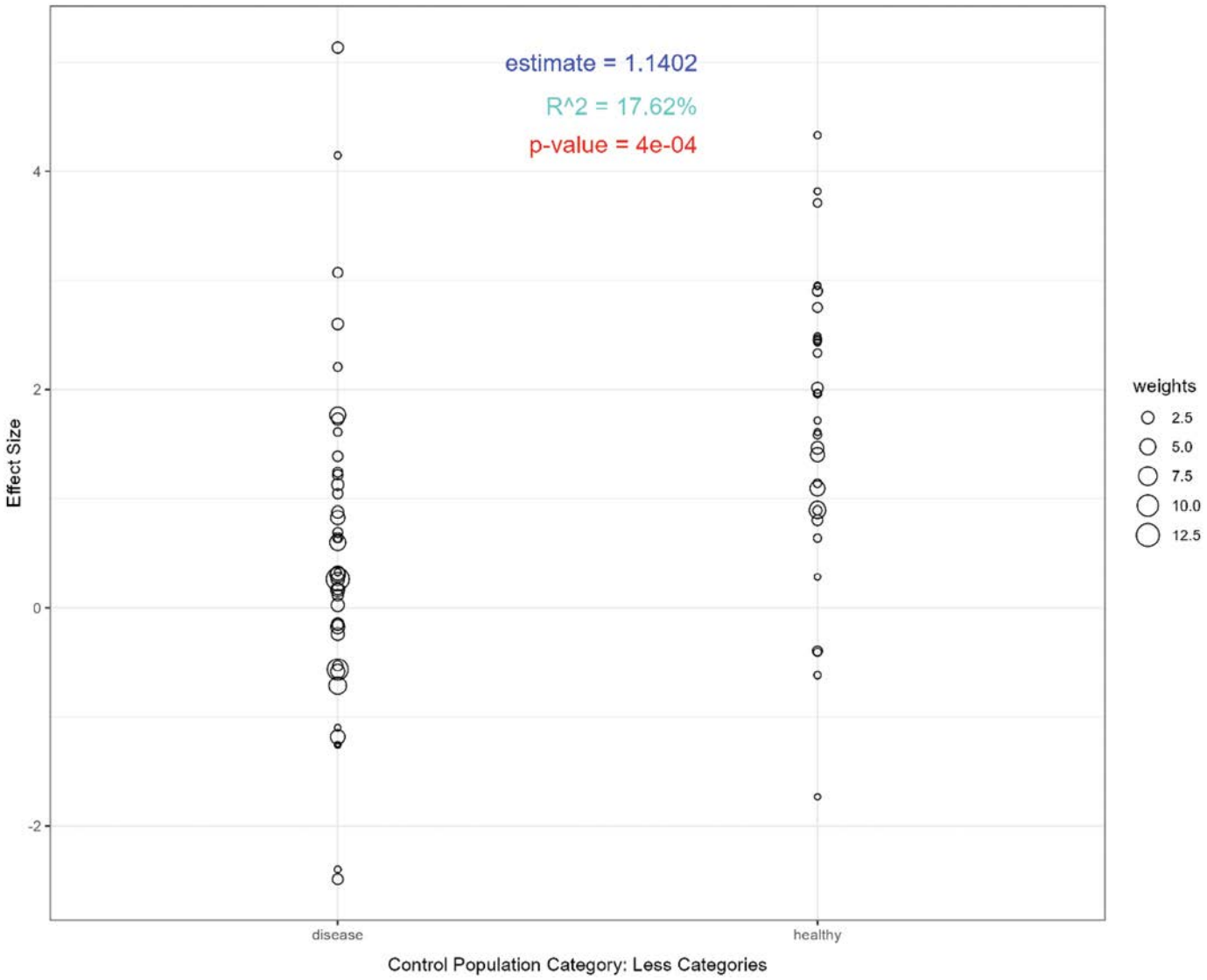


Figure 5. Meta-regression to determine sources of heterogeneity between studies. Studies compared the prevalence of anti-Col2 antibodies using healthy controls rather than disease controls reported significantly higher effect estimates. Col2, collagen II. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25697/abstract.

supporting the concept that anti-Col2 antibodies are not disease specific and can be seen in a range of other conditions.

The current cohort study has several strengths. First, compared to prior studies, this is the largest cohort of patients with RP for whom anti-Col2 antibodies were tested. Second, the cohort study compared RP to a spectrum of related diseases rather than healthy controls. Choosing an appropriate comparator population is crucial because the clinical utility of any diagnostic test is only as good as its ability to differentiate diseases with similar presentations. Third, this study contains the only systematic review and meta-analysis performed to date to determine the clinical utility of anti-Col2 antibodies as a diagnostic biomarker. The study was performed in accordance with PRISMA recommendations. Results from the meta-analysis aligned with the cohort study to provide robust cumulative evidence about the poor performance characteristics of anti-Col2 antibodies to diagnose RP due to low sensitivity and specificity. Fourth, we detected a novel association between anti-Col2 antibodies and cartilage inflammation that was not disease-specific and could inform future experiments regarding the role of anti-Col2 antibodies in the context of cartilage inflammation.

This study also has potential limitations to consider. In the cohort study, anti-Col2 antibody levels were typically measured later into disease in patients who were receiving immunosuppressive therapies. Anti-Col2 antibodies may be higher earlier on in disease onset, such as in juvenile idiopathic arthritis.[21] However, the performance characteristics of anti-Col2 antibodies were poor even when excluding patients on B cell–depleting therapies, and

no association was seen between anti-Col2 and disease activity. Disease activity scores differed at the time of antibody assessment, with the highest scores observed in patients with RP, MAGIC, and VEXAS. Despite these differences, an increased prevalence of anti-Col2 antibodies was not observed in these conditions relative to other disease comparators. The Mayo Clinic (test identifier: FFTYC) Collagen II Antibodies test recognizes a specific epitope, and it is possible that another epitope would show a more significant correlation. Nevertheless, this is the major commercially available test for anti-Col2 antibodies, and thus it is a reasonable epitope on which to analyze performance characteristics. Additionally, the studies analyzed in the systematic review and meta-analysis used a wide range of different antigens and epitopes, yet no consistent pattern of anti-Col2 antibody positivity was associated with a particular disease or cartilage antigen. For the systematic review and meta-analysis, diseases were categorized based on study investigator agreement, so there was potential for misclassification, particularly when differentiating between inflammatory diseases with or without cartilage involvement. The number of articles included within each disease category was variable, and few articles were identified in the predisease group, which may limit the precision of the findings and introduce bias. The definitions for normal varied across the studies, so it can be challenging to know precisely how the deviations in the definitions may have affected positivity rates in the individual study cohorts. However, meta-regression analysis based on study definitions of normal did not reveal a significant contribution to interstudy heterogeneity.

In conclusion, anti-Col2 antibodies have limited utility in clinical practice. These antibodies are neither sensitive nor specific to diagnose RP, and they do not correlate with disease activity or specific clinical features. Anti-Col2 antibodies may be indiscriminately detected in a wide range of diseases in which there is inflammation of cartilage, indicating that they lack the specificity required for them to be used for clinical diagnosis or disease monitoring. Despite a lack of clinical utility, there is continued rationale to study the development and function of anti-Col2 antibodies in research settings. Identification of a diagnostic biomarker that can differentiate RP from related conditions remains a research priority and a major unmet clinical need.

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 Grayson 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. Damiani JM, Levine HL. Relapsing polychondritis--report of ten cases. Laryngoscope 1979;89(6 Pt 1):929–946.
2. McAdam LP, O'Hanlan MA, Bluestone R, et al. Relapsing polychondritis: prospective study of 23 patients and a review of the literature. Medicine (Baltimore) 1976;55(3):193–215.
3. Michet CJ Jr, McKenna CH, Luthra HS, et al. Relapsing polychondritis. Survival and predictive role of early disease manifestations. Ann Intern Med 1986;104(1):74–78.
4. Ebringer R, Rook G, Swana GT, et al. Autoantibodies to cartilage and type II collagen in relapsing polychondritis and other rheumatic diseases. Ann Rheum Dis 1981;40(5):473–479.
5. Foidart JM, Abe S, Martin GR, et al. Antibodies to type II collagen in relapsing polychondritis. N Engl J Med 1978;299(22):1203–1207.
6. McCune WJ, Schiller AL, Dynesius-Trentham RA, et al. Type II collagen-induced auricular chondritis. Arthritis Rheum 1982;25(3): 266–273.
7. Cremer MA, Rosloniec EF, Kang AH. The cartilage collagens: a review of their structure, organization, and role in the pathogenesis of experimental arthritis in animals and in human rheumatic disease. J Mol Med (Berl) 1998;76(3-4):275–288.
8. Cremer MA, Pitcock JA, Stuart JM, et al. Auricular chondritis in rats. An experimental model of relapsing polychondritis induced with type II collagen. J Exp Med 1981;154(2):535–540.
9. Brand DD, Latham KA, Rosloniec EF. Collagen-induced arthritis. Nat Protoc 2007;2(5):1269–1275.
10. Ferrada M, Rimland CA, Quinn K, et al. Defining clinical subgroups in relapsing polychondritis: a prospective observational cohort study. Arthritis Rheumatol 2020;72(8):1396–1402.
11. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 2011;12(1):77.
12. Ouzzani, M., Hammady, H., Fedorowicz, Z. and Elmagarmid, A., Rayyan—a web and mobile app for systematic reviews. Systematic Reviews 2016;5(1): 210.
13. Viechtbauer W. Conducting meta-analyses in R with the metafor package. J Stat Softw 2010;36(3):1–48.
14. Wickham H. ggplot2: Elegant Graphics for Data Analysis. Springer-VerlagNew York; 2016.
15. Balduzzi S, Rücker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. Evid Based Ment Health 2019;22(4): 153–160.
16. Nakamaru Y, Fukuda S, Maguchi S, et al. Relapsing polychondritis: a report of eight cases. Auris Nasus Larynx 2001;28(suppl):S107–S110.
17. Yang CL, Brinckmann J, Rui HF, et al. Autoantibodies to cartilage collagens in relapsing polychondritis. Arch Dermatol Res 1993;285(5): 245–249.
18. Kim HA, Seo YI, Lee J, et al. Detection of anti-type II collagen antibodies in patients with chronic gouty arthritis: findings from a pilot study. J Clin Rheumatol 2016;22(7):360–363.
19. Strollo R, Ponchel F, Malmström V, et al. Autoantibodies to posttranslationally modified type II collagen as potential biomarkers for rheumatoid arthritis. Arthritis Rheum 2013;65(7):1702–1712.
20. Raza K, Mullazehi M, Salmon M, et al. Anti-collagen type II antibodies in patients with very early synovitis. Ann Rheum Dis 2008;67(9):1354–1355.
21. Araujo GR, Fonseca JE, Fujimura PT, et al. Anti-type II collagen antibodies detection and avidity in patients with oligoarticular and polyarticular forms of juvenile idiopathic arthritis. Immunol Lett 2015;165(1): 20–25.

22. Steffen C, Sänger L, Menzel J. Demonstration of antibodies to denatured type I and type II collagen in juvenile rheumatoid arthritis, Still's syndrome and controls by [14C] collagen radioimmunoassay. Scand J Rheumatol 1980;9(2):69–76.
23. Beard HK, Lea DJ, Ryvar R. Anomalous reactions in the haemagglutination assay for anti-collagen antibodies: studies on patients with rheumatoid arthritis or chronic low back pain. J Immunol Methods 1979;31(1-2):119–128.
24. Manivel VA, van der Woude D, Toes R, et al. Antibodies against collagen type II are not a general marker of acute arthritis onset. Ann Rheum Dis 2018;77(6):954–956.
25. Bartholomew JS, Evanson JM, Woolley DE. Serum IgE anti-cartilage collagen antibodies in rheumatoid patients. Rheumatol Int 1991; 11(1):37–40.
26. Boissier MC, Chiocchia G, Texier B, et al. Pattern of humoral reactivity to type II collagen in rheumatoid arthritis. Clin Exp Immunol 1989; 78(2):177–183.
27. Charrière G, Hartmann DJ, Vignon E, et al. Antibodies to types I, II, IX, and XI collagen in the serum of patients with rheumatic diseases. Arthritis Rheum 1988;31(3):325–332.
28. Choi EK, Gatenby PA, McGill NW, et al. Autoantibodies to type II collagen: occurrence in rheumatoid arthritis, other arthritides, autoimmune connective tissue diseases, and chronic inflammatory syndromes. Ann Rheum Dis 1988;47(4):313–322.
29. Clague RB, Firth SA, Holt PJ, et al. Serum antibodies to type II collagen in rheumatoid arthritis: comparison of 6 immunological methods and clinical features. Ann Rheum Dis 1983;42(5):537–544.
30. Clague RB, Morgan K, Reynolds I, et al. The prevalence of serum IgG antibodies to type II collagen in American patients with rheumatoid arthritis. Br J Rheumatol 1994;33(4):336–338.
31. Clague RB, Shaw MJ, Holt PJ. Incidence and correlation between serum IgG and IgM antibodies to native type II collagen in patients with chronic inflammatory arthritis. Ann Rheum Dis 1981;40(1):6–10.
32. Clague RB, Shaw MJ, Holt PJ. Incidence of serum antibodies to native type I and type II collagens in patients with inflammatory arthritis. Ann Rheum Dis 1980;39(3):201–206.
33. Cook AD, Gray R, Ramshaw J, et al. Antibodies against the CB10 fragment of type II collagen in rheumatoid arthritis. Arthritis Res Ther 2004;6(5):R477–R483.
34. Cooper AL, Snowden N, Woolley DE. IgE antibodies specific for cartilage collagens type II, IX and XI in rheumatic diseases. Scand J Rheumatol 1993;22(5):207–214.
35. Fujii K, Tsuji M, Kitamura A, et al. The diagnostic significance of anti-type II collagen antibody assay in rheumatoid arthritis. Int Orthop 1992;16(3):272–276.
36. Goules JD, Goules AV, Tzioufas AG. Fine specificity of anti-citrullinated peptide antibodies discloses a heterogeneous antibody population in rheumatoid arthritis. Clin Exp Immunol 2013;174(1): 10–17.
37. Iribe H, Kabashima H, Ishii Y, et al. Epitope specificity of antibody response against human type II collagen in the mouse susceptible to collagen-induced arthritis and patients with rheumatoid arthritis. Clin Exp Immunol 1988;73(3):443–448.
38. Kammer GM, Trentham DE. HLA-DR4 is not a requisite for autoimmunity to collagen in rheumatoid arthritis. Arthritis Rheum 1984;27(5): 489–494.
39. Kim HY, Kim WU, Cho ML, et al. Enhanced T cell proliferative response to type II collagen and synthetic peptide CII (255-274) in patients with rheumatoid arthritis. Arthritis Rheum 1999;42(10): 2085–2093.
40. Koivula MK, Aman S, Karjalainen A, et al. Are there autoantibodies reacting against citrullinated peptides derived from type I and type II collagens in patients with rheumatoid arthritis? Ann Rheum Dis 2005;64(10):1443–1450.
41. 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.
42. 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.
43. Morgan K, Clague RB, Reynolds I, et al. Antibodies to type II collagen in early rheumatoid arthritis. Br J Rheumatol 1993;32(4):333–335.
44. Mullazehi M, Wick MC, Klareskog L, et al. Anti-type II collagen antibodies are associated with early radiographic destruction in rheumatoid arthritis. Arthritis Res Ther 2012;14(3):R100.
45. Nissim A, Winyard PG, Corrigall V, et al. Generation of neoantigenic epitopes after posttranslational modification of type II collagen by factors present within the inflamed joint. Arthritis Rheum 2005;52(12): 3829–3838.
46. Pesoa SA, Vullo CM, Onetti CM, et al. Rheumatoid arthritis and its association with HLA-DR antigens. II. Antibodies to native connective tissue antigens detected by enzyme linked immunosorbent assay. Autoimmunity 1989;4(3):171–179.
47. Rowley M, Tait B, Mackay IR, et al. Collagen antibodies in rheumatoid arthritis. Significance of antibodies to denatured collagen and their association with HLA-DR4. Arthritis Rheum 1986;29(2):174–184.
48. 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.
49. Terato K, DeArmey DA, Ye XJ, et al. The mechanism of autoantibody formation to cartilage in rheumatoid arthritis: possible cross-reaction of antibodies to dietary collagens with autologous type II collagen. Clin Immunol Immunopathol 1996;79(2):142–154.
50. Terato K, Shimozuru Y, Katayama K, et al. Specificity of antibodies to type II collagen in rheumatoid arthritis. Arthritis Rheum 1990;33(10): 1493–1500.
51. Turunen S, Hannonen P, Koivula MK, et al. Separate and overlapping specificities in rheumatoid arthritis antibodies binding to citrulline- and homocitrulline-containing peptides related to type I and II collagen telopeptides. Arthritis Res Ther 2015;17(1):2.
52. Yoshida M, Tsuji M, Kurosaka D, et al. Autoimmunity to citrullinated type II collagen in rheumatoid arthritis. Mod Rheumatol 2006;16(5): 276–281.
53. Zeng P, Pertsinidou E, Brynedal B, et al. Occupational physical workload and development of anti-collagen type II antibodies in rheumatoid arthritis: results from the Swedish EIRA population-based case-control study. Clin Exp Rheumatol 2020;38(5):1029–1030.
54. Stuart JM, Huffstutter EH, Townes AS, et al. Incidence and specificity of antibodies to types I, II, III, IV, and V collagen in rheumatoid arthritis and other rheumatic diseases as measured by 125I-radioimmunoassay. Arthritis Rheum 1983;26(7):832–840.
55. Too CL, Murad S, Hansson M, et al. Differences in the spectrum of anti-citrullinated protein antibody fine specificities between Malaysian and Swedish patients with rheumatoid arthritis: implications for disease pathogenesis. Arthritis Rheumatol 2017;69(1):58–69.
56. Vergouwen DPC, van Beek AA, de Hoog J, et al. The enigma of sclera-specific autoimmunity in scleritis. J Autoimmun 2024;144: 103178.
57. Vinci C, Infantino M, Raturi S, et al. Immunoglobulin A antibodies to oxidized collagen type II as a potential biomarker for the stratification of spondyloarthritis from rheumatoid arthritis. Scand J Rheumatol 2020;49(4):281–291.

58. Gay S, Fine JD, Storer JS. Autoantibodies to extracellular collagen matrix components in epidermolysis bullosa and other bullous diseases. Arch Dermatol Res 1988;280(6):333–337.
59. Kobayashi S, Wada N, Kubo M. Antibodies to native type III collagen in the serum of patients with Kawasaki disease. Eur J Pediatr 1992; 151(3):183–187.
60. He C, Mao T, Feng Y, et al. Anti-CII antibody as a novel indicator to assess disease activity in systemic lupus erythematosus. Lupus 2015;24(13):1370–1376.
61. Strollo R, Rizzo P, Spoletini M, et al. HLA-dependent autoantibodies against post-translationally modified collagen type II in type 1 diabetes mellitus. Diabetologia 2013;56(3):563–572.
62. Moreland LW, Gay RE, Gay S. Collagen autoantibodies in patients with vasculitis and systemic lupus erythematosus. Clin Immunol Immunopathol 1991;60(3):412–418.
63. Elshafie AI, Mullazehi M, Rönnelid J. General false positive ELISA reactions in visceral leishmaniasis. Implications for the use of enzyme immunoassay analyses in tropical Africa. J Immunol Methods 2016; 431:66–71.
64. Stolovitzky JP, Todd NW, Cotton RT, et al. Autoimmune hypothesis of acquired subglottic stenosis: lack of support at time of surgical repair in children. Int J Pediatr Otorhinolaryngol 1997;38(3):255–261.
65. Helfgott SM, Mosciscki RA, San Martin J, et al. Correlation between antibodies to type II collagen and treatment outcome in bilateral progressive sensorineural hearing loss. Lancet 1991;337(8738): 387–389.
66. Levine JJ, Lin HC, Rowley M, et al. Lack of autoantibody expression in children born to mothers with silicone breast implants. Pediatrics 1996;97(2):243–245.
67. Neumüller J, Menzel J, Millesi H. Prevalence of HLA-DR3 and autoantibodies to connective tissue components in Dupuytren's contracture. Clin Immunol Immunopathol 1994;71(2):142–148.
68. Clague RB, Morgan K, Collins I, et al. Absence of autoimmunity to type II collagen in generalised nodal osteoarthritis. Ann Rheum Dis 1991;50(11):769–771.
69. López-González MA, Lucas M, Sánchez B, et al; Autoimmune deafness is not related to hyperreactivity to type II collagen. Acta Otolaryngol 1999;119(6):690–694.
70. Nakos G, Adams A, Andriopoulos N. Antibodies to collagen in patients with idiopathic pulmonary fibrosis. Chest 1993;103(4): 1051–1058.
71. Herdman RC, Morgan K, Holt PJ, et al. Type II collagen autoimmunity and Ménière's disease. J Laryngol Otol 1993;107(11):994–998.
72. Koo JW, Oh SH, Chang SO, et al. Association of HLA-DR and type II collagen autoimmunity with Meniere's disease. Tissue Antigens 2003;61(1):99–103.
73. Paróczai C, Németh-Csóka M. Estimation of serum anticollagen and the antibodies against chondrocyte membrane fraction: their clinical diagnostic significance in osteoarthritis. Clin Biochem 1988;21(2): 117–121.
74. Ruthard J, Hermes G, Hartmann U, et al. Identification of antibodies against extracellular matrix proteins in human osteoarthritis. Biochem Biophys Res Commun 2018;503(3):1273–1277.
75. Jesić S, Radulović R, Arsović N. Altered immunoregulations in otosclerosis: presence of autoantibodies in otosclerotic sera samples. Eur Arch Oto-Rhino-Laryngol 1997;254(suppl 1):S50–S52.
76. Yoo TJ, Stuart JM, Kang AH, et al. Type II collagen autoimmunity in otosclerosis and Meniere's Disease. Science 1982;217(4565):1153–1155.
77. Bar-Meir E, Teuber SS, Lin HC, et al. Multiple autoantibodies in patients with silicone breast implants. J Autoimmun 1995;8(2): 267–277.
78. Rowley MJ, Cook AD, Teuber SS, et al. Antibodies to collagen: comparative epitope mapping in women with silicone breast implants, systemic lupus erythematosus and rheumatoid arthritis. J Autoimmun 1994;7(6):775–789.
79. Teuber SS, Rowley MJ, Yoshida SH, et al. Anti-collagen autoantibodies are found in women with silicone breast implants. J Autoimmun 1993;6(3):367–377.
80. Brink M, Hansson M, Mathsson L, et al. Multiplex analyses of antibodies against citrullinated peptides in individuals prior to development of rheumatoid arthritis. Arthritis Rheum 2013;65(4):899–910.
81. Kokkonen H, Brink M, Hansson M, et al. Associations of antibodies against citrullinated peptides with human leukocyte antigen-shared epitope and smoking prior to the development of rheumatoid arthritis. Arthritis Res Ther 2015;17(1):125.
82. Möttönen T, Hannonen P, Oka M, et al. Antibodies against native type II collagen do not precede the clinical onset of rheumatoid arthritis. Arthritis Rheum 1988;31(6):776–779.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 78, No. 6, June 2026, pp 810–812

LETTER

DOI 10.1002/acr.25688

Bridging clinical relevance and equity in the cost-effectiveness of standardized exercise therapy for osteoarthritis: comment on the article by Mazzei et al

To the Editor:

We read with great interest the recent article entitled, "Real-world cost-effectiveness of a standardized education and exercise therapy program for hip and knee osteoarthritis compared to usual care,"[1] which provides valuable insights into the economic and clinical implications of implementing standardized education and exercise programs within community health systems. This study is commendable for addressing the real-world performance of such interventions, an area often underexplored compared to controlled trial settings. From a clinical practice perspective, however, several issues merit further reflection to optimize translation of these findings into health care policy and routine care.

First, although the study highlights that the standardized program was associated with improvements in pain, function, and quality of life, the magnitude of these benefits did not reach the minimal clinically important difference (MCID). This raises an important question for clinicians: how can modest statistical improvements be transformed into outcomes that are meaningful for patients in everyday practice? One potential solution is to integrate structured adherence monitoring and individualized support within the program. Patient engagement in exercise interventions is notoriously variable, and suboptimal adherence may attenuate clinical gains. Incorporating digital health tools, such as wearable activity trackers or mobile applications, could provide real-time feedback and enhance patient accountability.[2] Additionally, tailoring exercise intensity and progression to individual capabilities—rather than relying solely on a standardized schedule—might yield greater improvements that cross the MCID threshold. These enhancements would not only strengthen clinical effectiveness but also amplify the cost-effectiveness of the intervention from both public and private payer perspectives.

Second, the comparator group of "usual care" encompassed a heterogeneous mix of community-based resources, many of which are not guideline-recommended first-line treatments. For example, the reliance on massage, chiropractic, or acupuncture by participants substantially inflated out-of-pocket expenditures, thereby diluting the comparative value of the standardized program when a health care perspective was considered. From a clinical standpoint, this heterogeneity complicates the interpretation of findings because it blurs the contrast between evidence-based and non–evidence-based care. A more nuanced approach in future studies would be to stratify "usual care" according to guideline concordance. This would enable a clearer assessment of whether structured education and exercise programs outperform truly recommended alternatives or whether their observed advantage primarily reflects displacement of less effective practices. Such differentiation would provide clinicians and policymakers with more precise evidence on which interventions deserve prioritization in resource-limited health care environments.

Third, the lifetime cost-effectiveness estimates relied on Markov modeling assumptions that all-cause mortality risks were equal across health states and that prosthetic joint outcomes were homogenous over time. These simplifications, although necessary for modeling feasibility, may understate the long-term clinical risks associated with joint replacement, such as revision surgery, infection, or reduced functional gains in older adults. From a clinical practice perspective, these assumptions limit the ability to fully appreciate the potential role of first-line treatments in delaying or preventing surgery. More robust modeling could incorporate differential mortality risks and surgical complications, ideally informed by national joint registries. In parallel, linking exercise adherence and weight management to surgical incidence would yield more realistic estimates of how such interventions influence long-term patient trajectories. By doing so, clinicians would gain a more reliable understanding of how nonsurgical strategies can meaningfully alter the natural history of osteoarthritis, rather than merely deferring surgical intervention.

Finally, an important implication for clinical practice lies in the equity of access. The study revealed that nearly 60% of osteoarthritis-related health care costs were borne by patients out-of-pocket.[1] In practice, this creates a socioeconomic gradient in who can realistically benefit from structured first-line programs. For clinicians, recommending such interventions without acknowledging their financial burden may inadvertently widen health disparities because patients with lower income are less able to participate. This underscores the necessity of policy reform to include standardized education and exercise programs within the publicly funded basket of services. In the meantime, clinicians should proactively screen for financial barriers during consultations and consider referral pathways that connect patients to subsidized community programs or virtual delivery models, which can reduce costs associated with travel and facility-based care. Addressing affordability

directly in clinical encounters will help mitigate inequities until system-level solutions are implemented.

In conclusion, this study adds substantially to the evidence base supporting standardized education and exercise therapy for hip and knee osteoarthritis. Yet, for its findings to translate into meaningful clinical benefit, further refinement is needed in program adherence strategies, comparator group characterization, long-term modeling, and equitable access. By tackling these areas, future research and clinical implementation can ensure that standardized first-line interventions not only deliver statistical improvements but also generate clinically meaningful, equitable, and sustainable outcomes for patients living with osteoarthritis.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25688.

Yadi Li, MD
yadili9232@gmail.com
Zheng Wei, MD
Deyang People's Hospital
Deyang, China
Jianlong Zhou, MD
Chengdu University of Traditional Chinese Medicine
Chengdu, China

1. Mazzei DR, Whittaker JL, Faris P, et al. Real-world cost-effectiveness of a standardized education and exercise therapy program for hip and knee osteoarthritis compared to usual care. Arthritis Care Res (Hoboken) 2025;77:1194–1203.
2. Kang HS, Exworthy M. Wearing the future—wearables to empower users to take greater responsibility for their health and care: scoping review. JMIR Mhealth Uhealth 2022;10(7):e35684.

DOI 10.1002/acr.25689

Reply

To the Editor:

We thank Drs Li et al for their interest and commendations about our cost-effectiveness analysis (CEA) of standardized exercise and education programs compared to joint replacement as a plausible first-line guideline-supported nonsurgical option for patients waiting for surgical consultation and who may benefit.[1] They raise many good points, to which we have responded below.

(1) Minimal clinically important difference (MCID): There remains no consensus about the MCID for osteoarthritis interventions such as joint replacement surgery. There is substantial variability in both the definition of MCID (distribution-based vs anchor-based approaches) and the population in which MCIDs are measured. For example, a systematic review reported a median anchor-based MCID for Hip Disability and Osteoarthritis Outcome Score (HOOS) quality of life of 25 (interquartile range [IQR] 14–27) and a median distribution-based MCID of 13 (IQR 10–14) for hip arthroplasty.[2] Alternative measures such as minimal important difference, minimal important change (MIC), and patient acceptable symptom state have also been suggested. In our study, all outcome measures yielded positive improvements for GoodLife with osteoArthritis in Denmark (GLA:D®) compared to usual care. Using an MIC of 5.5, our findings (mean difference of 7.14) for the HOOS/Knee Injury and Osteoarthritis Outcome Score quality of life measures are an important change.

(2) Solutions to improve clinical effectiveness and cost-effectiveness: We welcome the suggested solutions to transform statistical improvements into outcomes that are clinically meaningful. Poor adherence may attenuate gains from exercise, and enhancements such as digital health tools to improve attendance would only increase the clinical benefits of these interventions and improve cost-effectiveness. Similarly, tailoring and virtual adaptations could be promising enhancements. Studies combining wearable devices, apps, and coaching have demonstrated modest but significant improvements in physical activity and related patient outcomes.[3]

(3) Improving estimates given heterogeneity of usual care comparator: Our "usual care" comparator included a heterogeneous mix of treatments, some of which are likely not evidence-based—reflecting the "real-world" context in which patients access a mixture of options to manage their osteoarthritis. If we were to limit the usual care comparator to only guideline-recommended treatments, we expect this would affect not only the difference in out-of-pocket expenditures by potentially diluting the comparative cost but also may impact the difference in outcomes and comparative clinical effectiveness. Thus, it would be difficult to predict the net effect on the cost-effectiveness estimate.

(4) Dynamic simulation modeling approaches and optimization to prioritize resource allocation: Our Markov model is a simplification that likely understates the risks associated with joint replacement and the potential clinical benefits of first-line treatments in delaying or preventing surgery. A survivor effect could offset some of these surgical risks because people with higher mortality risks may not survive to receive joint replacement surgery. A dynamic simulation model would better capture the evolving benefits, risks, and outcomes of treatment interventions across the osteoarthritis disease trajectory. Dynamic simulation models (eg, discrete event simulation or agent-based modeling) are well suited in a health system context to enable modeling of individual behaviors (eg, exercise adherence) and broader population-level effects over time.[4] These models can capture patient heterogeneity and time-varying risks (eg, mortality risk by age, comorbidities, and surgical status). These features are relevant for system-level evaluation of real-world implementation of health care interventions and long-term impacts. To expand consideration to all possible alternative interventions simultaneously and inform decision makers about resource allocation across the range of all possible alternatives, constrained optimization modeling methods would be helpful.[5]

directly in clinical encounters will help mitigate inequities until system-level solutions are implemented.

In conclusion, this study adds substantially to the evidence base supporting standardized education and exercise therapy for hip and knee osteoarthritis. Yet, for its findings to translate into meaningful clinical benefit, further refinement is needed in program adherence strategies, comparator group characterization, long-term modeling, and equitable access. By tackling these areas, future research and clinical implementation can ensure that standardized first-line interventions not only deliver statistical improvements but also generate clinically meaningful, equitable, and sustainable outcomes for patients living with osteoarthritis.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25688.

Yadi Li, MD
yadili9232@gmail.com
Zheng Wei, MD
Deyang People's Hospital
Deyang, China
Jianlong Zhou, MD
Chengdu University of Traditional Chinese Medicine
Chengdu, China

1. Mazzei DR, Whittaker JL, Faris P, et al. Real-world cost-effectiveness of a standardized education and exercise therapy program for hip and knee osteoarthritis compared to usual care. Arthritis Care Res (Hoboken) 2025;77:1194–1203.
2. Kang HS, Exworthy M. Wearing the future—wearables to empower users to take greater responsibility for their health and care: scoping review. JMIR Mhealth Uhealth 2022;10(7):e35684.

DOI 10.1002/acr.25689

Reply

To the Editor:

We thank Drs Li et al for their interest and commendations about our cost-effectiveness analysis (CEA) of standardized exercise and education programs compared to joint replacement as a plausible first-line guideline-supported nonsurgical option for patients waiting for surgical consultation and who may benefit.[1] They raise many good points, to which we have responded below.

(1) Minimal clinically important difference (MCID): There remains no consensus about the MCID for osteoarthritis interventions such as joint replacement surgery. There is substantial variability in both the definition of MCID (distribution-based vs anchor-based approaches) and the population in which MCIDs are measured. For example, a systematic review reported a median anchor-based MCID for Hip Disability and Osteoarthritis Outcome Score (HOOS) quality of life of 25 (interquartile range [IQR] 14–27) and a median distribution-based MCID of 13 (IQR 10–14) for hip arthroplasty.[2] Alternative measures such as minimal important difference, minimal important change (MIC), and patient acceptable symptom state have also been suggested. In our study, all outcome measures yielded positive improvements for GoodLife with osteoArthritis in Denmark (GLA:D®) compared to usual care. Using an MIC of 5.5, our findings (mean difference of 7.14) for the HOOS/Knee Injury and Osteoarthritis Outcome Score quality of life measures are an important change.

(2) Solutions to improve clinical effectiveness and cost-effectiveness: We welcome the suggested solutions to transform statistical improvements into outcomes that are clinically meaningful. Poor adherence may attenuate gains from exercise, and enhancements such as digital health tools to improve attendance would only increase the clinical benefits of these interventions and improve cost-effectiveness. Similarly, tailoring and virtual adaptations could be promising enhancements. Studies combining wearable devices, apps, and coaching have demonstrated modest but significant improvements in physical activity and related patient outcomes.[3]

(3) Improving estimates given heterogeneity of usual care comparator: Our "usual care" comparator included a heterogeneous mix of treatments, some of which are likely not evidence-based—reflecting the "real-world" context in which patients access a mixture of options to manage their osteoarthritis. If we were to limit the usual care comparator to only guideline-recommended treatments, we expect this would affect not only the difference in out-of-pocket expenditures by potentially diluting the comparative cost but also may impact the difference in outcomes and comparative clinical effectiveness. Thus, it would be difficult to predict the net effect on the cost-effectiveness estimate.

(4) Dynamic simulation modeling approaches and optimization to prioritize resource allocation: Our Markov model is a simplification that likely understates the risks associated with joint replacement and the potential clinical benefits of first-line treatments in delaying or preventing surgery. A survivor effect could offset some of these surgical risks because people with higher mortality risks may not survive to receive joint replacement surgery. A dynamic simulation model would better capture the evolving benefits, risks, and outcomes of treatment interventions across the osteoarthritis disease trajectory. Dynamic simulation models (eg, discrete event simulation or agent-based modeling) are well suited in a health system context to enable modeling of individual behaviors (eg, exercise adherence) and broader population-level effects over time.[4] These models can capture patient heterogeneity and time-varying risks (eg, mortality risk by age, comorbidities, and surgical status). These features are relevant for system-level evaluation of real-world implementation of health care interventions and long-term impacts. To expand consideration to all possible alternative interventions simultaneously and inform decision makers about resource allocation across the range of all possible alternatives, constrained optimization modeling methods would be helpful.[5]

(5) Affordability and equity considerations: This article focused specifically on the cost-effectiveness (measured as the incremental cost per quality-adjusted life-year gained and net monetary benefit) of standardized exercise and education as a guideline-recommended first-line intervention, given that we had launched the GLA:D® program across the province. However, many patients had to pay out of pocket for the program—thus discouraging uptake and creating inequities. To support translation of these findings to inform health care policy, we also examined affordability in our companion budget impact analysis (BIA) article to estimate the annual costs and benefits of providing the program to people waiting for a surgical consult.[6] Based on observed practice in our clinics and published evidence, we conservatively estimated that 40% of patients would attend the sessions and 11% would avoid surgery based on improvements experienced from the program. Even these conservative parameter estimates resulted in estimates of cost-saving. We concluded that public funding of these programs could help reduce surgeries, improve access to evidence-based care, and yield net savings.

Equity of access to treatment is another consideration. GLA:D® is currently offered in 73 registered Alberta locations, of which 47 are publicly funded. The remaining 26 sites are private, and participants pay out of pocket for the program. One of the mechanisms to address the equity issue was to present a business case for publicly funded clinic options to be made available across the province as a means of minimizing the imbalance. This was a key motivation for undertaking this research. Our BIA demonstrates the potential value of public investment in standardized education and exercise interventions from the payer perspective as well as the patient perspective.

In conclusion, we hope that our CEA and BIA findings, along with proposed further enhancements to improve clinical effectiveness and cost-effectiveness, will support the implementation of evidence-based first-line solutions to improve access to treatment for patients across all socioeconomic classes and provide relief to a highly strained health care system with long wait times for surgical consultation.

Deborah A. Marshall, PhD
damarsha@ucalgary.ca
Department of Community Health Sciences and McCaig Institute for Bone and Joint Health, Cumming School of Medicine, University of Calgary
Calgary, Alberta, Canada
Tracy Wasylak, MSc
Faculty of Nursing, University of Calgary
Calgary, Alberta, Canada
Peter Faris, PhD
Department of Community Health Sciences, Cumming School of Medicine, University of Calgary Calgary, Alberta, Canada and Alberta Health Services
Edmonton, Alberta, Canada
Darren Mazzei, PhD
Department of Community Health Sciences, Cumming School of Medicine, University of Calgary
Calgary, Alberta, Canada
Jackie Whittaker, PhD
Department of Physical Therapy, University of British Columbia
Vancouver, British Columbia, Canada

1. Mazzei DR, Whittaker JL, Faris P, et al. Real-world cost-effectiveness of a standardized education and exercise therapy program for hip and knee osteoarthritis compared to usual care. Arthritis Care Res (Hoboken) 2025;77:1194–1203.
2. Deckey DG, Verhey JT, Christopher ZK, et al. Discordance abounds in minimum clinically important differences in THA: a systematic review. Clin Orthop Relat Res 2023;481:702–714.
3. Li LC, Feehan LM, Xie H, et al. Effects of a 12-week multifaceted wearable-based program for people with knee osteoarthritis: randomized controlled trial. JMIR Mhealth Uhealth 2020;8:e19116.
4. Marshall DA, Burgos-Liz L, Pasupathy et al. Transforming healthcare delivery: integrating dynamic simulation modeling and big data in health economics and outcomes research. PharmacoEconomics 2016;34(2): 115–126.
5. Crown W, Buyukkaramikli N, Sir MY, et al. Application of constrained optimization methods in health services research: report 2 of the ISPOR Optimization Methods Emerging Good Practices Task Force. Value Health 2018;21(9):1019–1028.
6. Mazzei DR, Whittaker JL, Faris P, et al. Estimating budget impact and joint replacement avoidance by implementing a standardized education and exercise therapy program for hip and knee osteoarthritis in a publicly insured health care system. Arthritis Care Res (Hoboken) 2025; 77:744–752.