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. 3 March 2026

Arthritis Care & Research

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

Arthritis Care & Research

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

VOLUME 78 • MARCH 2026 • NO. 3

Rheumatoid Arthritis

Evaluating the Longitudinal Association of Rheumatoid Arthritis Disease Activity With Phenotypic Frailty: Evidence for Secondary Frailty?
Hannah F. Brubeck, Kylie E. Riggles, Riley S. Bass, Elizabeth R. Wahl, George Mount, Dolores M. Shoback, James S. Andrews, Jose M. Garcia, Ariela R. Orkaby, Joshua F. Baker, Patricia P. Katz, Katherine D. Wysham, and Courtney N. Loecker 307

Adipokines and Associations With Incident Osteoporotic Fracture in Patients With Rheumatoid Arthritis
Joshua F. Baker, Bryant R. England, Michael D. George, Hannah Brubeck, Brian Sauer, Aleksander Lenert, Punyasha Roul, Geoffrey M. Thiele, Ted R. Mikuls, and Katherine D. Wysham 316

Review Article: Economic Burden of Rheumatoid Arthritis in Low- and Middle-Income Countries: Systematic Review and Meta-Analysis
Tadesse Gebrye, Chidozie E. Mbada, Clara T. Fatoye, Faatihah Niyi-Odumosu, Ushotanefe Useh, Zalmai Hakimi, and Francis Fatoye 325

Gout

Cost-Effectiveness of Low-Dose Colchicine Prophylaxis When Starting Allopurinol Using the "Start-Low Go-Slow" Approach for Gout: Evidence From a Noninferiority Randomized Double-Blind Placebo-Controlled Trial
Yana Pryymachenko, Ross Wilson, Nicola Dalbeth, J. Haxby Abbott, and Lisa Stamp 337

Long-Term Opioids in Gout: A Matched Cohort Study From the Veterans Health Administration
Lindsay N. Helget, Bryant R. England, Punyasha Roul, Harlan Sayles, Tuhina Neogi, James R. O'Dell, Joshua F. Baker, and Ted R. Mikuls 344

Systemic Sclerosis

International Investigation of the Gut-Lung Axis in Systemic Sclerosis–Interstitial Lung Disease
Kristofer Andréasson, Arissa Young, Swapna Joshi, Jennifer S. Labus, Andrea Hsiu Ling Low, Vanessa Smith, Zsuzsanna McMahan, Susanna Proudman, Antonia Valenzuela, Phoebe Hunter, Grace Hyun Kim, Gracijela Bozovic, Jonathan Goldin, Ezinne Aja, Jonathan P. Jacobs, and Elizabeth R. Volkmann. 352

Dermatomyositis

Identification of Adult Patients With Dermatomyositis Using Real-World Data Sources
Benjamin Martin, Will Kelly, Hannah Morgan-Cooper, Thomas Falconer, Elizabeth Park, Priya Desai, David Fiorentino, Lorinda Chung, Sean Yen, Zachary Wang, Didem Saygin, Michael George, Gowtham A. Rao, Joel Swerdel, Azza Shoaibi, and Christopher A. Mecoli 362

Systemic Lupus Erythematosus

Identifying High-Impact Solutions to Address Racial and Ethnic Health Disparities in Lupus: A Consensus-Based Approach
Joy Buie, Michael P. Fisher, Kristen Backor, Hannah Tyldsley, Ashira Blazer, Candace Feldman, Andrea Knight, S. Sam Lim, Barbara Ann Polk, Ed Yelin, Edith Williams, and Karen H. Costenbader 371

Prevalence, Determinants, and Outcomes of Low Disease Activity and Remission Attainment in Patients With Systemic Lupus Erythematosus That Is Clinically Active
Yanjie Hao, Dylan Hansen, Rangi Kandane-Rathnayake, Worawit Louthrenoo, Yi-Hsing Chen, Jiacai Cho, Aisha Lateef, Laniyati Hamijoyo, Shue Fen Luo, Yeong-Jian Jan Wu, Sandra Navarra, Leonid Zamora, Zhanguo Li, Sargunan Sockalingam, Yasuhiro Katsumata, Masayoshi Harigai, Zhuoli Zhang, Madelynn Chan, Jun Kikuchi, Tsutomu Takeuchi, Sang-Cheol Bae, Fiona Goldblatt, Sean O'Neill, Kristine (Pek Ling) Ng, Annie Law, B. M. D. B. Basnayake, Nicola Tugnet, Sunil Kumar, Cherica Tee, Michael Tee, Naoaki Ohkubo, Yoshiya Tanaka, Shirley Chan, C. S. Lau, Vera Golder, Alberta Hoi, Shereen Oon, Eric Morand, and Mandana Nikpour, for the Asia-Pacific Lupus Collaboration 378

Osteoarthritis

Virtual or In-Person: Does It Matter? Comparing Pain, Function, Quality of Life, Self-Efficacy, and Physical Function Outcomes of Virtual, Hybrid, and In-Person Education and Exercise Program Participants
Jill Van Damme, Vanina Dal Bello-Haas, Ayse Kuspinar, Patricia Strachan, Michael Zywiel, James Young, Rhona McGlasson, and Lisa C. Carlesso 388

Knee Crepitus and Osteoarthritis Features in Young Adults Following Traumatic Knee Injury
Jamon L. Couch, Brooke E. Patterson, Kay M. Crossley, Ali Guermazi, Matthew G. King, Danilo De Oliveira Silva, Jackie L. Whittaker, Michael A. Girdwood, and Adam G. Culvenor 398

Other and Mixed Rheumatic Diseases

Brief Report: Serum Uric Acid Levels in Older Adults: Associations With Clinical Outcomes and Implications for Reference Intervals in Those Aged 70 Years and Over
Amanda J. Rickard, Cammie Tran, Hans G. Schneider, Flavia M. Cicuttini, Anita E. Wluka, Ego Seeman, Johannes T. Neumann, Md Nazmul Karim, Zhen Zhou, Sultana Monira Hussain, David P. Q. Clark, Daniel Clayton-Chubb, Andrew M. Tonkin, Lawrence J. Beilin, Robyn L. Woods, and John J. McNeil 407

Perceptions About Asymptomatic Hyperuricemia and Views About Urate-Lowering Therapy in People With Asymptomatic Hyperuricemia
Nicola Dalbeth, Sarah Stewart, Gregory D. Gamble, Borislav Mihov, Lisa K. Stamp, Janine Haslett, William J. Taylor, Tony R. Merriman, Adwoa Dansoa Tabi-Amponsah, Anne Horne, Tuhina Neogi, Tristan Pascart, Mariano Andrés, Maria-Luisa Peral-Garrido, Eleonora Norkuviene, Janitzia Vazquez-Mellado, Till Uhlig, Mingshu Sun, Changgui Li, and Keith J. Petrie 417

Letter

Reframing Pain Modulation and Disease-Modifying Antirheumatic Drugs De-Escalation: Lessons from a Lifestyle Intervention in Rheumatoid Arthritis and Osteoarthritis: Comment on the Article by Wagenaar et al
Yujie Xu and Hua Xu 424

Reply

Carlijn A. Wagenaar, Dirkjan van Schaardenburg, Wendy Walrabenstein, Marike van der Leeden, Franktien Turkstra, Martijn Gerritsen, Jos W. R. Twisk, Maarten Boers, Martin van der Esch, Henriët van Middendorp, and Peter J. M. Weijs 425

Erratum

Correction to "Effects of Social Vulnerability and Environmental Burden on Care Fragmentation and Social Needs Among Individuals With Rheumatic Conditions" 427

Cover image: The image on the cover (Andréasson et al; pages 352–361) shows results of a beta diversity analysis (e.g., differences in overall gastrointestinal microbial composition), performed using robust Atchinson distance. Each dot represents a patient sample and is color-coded based on study site.

Arthritis Care & Research
Vol. 78, No. 3, March 2026, pp 307–315
DOI 10.1002/acr.25615

Evaluating the Longitudinal Association of Rheumatoid Arthritis Disease Activity With Phenotypic Frailty: Evidence for Secondary Frailty?

Hannah F. Brubeck,[1] Kylie E. Riggles,[1] Riley S. Bass,[1] Elizabeth R. Wahl,[2] George Mount,[2] Dolores M. Shoback,[3] James S. Andrews,[4] Jose M. Garcia,[5] Ariela R. Orkaby,[6] Joshua F. Baker,[7] Patricia P. Katz,[8] Katherine D. Wysham,[5] and Courtney N. Loecker[9]

Objective. The cross-sectional association between rheumatoid arthritis (RA) disease activity and frailty has been described; however, the longitudinal relationship is less well understood. We evaluated the association between disease activity and frailty over time.

Methods. We used a longitudinal RA cohort established at the Department of Veterans Affairs Puget Sound Health Care System. RA disease activity (Disease Activity Score in 28 joints using the C-reactive protein level [DAS28-CRP]) and frailty (measured by the Fried Frailty Phenotype [FFP]) were evaluated at baseline and at one year. Frailty was categorized as robust, prefrail, or frail. Ordinal logistic regressions assessed the cross-sectional associations of DAS28-CRP and frailty at baseline and at one year. Paired *t*-tests and multivariable mixed ordinal logistic regressions assessed the longitudinal associations of DAS28-CRP and frailty. Models were adjusted for age, sex, disease duration, prednisone use, conventional synthetic disease-modifying antirheumatic drug (DMARD) use, and biologic DMARD use.

Results. A total of 132 patients with RA aged 64.2 ± 11.3 years were included; 73% were male, 69% were White, 11% were Black, and 12% reported multiple races. The mean ± SD baseline DAS28-CRP was 3.9 ± 1.3, and frailty was categorized as robust in 35 (27%) patients, prefrail in 77 (58%) patients, and frail in 20 (15%) patients. DAS28-CRP (per 1-unit increase) was associated with a higher FFP category at baseline (adjusted odds ratio [aOR] 1.98, $P < 0.0001$). Disease activity increased in those whose frailty score worsened at one year (mean ± SD change score 0.61 ± 0.96, $P = 0.0121$). An increased DAS28-CRP was independently associated with a higher frailty category over one year (aOR 3.31, $P < 0.0001$; n = 65).

Conclusion. Disease activity is independently associated with phenotypic frailty. Increased disease activity over time is associated with worsening frailty status. Future studies are needed to explore this longitudinal relationship and to determine if controlling disease activity can mitigate frailty.

The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs or the US government.

Supported by the Department of Veterans Affairs (VA) Clinical Science Research and Development (grant IK2-CX-002351) and the VA Puget Sound Health Care System Geriatrics Research Education and Clinical Center. Dr Garcia's work was supported by the US Department of Veterans Affairs (grant BX-002807), the Congressionally Directed Medical Research Programs (grant PC-170059), and the NIH (grants R01-CA-239208, R01-AG-061558, and R01-CA-279220). Dr Baker's work was supported by two Rehabilitation Research & Development Merit Awards (grants RX-003644 and RX-004770).

[1]Hannah F. Brubeck, BS, Kylie E. Riggles, BS, Riley S. Bass, BS: VA Puget Sound Health Care System, Seattle, Washington; [2]Elizabeth R. Wahl, MD, MAS, George Mount, MD: VA Puget Sound Health Care System and University of Washington, Seattle, Washington; [3]Dolores M. Shoback, MD: San Francisco VA Medical Center and University of California San Francisco, San Francisco, California; [4]James S. Andrews, MD: University of Alabama at Birmingham, Birmingham VA Health Care System, and Atlanta-Birmingham VA Geriatric Research Education and Clinical Center, Birmingham, Alabama; [5]Jose M. Garcia, MD, PhD, Katherine D. Wysham, MD: VA Puget Sound Health Care System, Puget Sound Geriatric Research Education and Clinical Center, and University of Washington, Seattle, Washington; [6]Ariela R. Orkaby, MD: New England Geriatric Research, Education, and Clinical Center, VA Boston Healthcare System, Brigham and Women's Hospital, and Harvard Medical School, Boston, Massachusetts; [7]Joshua F. Baker, MD, MSCE: Corporal Crescenz VA Medical Center and University of Pennsylvania, Philadelphia, Pennsylvania; [8]Patricia P. Katz, PhD: University of California San Francisco, San Francisco, California; [9]Courtney N. Loecker, PhD, MSN, APRN-NP: University of Nebraska Medical Center and Omaha VA Medical Center, Omaha, Nebraska.

Drs Wysham and Loecker are co-last authors and contributed equally to this work.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25615).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25615.

Address correspondence via email to Katherine D. Wysham, MD, katherine.wysham@va.gov.

Submitted for publication February 3, 2025; accepted in revised form July 10, 2025.

SIGNIFICANCE & INNOVATIONS

- Rheumatoid arthritis (RA) disease activity had a stronger association with self-reported exhaustion, followed by weight loss and low hand grip strength—frailty features that can be affected by active RA disease.
- Disease activity increased in those whose frailty scores worsened at one year.
- We demonstrated an independent association between disease activity and frailty over time, supporting the concept of "secondary frailty" or frailty driven by an active underlying disease.

INTRODUCTION

Rheumatoid arthritis (RA) is an autoimmune disease characterized by joint pain, stiffness, fatigue, and chronic inflammation.[1] Sustained RA disease activity may lead to joint destruction and disability, negatively impacting overall well-being and quality of life.[2] Frailty is a state of decreased physiologic reserve, which places the individual at increased risk of poor health outcomes. Frailty can be measured using a deficit accumulation approach or through direct measurement of phenotypic frailty.[3,4] Phenotypic frailty is based on reduced physical function, exhaustion, and weight loss, many of which can be linked to RA disease activity and damage.[5–7] Compared to the general population of older adults, individuals with RA have an earlier onset and greater prevalence of frailty.[8] Additionally, frailty in RA has been found to be associated with adverse health outcomes, including reduced physical function, fractures, hospitalization, and mortality.[3,9] As the number of aging adults with RA continues to grow, improving our understanding of mechanisms contributing to frailty and interventions targeting it are needed.[3]

Although cross-sectional associations have been previously shown,[10–14] given the fluctuating nature of RA disease activity over time, longitudinal studies are also needed to investigate whether changes in disease activity impact frailty over time. However, studies evaluating this dynamic relationship are scarce and currently limited to early RA, in which disease activity is more amenable to treatment. Adding further complexity to this relationship is the concept of "secondary frailty," which occurs when an underlying condition is undertreated or uncontrolled and drives the clinical presentation of frailty[15,16] or acutely worsens frailty status.[15,17] In RA, some of the physiologic impacts of active disease are included in frailty assessment, suggesting that these concepts are intertwined and that RA disease might exacerbate aspects of frailty, including hand grip strength and fatigue.

The aim of this study is to evaluate the temporal relationship between disease activity and frailty in RA. We hypothesized that RA disease activity would be associated with frailty, independent of known risk factors, and participants who experienced higher disease activity over time would demonstrate a greater frailty burden, supporting the concept of secondary frailty.

METHODS

Participants. Data were from the Department of Veterans Affairs (VA) Rheumatoid Arthritis Frailty and Osteoporosis (VA-FROst) study, which is an ongoing longitudinal single-center cohort study of veterans with RA. Participants were enrolled from August 2022 to October 2024, and study visits were conducted annually by a rheumatologist-led research team (Supplementary Figure 1). Two waves of data were analyzed: baseline data (N = 132) and the group of participants who had reached their follow-up visit window by the time of this analysis (n = 65). The remaining participants had not yet reached their follow-up window and thus were not yet able to contribute follow-up data (n = 59). Additionally, between baseline and follow-up, two participants withdrew, two participants passed away, and four moved away and could not complete in-person follow-up. Participants were at least 35 years old and met 2010 American College of Rheumatology/EULAR classification criteria.[18] Exclusion criteria were concurrent diagnosis of another inflammatory arthritis, diseases that impact bone density (gastric bypass, hyperparathyroidism, end-stage renal disease, malabsorption, inflammatory bowel disease), and diseases that impact muscle function (multiple sclerosis, Parkinson disease). Pregnant individuals and people with bilateral hip replacements were excluded. Individuals taking sex hormone replacement (testosterone or estrogen), androgen deprivation therapy, systemic chemotherapy, anabolic medications for osteoporosis, or prednisone (>10 mg/day) were excluded. Those with weight >450 pounds, life expectancy less than two years, or unable to provide informed consent were excluded. All participants provided written informed consent. The study was approved by the VA Puget Sound Health Care System Institutional Review Board (no. 1872726).

Primary outcome measure. The primary outcome measure was frailty, which was measured at baseline and one-year follow-up and assessed using the validated Fried Frailty Phenotype (FFP).[4] This frailty measure has been well validated in the general population as a clinical presentation of frailty[4,19] and has also been studied in RA.[20–22] It includes measures of walk speed, hand grip strength, weight loss, exhaustion, and physical activity. Fifteen-foot walking speed and hand grip strength were categorized based on established cutoffs.[4] Weight loss of >5% in the last year was ascertained at baseline by reviewing the medical record for weight one year ago and at follow-up by comparing against weight at baseline. Exhaustion was assessed through self-reported exhaustion using the Center for Epidemiologic Studies Depression Scale,[23] and physical activity was queried using the International Physical Activity Questionnaire.[24] One point was assigned for each FFP component that met the predetermined

criteria, which produces FFP scores ranging from 0 to 5. Frailty was categorized as robust (0), prefrail (1–2), or frail (3–5) based on previously established score cutoffs.[15] Detailed definitions and cutoffs for each frailty component are described in Supplementary Table 1. Incident frailty at follow-up was not investigated due to being underpowered (n = 7).

Primary predictor. Physician-measured RA disease activity was measured at baseline and one-year follow-up with the Disease Activity Score in 28 joints using the C-reactive protein level (DAS28-CRP) (range 0–9.4).[25] The DAS28-CRP was calculated from the following subcomponents: tender joint count (TJC) (0–28), swollen joint count (SJC) (0–28), patient global assessment (PtGA) (0–10), and CRP level (0–100 mg/dL). DAS28-CRP categories were defined as follows: <2.6, remission; 2.6 to <3.2, low RA disease activity; 3.2 to ≤5.1, moderate RA disease activity; and >5.1, high RA disease activity.[25]

Other measures. Sociodemographic data were collected at baseline by patient self-report and included age, sex, race, ethnicity, years of education, and smoking history. Anti–cyclic citrullinated peptide level was obtained from the medical record. Radiographs of the hands and feet were reviewed by a physician rheumatologist to evaluate for erosions and joint space narrowing and were recorded as a binary yes or no. RA disease duration was assessed using the RA diagnosis date collected via self-report. Body mass index (BMI) was calculated using height and weight data collected during the baseline study visit. Prednisone use was collected by self-report. Conventional synthetic disease-modifying antirheumatic drug (csDMARD) and biologic and targeted synthetic DMARD (b/tsDMARD) use was collected from the medical record. csDMARDs (hydroxychloroquine, leflunomide, methotrexate, and sulfasalazine) and b/tsDMARDs (abatacept, adalimumab, etanercept, golimumab, infliximab, rituximab, tocilizumab, and tofacitinib) were recorded as a binary yes or no for active use at the time of the baseline or follow-up study visit, respectively. Comorbidity data were collected by chart review of the participants' medical problems lists at the time of the study visit. The comorbidities included were those that have demonstrated the strongest relationship with frailty in the literature: coronary artery disease, cerebrovascular accident, hypertension, peripheral vascular disease, lung disease, diabetes, and depression.[2,26]

Statistical analysis. Descriptive statistics were used to assess cohort characteristics, including sociodemographic characteristics, RA disease characteristics, and FFP subcomponents of the whole cohort by frailty category. Frequencies and percentages were reported for categorical and dichotomous variables. Means and SDs were reported for continuous variables. The participants who completed one-year follow-up visits were similarly described.

Unadjusted and adjusted ordinal logistic regression models were used to assess the relationship between RA disease activity (and each subcomponent) and frailty category (robust, prefrail, and frail) at baseline. Models were adjusted for age, sex, disease duration, prednisone use (dose in milligrams), csDMARD use, and b/tsDMARD use. These models were repeated to examine the association in veterans who completed the one-year follow-up. We tested the assumption of the proportion odds using the parallel lines proportion test. To assess which components of the FFP were most impacted by disease activity, we performed separate univariable logistic regression models to evaluate the association between baseline RA disease activity and the presence of each FFP subcomponent (low hand grip strength, slow walking, exhaustion, low physical activity, and weight loss).

Exploratory analyses. To assess the longitudinal relationship of RA disease activity with frailty category from baseline to one-year follow-up, we evaluated whether RA disease activity was associated with change in FFP score. FFP change scores (0–5) were calculated as the difference between FFP at follow-up and FFP at baseline. Based on FFP change scores, participants were categorized as having either the same or an improved frailty score (FFP score at one year ≤ FFP score at baseline) or a worse frailty score (FFP score at one year > FFP score at baseline). DAS28-CRP change scores were calculated as the difference between the mean DAS28-CRP at follow-up and the mean DAS28-CRP at baseline. Change scores were similarly calculated for PtGA, TJC, SJC, and CRP. Paired *t*-tests and Cohen's *d* values were used to test the difference and evaluate the effect size between the mean change scores.

The longitudinal relationship was also examined using univariable and multivariable mixed ordinal logistic regressions, in which the three-level frailty variable is the outcome (robust, prefrail, frail). Models were adjusted for age category, sex, RA disease duration, prednisone use (dose in milligrams), csDMARD and b/tsDMARD use, and time to follow-up. Models were repeated to test the association of each disease activity subcomponent with frailty category. The interaction between RA disease activity and time to follow-up (one year) was tested and not significant, so it was not included in the final model. The significance level for all analyses was set at 0.05, and STATA[27] version 18.0 was used to perform data analyses.

RESULTS

Cohort characteristics. A total of 132 participants with a mean age of 64.2 ± 11.3 years were included (Table 1). The participants were majority male (n = 96, 73%), 69% were White, 11% were Black, and 12% reported multiple races. The cohort had moderate disease activity on average (mean ± SD score 3.9 ± 1.3), with a mean RA disease duration of 14.3 ± 13.3 years. Frailty was characterized as robust in 35 (27%) patients, as

Table 1. Participant demographics and RA disease characteristics at baseline, stratified by frailty category*

	Total (N = 132)	Robust (n = 35, 27%)	Prefrail (n = 77, 58%)	Frail (n = 20, 15%)
Sociodemographic				
Age, y	64.2 (11.3)	61.5 (10.1)	64.5 (12.2)	67.6 (9.0)
Male	96 (73%)	21 (60%)	57 (74%)	18 (90%)
Race				
Black	15 (11%)	7 (20%)	8 (10%)	0 (0%)
Other[a]	10 (8%)	4 (12%)	6 (8%)	0 (0%)
White	91 (69%)	20 (57%)	55 (71%)	16 (80%)
Multiple races	16 (12%)	4 (11%)	8 (10%)	4 (20%)
Hispanic/Latino	8 (6%)	3 (9%)	5 (7%)	0 (0%)
Schooling, y	14.7 (2.4)	15.2 (2.6)	14.4 (2.3)	15.3 (2.1)
Smoking, pack years	18.4 (27.3)	8.1 (13.2)	18.9 (28.4)	34.8 (33.4)
BMI	30.1 (5.3)	31.4 (5.4)	30.0 (5.2)	28.1 (5.4)
RA disease characteristics				
Anti-CCP positivity	84 (64%)	21 (60%)	52 (68%)	11 (55%)
Disease duration, y	14.3 (13.3)	12.4 (10.9)	15.5 (14.2)	13.0 (13.8)
Erosions	53 (40%)	15 (43%)	29 (38%)	9 (45%)
Prednisone use, mg/day	1.3 (3.4)	2.2 (4.7)	1.1 (3.0)	0.7 (1.7)
csDMARD use	87 (66%)	22 (63%)	50 (65%)	15 (75%)
b/tsDMARD use	63 (48%)	18 (51%)	38 (49%)	7 (35%)
DAS28-CRP (0–9.4)	3.9 (1.3)	3.3 (1.0)	4.0 (1.3)	4.7 (1.3)
Remission (<2.6)	24 (18%)	9 (26%)	13 (17%)	2 (10%)
Low (2.6 to <3.2)	18 (14%)	9 (26%)	8 (10%)	1 (5%)
Moderate (3.2 to ≤5.1)	67 (51%)	16 (46%)	41 (53%)	10 (50%)
High (>5.1)	23 (17%)	1 (3%)	15 (20%)	7 (35%)
RA disease activity subcomponents				
Patient global assessment (0–10)	3.7 (2.2)	2.6 (1.7)	4.1 (2.2)	4.1 (2.5)
Tender joint count (0–28)	3.7 (4.9)	1.9 (2.5)	3.8 (4.7)	6.7 (7.3)
Swollen joint count (0–28)	3.8 (3.9)	2.8 (2.7)	3.8 (4.0)	5.6 (4.4)
CRP, mg/dL	7.1 (10.4)	4.7 (7.1)	7.4 (11.5)	10.0 (10.5)
Fried Frailty Phenotype subcomponents				
Weight loss	18 (14%)	0 (0%)	10 (13%)	8 (40%)
Low strength	53 (40%)	0 (0%)	36 (47%)	17 (85%)
Exhaustion	68 (52%)	0 (0%)	50 (65%)	18 (90%)
Slow walking	10 (8%)	0 (0%)	3 (4%)	7 (35%)
Low physical activity	26 (20%)	0 (0%)	11 (14%)	15 (75%)
Comorbidities				
Coronary artery disease	14 (11%)	2 (6%)	8 (10%)	4 (20%)
Cerebrovascular accident	6 (5%)	1 (3%)	4 (5%)	1 (5%)
Hypertension	63 (48%)	14 (40%)	39 (51%)	10 (50%)
Peripheral vascular disease	5 (4%)	0 (0%)	4 (5%)	1 (5%)
Lung disease	30 (23%)	6 (17%)	20 (26%)	4 (20%)
Diabetes	26 (20%)	6 (17%)	19 (25%)	1 (5%)
Depression	37 (28%)	7 (20%)	24 (31%)	6 (30%)

* Values are reported as mean (SD) or n (%). b/tsDMARD, biologic/targeted synthetic disease-modifying antirheumatic drug; BMI, body mass index; CCP, cyclic citrullinated peptide; CRP, C-reactive protein; csDMARD, conventional synthetic DMARD; DAS28-CRP, Disease Activity Score in 28 joints using the CRP level; RA, rheumatoid arthritis.
[a] American Indian, Alaska Native, Pacific Islander, Asian, and other/unknown.

prefrail in 77 (58%) patients, and as frail in 20 (15%) patients. Frail veterans were older and had greater smoking history and higher RA disease activity scores compared to robust and prefrail veterans. Prevalence of important frailty-associated comorbidities was similar across the frailty categories. At the time of this analysis, 65 participants had completed both baseline and one-year visits, whereas the remaining baseline participants had not yet reached their follow-up visit window. Similar trends of sociodemographic and RA disease characteristics were observed in the cohort at one-year follow-up (Table 2). There were no clinically significant differences between those who contributed follow-up data and those who did not (Supplementary Table 2).

Cross-sectional analysis: RA disease activity, subcomponents of RA disease activity, and frailty. In both the unadjusted and adjusted models at baseline, higher RA disease activity was associated with a higher frailty category (odds ratio [OR] 1.79, 95% confidence interval [CI] 1.34–2.39, $P < 0.0001$; adjusted OR [aOR] 1.98, 95% CI 1.45–2.69, $P < 0.0001$) (Table 3). Higher values for TJC and SJC were also associated with greater odds of having a higher level of frailty at

Table 2. Participant demographics and RA disease characteristics at 1-year follow-up for those who contributed longitudinal data, stratified by frailty category*

	Total (N = 65)	Robust (n = 19, 29%)	Prefrail (n = 36, 55%)	Frail (n = 10, 15%)
Sociodemographic				
Age, y	65.2 (10.4)	64.2 (7.3)	64.4 (11.8)	70.0 (9.3)
Male	49 (75%)	12 (63%)	28 (78%)	9 (90%)
Race				
Black	6 (9%)	5 (26%)	1 (3%)	0 (0%)
Other[a]	3 (5%)	0 (0%)	3 (8%)	0 (0%)
White	48 (74%)	11 (58%)	28 (78%)	9 (90%)
Multiple races	8 (12%)	3 (16%)	4 (11%)	1 (10%)
Hispanic/Latino	4 (6%)	2 (11%)	2 (6%)	0 (0%)
Schooling, y	14.8 (2.4)	15.5 (2.8)	14.6 (2.2)	14.2 (1.7)
Smoking, pack years	16.0 (19.1)	13.2 (17.5)	14.9 (18.7)	25.5 (22.1)
BMI	30.5 (5.5)	30.2 (4.8)	31.4 (6.1)	28.2 (3.8)
RA disease characteristics				
Anti-CCP positivity	44 (68%)	12 (63%)	22 (61%)	10 (100%)
Disease duration, y	15.8 (14.1)	13.0 (9.9)	14.8 (13.5)	24.5 (20.0)
Erosions	25 (39%)	7 (37%)	14 (39%)	4 (40%)
Prednisone use, mg/day	0.8 (2.2)	0.4 (1.3)	1.1 (2.7)	0.7 (1.6)
csDMARD use	65 (100%)	19 (100%)	36 (100%)	10 (100%)
b/tsDMARD use	34 (52%)	9 (47%)	20 (56%)	5 (50%)
DAS28-CRP (0–9.4)	4.0 (1.1)	3.4 (0.9)	4.2 (1.1)	4.5 (1.1)
Remission (<2.6)	7 (11%)	3 (16%)	3 (8%)	1 (10%)
Low (2.6 to <3.2)	10 (15%)	6 (32%)	4 (11%)	0 (0%)
Moderate (3.2 to ≤5.1)	36 (55%)	10 (53%)	20 (56%)	6 (60%)
High (>5.1)	12 (19%)	0 (0%)	9 (25%)	3 (30%)
RA disease activity subcomponents				
Patient global assessment (0–10)	4.2 (2.3)	3.3 (2.6)	4.4 (2.1)	5.2 (1.7)
Tender joint count (0–28)	4.0 (5.0)	2.2 (2.4)	4.7 (5.7)	5.0 (5.5)
Swollen joint count (0–28)	4.1 (3.8)	3.0 (2.5)	4.4 (4.0)	4.9 (4.9)
CRP, mg/dL	5.1 (5.3)	3.3 (2.2)	5.8 (6.1)	6.4 (6.0)
Fried Frailty Phenotype subcomponents				
Weight loss	10 (15%)	0 (0.0%)	4 (11.1%)	6 (60.0%)
Low strength	21 (32%)	0 (0.0%)	13 (36.1%)	8 (80.0%)
Exhaustion	33 (51%)	0 (0.0%)	24 (66.7%)	9 (90.0%)
Slow walking	8 (12%)	0 (0.0%)	2 (5.6%)	6 (60.0%)
Low physical activity	15 (23%)	0 (0.0%)	10 (27.8%)	5 (50.0%)
Comorbidities				
Coronary artery disease	6 (9%)	3 (16%)	3 (8%)	0 (0%)
Cerebrovascular accident	2 (3%)	1 (5%)	0 (0%)	1 (10%)
Hypertension	33 (51%)	9 (47%)	20 (56%)	4 (40%)
Peripheral vascular disease	2 (3%)	0 (0%)	2 (6%)	0 (0%)
Lung disease	15 (23%)	4 (21%)	10 (28%)	1 (10%)
Diabetes	13 (20%)	5 (26%)	7 (19%)	1 (10%)
Depression	20 (31%)	3 (16%)	14 (39%)	3 (30%)

* Values are reported as mean (SD) or n (%). b/tsDMARD, biologic/targeted synthetic disease-modifying antirheumatic drug; BMI, body mass index; CCP, cyclic citrullinated peptide; CRP, C-reactive protein; csDMARD, conventional synthetic DMARD; DAS28-CRP, Disease Activity Score in 28 joints using the CRP level; RA, rheumatoid arthritis.
[a] American Indian, Alaska Native, Pacific Islander, Asian, and other/unknown.

baseline after adjustment for potential confounders (aOR 1.15, 95% CI 1.07–1.24, $P < 0.0001$; aOR 1.13, 95% CI 1.04–1.24, $P = 0.006$, respectively) (Table 3). PtGA also met significance; however, testing revealed violation of the proportional odds assumption. The association between increased RA disease activity and higher frailty category was also evident at one-year follow-up in those who contributed follow-up data (n = 65) (OR 1.88, 95% CI 1.19–2.98, $P = 0.007$; aOR 2.59, 95% CI 1.52–4.44, $P = 0.001$). Testing revealed there was no violation of the proportional odds assumption for any of the variables aside from PtGA. When examining the relationship between RA disease activity and the FFP components, higher disease activity was significantly associated with increased odds of low hand grip strength (OR 1.45, 95% CI 1.09–1.94, $P = 0.012$), exhaustion (OR 1.86, 95% CI 1.35–2.56, $P < 0.0001$), and weight loss (OR 1.57, 95% CI 1.05–2.34, $P = 0.027$) (Figure 1).

Exploratory analyses: Association of RA disease activity and frailty over time. Frailty scores were higher in 29% (19 of 65) of veterans after one year, suggesting worse frailty compared to baseline (Table 4). On the other hand, 71% (46 of 65) of veterans' frailty scores remained the same or improved. In

Table 3. Unadjusted and adjusted ordinal logistic regression models evaluating the association of DAS28-CRP and its subcomponents with frailty category of the whole cohort at baseline (N = 132)*

	Unadjusted model			Adjusted model		
	OR[a]	95% CI	*P* value	aOR[b]	95% CI	*P* value
DAS28-CRP	**1.79**	**1.34–2.39**	**<0.0001**	**1.98**	**1.45–2.69**	**<0.0001**
DAS28-CRP subcomponents						
Patient global assessment[c]	1.27	1.08–1.49	0.004	1.41	1.18–1.68	<0.0001
TJC	**1.13**	**1.05–1.22**	**0.001**	**1.15**	**1.07–1.24**	**<0.0001**
SJC	**1.12**	**1.02–1.22**	**0.014**	**1.13**	**1.04–1.24**	**0.006**
CRP	1.03	1.00–1.07	0.079	1.03	1.00–1.07	0.088

* Bold values indicate *P* value < 0.05. aOR, adjusted odds ratio; b/tsDMARD, biologic/targeted synthetic disease-modifying antirheumatic drug; CI, confidence interval; CRP, C-reactive protein; csDMARD, conventional synthetic DMARD; DAS28-CRP, Disease Activity Score in 28 joints using the CRP level; OR, odds ratio; SJC, swollen joint count; TJC, tender joint count.
[a] Values are ORs for greater degree of frailty per unit increase in DAS28-CRP or subcomponents.
[b] Models were adjusted for age category, sex, disease duration, prednisone use, csDMARD use, and b/DMARD use.
[c] Failed the proportional odds assumption.

the group who became more frail, the mean change in DAS28-CRP was an increase in 0.61 units at one-year follow-up (mean ± SD 0.61 ± 0.96, *P* = 0.0121). In participants whose frailty scores stayed the same or improved, change in CRP level on average was 1.50 mg/dL lower (mean ± SD −1.50 ± 4.43 mg/dL, *P* = 0.0266), but there was no significant change in overall DAS28-CRP in the same or improved frailty group.

Both unadjusted and adjusted mixed ordinal logistic regression models found higher RA disease activity to be independently associated with a greater frailty category over one year (OR 2.66, 95% CI 1.56–4.53, *P* < 0.0001; aOR 3.31, 95% CI 1.91–5.75, *P* < 0.0001) (Table 5). Higher values for all subcomponents of RA disease activity were also associated with increased odds of greater frailty over time (PtGA: aOR 1.41, *P* = 0.004; TJC: aOR 1.23, *P* = 0.001; SJC: aOR 1.22, *P* = 0.007; CRP: aOR 1.14, *P* = 0.015).

DISCUSSION

This study demonstrated significant, independent associations between frailty and RA disease activity in cross-section at both the time of enrollment and follow-up after one year. We also found that greater disease activity at baseline was associated with deficits in hand grip strength, exhaustion, and weight loss. Lastly, our longitudinal analyses showed that disease activity increased in those who had an increase in frailty. Additionally, we saw that higher disease activity was associated with greater frailty burden over time. These findings support the hypothesis that RA disease activity impacts frailty status.

There is a known association between RA disease activity and frailty, but most studies have been cross-sectional with limited covariates.[10–14] This is the first study to demonstrate a significant cross-sectional relationship between disease activity and frailty in a cohort of people with established RA disease

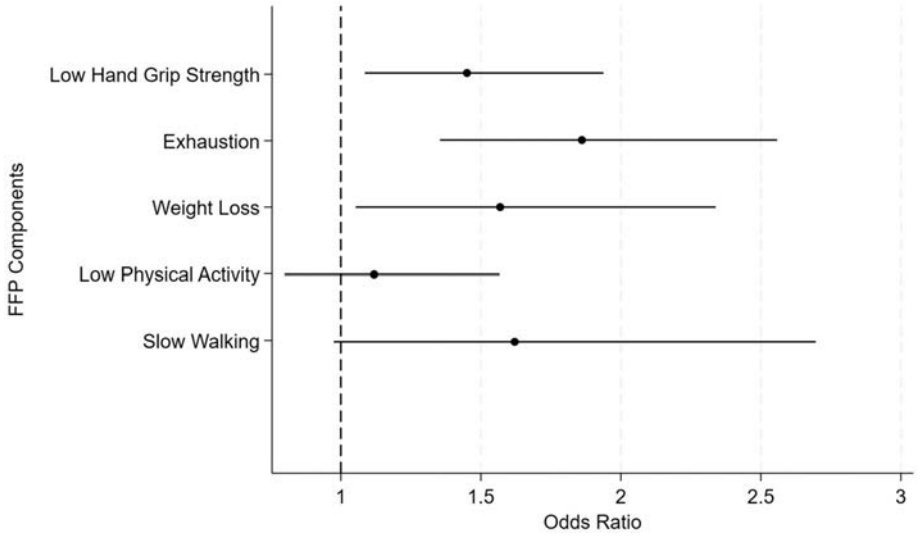

Figure 1. Univariable logistic regressions evaluating the associations between baseline DAS28-CRP and FFP components (N = 132). Odds ratios represent higher odds of having an FFP component deficit per unit increase in RA disease activity. DAS28-CRP, Disease Activity Score in 28 joints using the C-reactive protein level; FFP, Fried Frailty Phenotype; RA, rheumatoid arthritis.

Table 4. Difference between RA disease activity and subcomponents at baseline and after 1 year, stratified by those whose frailty scores worsened versus those whose frailty scores were the same/improved (N = 65)*

	Mean change score (SD)	*P* value	Cohen's *d*
Worse frailty score (n = 19)			
DAS28-CRP	**0.61 (0.96)**	**0.0121**	**0.53**[a]
Patient global assessment	1.28 (2.82)	0.0628	0.58[a]
Tender joint count	0.39 (3.06)	0.5806	0.08
Swollen joint count	1.00 (3.37)	0.2118	0.23
CRP	3.20 (7.58)	0.0823	0.50[a]
Same/improved frailty score (n = 46)			
DAS28-CRP	−0.05 (1.00)	0.7226	0.05
Patient global assessment	0.47 (2.28)	0.1715	0.20
Tender joint count	−0.07 (3.74)	0.9063	0.01
Swollen joint count	−0.04 (3.61)	0.9352	0.01
CRP	**−1.50 (4.43)**	**0.0266**	**0.36**

* Paired *t*-tests were used to analyze the change in disease activity scores and the individual components of disease activity at baseline and 1-year follow-up. These change scores were stratified by worse frailty score (FFP score at 1 year > FFP score at baseline) or same/improved frailty score (FFP score at 1 year ≤ FFP score at baseline). A larger Cohen's *d* value indicates a greater difference between the mean scores at baseline and the mean scores at 1-year follow-up. Small effect size *d* ≤ 0.50; medium effect size *d* 0.50 to <0.80; large effect size d ≥ 0.80. Significant values are bolded (*P* ≤ 0.05). CRP, C-reactive protein; DAS28-CRP, Disease Activity Score in 28 joints using the CRP level; FFP, Fried Frailty Phenotype; RA, rheumatoid arthritis.
[a] Medium or large effect size for *P* values at or nearing significance.

independent of important confounders, including RA disease duration and medication use. We found disease activity was higher in those whose frailty status worsened and that disease activity was independently associated with greater frailty at both time points. Increases in all disease activity subcomponent scores were also associated with worse frailty over time. Our study adds to the literature evaluating longitudinal relationships between RA disease activity and frailty. A study by Ohashi et al of prefrail patients with RA showed that higher disease activity by DAS28 predicted future frailty using an adapted FFP evaluating only the transition from prefrail to frail.[28] We added to these findings by assessing all three levels of frailty (robust, prefrail, and frail) and additionally analyzed the subcomponents of both disease activity and frailty, which highlighted areas that may be targets for intervention to mitigate frailty in RA. Another study by Hanlon et al[29] in a cohort of adults with early RA (RA disease duration of zero months [recruited at time of RA diagnosis]; median symptom duration of six months) demonstrated improvement in both mean frailty and disease activity scores at six months after enrollment. Their study constructed a deficit accumulation frailty index based on morbidities, baseline laboratory values, and functional measures rather than phenotypic frailty. Importantly, the Hanlon et al study may not be generalizable to people with established RA, who may not have as robust of a response to treatments and may have accrued damage from their underlying disease.[29] It is possible that frailty may worsen RA disease activity through increased pain and disability, impacting patient-reported outcomes and lesser use of potent immunosuppressive drugs. The impact of frailty on RA disease activity has been previously investigated in two separate studies. Hanlon et al found disease activity levels were higher in the mild and moderate/severe groups; however, in this study of early RA, disease activity improved in all

Table 5. Unadjusted and adjusted mixed ordinal logistic regression models evaluating the association of DAS28-CRP and its subcomponents with frailty category over 1 year (N = 65)*

	Unadjusted model			Adjusted model		
	OR[a]	95% CI	*P* value	aOR[b]	95% CI	*P* value
DAS28-CRP	2.66	1.56–4.53	<0.0001	3.31	1.91–5.75	<0.0001
DAS28-CRP subcomponents						
Patient global assessment[c]	1.30	1.04–1.64	0.023	1.41	1.12–1.78	0.004
TJC	1.19	1.06–1.35	0.004	1.23	1.09–1.39	0.001
SJC	1.17	1.02–1.35	0.027	1.22	1.06–1.41	0.007
CRP	1.13	1.02–1.25	0.024	1.14	1.02–1.26	0.015

* aOR, adjusted odds ratio; b/tsDMARD, biologic/targeted synthetic disease-modifying antirheumatic drug; CI, confidence interval; CRP, C-reactive protein; csDMARD, conventional synthetic DMARD; DAS28-CRP, Disease Activity Score in 28 joints using the CRP level; OR, odds ratio; SJC, swollen joint count; TJC, tender joint count.
[a] Values are ORs for greater degree of frailty per unit increase in DAS28-CRP or subcomponents over 1 year's time.
[b] Models were adjusted for age category, sex, disease duration, prednisone use, csDMARD use, b/tsDMARD use, and time.
[c] Failed the proportional odds assumption in cross-sectional model.

groups and to a greater degree in those who were most frail.[29] Additionally, Andrews et al found a trend toward worsening disease activity in those who were frail at baseline; however, the findings did not achieve statistical significance.[30] Although a bidirectional relationship between frailty and RA disease activity likely exists, we believe that RA disease primarily drives frailty through accumulation of inflammation and damage and is likely an ideal target for frailty prevention.[31]

Our findings show that worsening frailty is coupled with worsening disease activity over one year's time, suggesting that uncontrolled RA disease activity may have acute impacts on frailty ascertainment and confound the measurement of underlying physiologic frailty. We believe our findings support the concept of "secondary frailty," namely frailty caused by active underlying disease.[15] Further emphasizing the role of secondary frailty, we found that RA disease activity appears to have particularly strong relationships with exhaustion, weight loss, and low hand grip strength in our cohort, all of which can be directly impacted by RA disease.[6,7,32] This suggests that uncontrolled disease activity may impact the measurements that define phenotypic frailty (eg, exhaustion, weight loss, low hand grip strength), which in turn posits whether attaining low disease activity or remission could improve these frailty components that were merely clouded by the heightened disease activity, thereby returning the individual back to their background underlying physiologic frailty state.

Our study is one of the first to examine the components of frailty to evaluate how disease activity impacts them individually and highlights the overlapping nature of these two constructs. Given the differential impacts of RA disease seen on the components of phenotypic frailty, there could be elements of RA disease that are interfering with frailty ascertainment using these traditional components. However, not all frailty components that we found to be associated with RA disease activity are acutely impacted by RA disease, such as weight loss, which requires time to accrue. This suggests that the relationship between frailty and RA disease activity is more complex than just due to an exacerbation of RA disease. Additional studies may want to investigate this relationship further by using instruments that do not depend on measures such as hand grip strength or fatigue that are highly impacted by RA disease activity. Future work is needed to focus on improving RA disease activity and to determine if using treat-to-target methods of low disease activity or remission seen in recent RA treatment guidelines[33–35] can improve frailty and even prevent onset.

This study has important strengths. All disease activity measures were performed by a board-certified rheumatologist, our cohort is composed of people with both new and established RA, and, most importantly, we have robust and detailed clinical measurements repeated at two time points. Additionally, our cohort adds novelty because though it is a predominantly male cohort, males are historically underrepresented in RA research. This study is limited by the small sample size, and the nature of the study design requiring in-person visits may have resulted in a lower capture of severely frail individuals. Our frailty category distribution, however, is consistent with other cohorts.[2,30] Though we do evaluate the longitudinal relationship between disease activity and frailty, we are unable to determine directionality or if this relationship is causal. The majority of our cohort exhibited moderate disease activity, which may overestimate the prevalence of frailty in our cohort due to secondary frailty from active RA disease. Finally, our cohort was predominantly older, White, male veterans from one geographic region, which may limit generalizability. Future work should evaluate whether these findings hold true in a cohort with a higher prevalence of female participants.

In conclusion, this study demonstrates a longitudinal association between RA disease activity and frailty in a well-phenotyped RA cohort. Importantly, we found that higher disease activity was associated with greater frailty over time, suggesting that active RA disease may impact frailty status (eg, secondary frailty). Additional work is needed to further evaluate the concept of secondary frailty and determine if improving disease activity in adults with established RA can improve and prevent frailty.

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 Wysham 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. Figus FA, Piga M, Azzolin I, et al. Rheumatoid arthritis: extra-articular manifestations and comorbidities. Autoimmun Rev 2021;20(4):102776.
2. Cook MJ, Verstappen SMM, Lunt M, et al. Increased frailty in individuals with osteoarthritis and rheumatoid arthritis and the influence of comorbidity: an analysis of the UK Biobank cohort. Arthritis Care Res (Hoboken) 2022;74(12):1989–1996.
3. Kim DH, Rockwood K. Frailty in older adults. N Engl J Med 2024;391(6):538–548.
4. Fried LP, Tangen CM, Walston J, et al; Cardiovascular Health Study Collaborative Research Group. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001;56(3):M146–M156.
5. Baker JF, Sauer BC, Cannon GW, et al. Changes in body mass related to the initiation of disease-modifying therapies in rheumatoid arthritis. Arthritis Rheumatol 2016;68(8):1818–1827.
6. Matura LA, Malone S, Jaime-Lara R, et al. A systematic review of biological mechanisms of fatigue in chronic illness. Biol Res Nurs 2018;20(4):410–421.
7. Pourhassan M, Sieske L, Janssen G, et al. The impact of acute changes of inflammation on appetite and food intake among older hospitalised patients. Br J Nutr 2020;124(10):1069–1075.

8. Ozeki S, Takeuchi K, Yasuoka M, et al. Comparison of frailty associated factors between older adult patients with rheumatoid arthritis and community dwellers. Arch Gerontol Geriatr 2021;96:104455.
9. Allison R II, Assadzandi S, Adelman M. Frailty: evaluation and management. Am Fam Physician 2021;103(4):219–226.
10. Furuya T, Oh K, Ikari K, et al. Factors associated with frailty in Japanese patients with rheumatoid arthritis: results from the Institute of Rheumatology Rheumatoid Arthritis cohort study. Clin Rheumatol 2022;41(2):405–410.
11. Salaffi F, Di Carlo M, Farah S, et al. Prevalence of frailty and its associated factors in patients with rheumatoid arthritis: a cross-sectional analysis. Clin Rheumatol 2019;38(7):1823–1830.
12. Yoshii I, Kondo M. Clinical characteristics of frailty in Japanese rheumatoid arthritis patients. J Frailty Aging 2020;9(3):158–164.
13. Ohashi Y, Takahashi N, Sobue Y, et al. Associations of frailty with RA-ILD and poor control of disease activity in patients with rheumatoid arthritis: a multi-center retrospective observational study. J Orthop Sci 2024;29(6):1496–1502.
14. Haider S, Grabovac I, Berner C, et al. Frailty in seropositive rheumatoid arthritis patients of working age: a cross-sectional study. Clin Exp Rheumatol 2019;37(4):585–592.
15. Fried LP, Cohen AA, Xue QL, et al. The physical frailty syndrome as a transition from homeostatic symphony to cacophony. Nat Aging 2021;1(1):36–46.
16. Goldwater DS, Pinney SP. Frailty in advanced heart failure: a consequence of aging or a separate entity? Clin Med Insights Cardiol 2015;9(suppl 2):39–46.
17. Chang SS, Weiss CO, Xue QL, et al. Association between inflammatory-related disease burden and frailty: results from the Women's Health and Aging Studies (WHAS) I and II. Arch Gerontol Geriatr 2012;54(1):9–15.
18. Aletaha D, Neogi T, Silman AJ, et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 2010;62(9):2569–2581.
19. Bandeen-Roche K, Xue QL, Ferrucci L, et al. Phenotype of frailty: characterization in the women's health and aging studies. J Gerontol A Biol Sci Med Sci 2006;61(3):262–266.
20. Hanlon P, Morrison H, Morton F, et al. Frailty in people with rheumatoid arthritis: a systematic review of observational studies. Wellcome Open Res 2022;6:244.
21. Gao RC, Wu ZG, Wu ZZ, et al. Frailty in rheumatoid arthritis: a systematic review and meta-analysis. Joint Bone Spine 2022;89(4):105343.
22. Loecker CN, Schmaderer MS, Wysham KD, et al. Psychometric properties of frailty instruments in adults with rheumatoid arthritis: a systematic review. ACR Open Rheumatol 2024;6(2):91–102.
23. Radloff LS. The CES-D Scale: a self-report depression scale for research in the general population. Appl Psychol Meas 1977;1(3):385–401.
24. Craig CL, Marshall AL, Sjöström M, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003;35(8):1381–1395.
25. England BR, Tiong BK, Bergman MJ, et al. 2019 Update of the American College of Rheumatology recommended rheumatoid arthritis disease activity measures. Arthritis Care Res (Hoboken) 2019;71(12):1540–1555.
26. Hanlon P, Nicholl BI, Jani BD, et al. Frailty and pre-frailty in middle-aged and older adults and its association with multimorbidity and mortality: a prospective analysis of 493 737 UK Biobank participants. Lancet Public Health 2018;3(7):e323–e332.
27. Stata Statistical Software: Release 18. StataCorp; 2023.
28. Ohashi Y, Takahashi N, Sobue Y, et al. Disease activity at baseline is an independent predictor of frailty at one year in pre-frail patients with rheumatoid arthritis; a multicenter retrospective observational study. J Orthop Sci 2024;29(1):315–320.
29. Hanlon P, Morton F, Siebert S, et al. Frailty in rheumatoidrmdopen-2021-002111 arthritis and its relationship with disease activity, hospitalisation and mortality: a longitudinal analysis of the Scottish Early Rheumatoid Arthritis cohort and UK Biobank. RMD Open 2022;8(1):e002111.
30. Andrews JS, Trupin L, Wysham KD, et al. The impact of frailty on changes in physical function and disease activity among adults with rheumatoid arthritis. ACR Open Rheumatol 2019;1(6):366–372.
31. Lieber SB, Wysham KD, Sattui SE, et al. Frailty and rheumatic diseases: evidence to date and lessons learned. Lancet Rheumatol 2024;6(12):e881–e891.
32. Arab Alkabeya H, Hughes AM, Adams J. Factors associated with hand and upper arm functional disability in people with rheumatoid arthritis: a systematic review. Arthritis Care Res (Hoboken) 2019;71(11):1473–1481.
33. Smolen JS, Aletaha D, Bijlsma JW, et al; T2T Expert Committee. Treating rheumatoid arthritis to target: recommendations of an international task force. Ann Rheum Dis 2010;69(4):631–637.
34. Smolen JS, Landewé RBM, Bergstra SA, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2022 update. Ann Rheum Dis 2023;82(1):3–18.
35. Fraenkel L, Bathon JM, England BR, et al. 2021 American College of Rheumatology guideline for the treatment of rheumatoid arthritis. Arthritis Rheumatol 2021;73(7):1108–1123.

Arthritis Care & Research
Vol. 78, No. 3, March 2026, pp 316–324
DOI 10.1002/acr.25632

American College of Rheumatology
Empowering Rheumatology Professionals

Adipokines and Associations With Incident Osteoporotic Fracture in Patients With Rheumatoid Arthritis

Joshua F. Baker,[1] Bryant R. England,[2] Michael D. George,[3] Hannah Brubeck,[4] Brian Sauer,[5] Aleksander Lenert,[6] Punyasha Roul,[2] Geoffrey M. Thiele,[2] Ted R. Mikuls,[2] and Katherine D. Wysham[4]

Objective. We assessed whether circulating adipokines are associated with incident fractures in patients with rheumatoid arthritis (RA).

Methods. Three adipokines (adiponectin, leptin, and fibroblast growth factor [FGF]-21) were measured using banked enrollment serum from participants in a longitudinal RA cohort. Adipokine levels were dichotomized as high/low using median values. Incident osteoporotic fracture was defined based on published algorithms using diagnostic codes and confirmed by chart review. Cox proportional hazard models evaluated adipokines and incident fracture risk adjusting for age, sex, race, smoking status, body mass index (BMI), prednisone use, disease activity, comorbidity score, calendar year, osteoporosis history, and previous fracture.

Results. A total of 2,527 participants were included (89% male, mean age 72 years). There were 228 incident fractures over 27,540 person-years of follow-up (8.3 fractures per 1,000 person-years). After adjustment, the risk of incident fracture was increased for high levels of leptin (hazard ratio [HR] = 1.47; 95% confidence interval [CI] 1.15–1.90; $P = 0.003$), FGF-21 [HR = 1.39; 95% CI 1.16–1.67; $P < 0.001$), and adiponectin (HR = 1.21; 95% CI 0.94–1.55), with the latter not achieving significance ($P = 0.13$). Participants who had elevated levels of all three adipokines experienced twice the risk of fracture compared with those in whom none was elevated (HR = 2.17; 95% CI 1.27–3.70; $P = 0.005$).

Conclusion. Elevations in adipokines are associated with an increased risk of fracture in patients with RA, independent of other established risk factors including BMI, smoking, and prednisone use. This supports further investigation to understand whether this association is related to altered body composition or disrupted metabolic pathways.

INTRODUCTION

People with rheumatoid arthritis (RA) are known to be at higher risk of osteoporosis and osteoporotic (OP) fracture.[1] The cause of this increased risk of fracture is multifactorial and thought to be related to the effects of chronic inflammation on bone, reduced physical activity, excess glucocorticoid use, and altered body composition.[2] Both low body mass index (BMI) and recent weight loss are considered risk factors for OP and fracture in a number of settings, including in RA.[2] In part, this relationship may be due to a loss of muscle mass that may directly contribute to bone loss through the loss of mechanical loading. Changes in body composition, including loss of muscle and excess adiposity, may also contribute to a higher risk of falls.[3] Finally, other metabolic derangements occurring in the context of systemic inflammation may have independent adverse effects on bone.[4]

Adipokines are fat and muscle-secreted protein hormones that function to regulate metabolism. Adipokines are closely tied to body composition abnormalities, most notably excess total and visceral adiposity. These measures therefore act as

The contents of this work do not represent the views of the Department of the Veterans Affairs or the United States Government.

Dr Baker's work was supported by Veterans Affairs (VA) Clinical Science Research & Development (CSR&D) Merit Award (grant I01 CX001703) and Rehabilitation Research & Development Merit Award (grant I01 RX003644). Dr England's work was supported by the VA CSR&D (grant IK2 CX002203), Rheumatology Research Foundation, and National Institute of General Medical Sciences, NIH (NIGMS; grant U54GM115458). Dr Mikuls's work was supported by VA Merit Award (grant I01 BX003635), US Department of Defense (grant PR200793), and NIGMS (grant U54GM115458). Dr Wysham's work was supported by CSR&D Career Development Award (grant IK2 CX002351).

[1]Joshua F. Baker, MD, MSCE: Corporal Michael J. Crescenz VA Medical Center and University of Pennsylvania, Philadelphia; [2]Bryant R. England, MD, PhD, Punyasha Roul, MS, Geoffrey M. Thiele, PhD, Ted R. Mikuls, MD, MSPH: VA Nebraska-Western Iowa Health Care System and University of Nebraska Medical Center, Omaha; [3]Michael D. George, MD, MSCE: University of Pennsylvania, Philadelphia, and VA Nebraska-Western Iowa Health Care System, Omaha; [4]Hannah Brubeck, BS, Katherine D. Wysham, MD: Puget Sound VA Health System and University of Washington, Seattle; [5]Brian Sauer, PhD: Salt Lake City VA Medical Center and University of Utah; [6]Aleksander Lenert, MD: VA Iowa City Health Care System, Iowa City, Iowa.

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.25632).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25632.

Address correspondence via email to Joshua F. Baker, MD, MSCE, at Joshua.baker@pennmedicine.upenn.edu.

Submitted for publication September 21, 2024; accepted in revised form July 22, 2025.

SIGNIFICANCE & INNOVATIONS

- Adipokines were associated with a greater prevalence of osteoporotic fracture and higher rates of incident fracture among patients with rheumatoid arthritis.
- Inclusion of adipokines into prediction models modestly, but significantly, improved prediction.
- High adiponectin levels were associated with fracture only among those with nonobese body mass index.

informative metabolic indicators and may render direct effects on both appetite and energy homeostasis. For example, adiponectin is observed to be higher in people who are thin and increases with rapid weight loss.[5,6] Adipokines may therefore serve as biomarkers of metabolic health and could help to predict adverse outcomes that are commonly observed in the setting of changes in energy homeostasis, weight loss, and fluctuations in weight. Adiponectin has been associated with fracture risk among men in the general population,[7–10] although its direct effects on bone are unclear.[11] Leptin, which is secreted by adipocytes and considered a satiety signal, has an unclear association with bone quality, with studies demonstrating conflicting evidence characterizing its relationship with bone density and fracture.[12–15] Another adipokine, fibroblast growth factor (FGF)-21, is also considered to be an indicator of metabolic stress, and higher levels have been associated with adverse bone morphology and reduced bone strength in other settings, such as among women with anorexia.[16]

Adipokines have been shown to be disrupted in patients with RA in association with disease activity[17] and have also been linked to poor outcomes. Elevated levels of circulating adiponectin, for example, have been associated with radiographic disease progression in RA.[18] We also previously demonstrated that FGF-21 was independently associated with unfavorable and deteriorating body composition as well as declining physical function in patients with RA.[19] Thus, adipokines may serve as a proxy for underlying metabolic derangements, and their measurement may provide prognostic information for a number of adverse outcomes among patients with RA. We are not aware of previous studies evaluating associations between adipokines and the long-term risk of fracture in patients with RA.

Understanding the link between metabolic dysregulation and fracture has the potential to inform novel therapies and improve long-term prediction of fracture. We aimed to determine if alterations in adipokine concentrations were independently associated with incident fracture in a cohort of patients with RA. We hypothesized that higher adiponectin, higher leptin, and higher FGF-21 levels would be associated with higher rates of incident fracture independent of BMI and other important health factors.

METHODS

Study setting. This is a prospective cohort study conducted in the Veterans Affairs Rheumatoid Arthritis (VARA) registry and biorepository. VARA is an ongoing national repository and multicenter RA registry that has been active for more than 20 years (initiated in 2003).[20–27] At the time this study was conducted, 13 VA sites had contributed data. Veterans who fulfill the 1987 American College of Rheumatology classification criteria for RA and are over 18 years of age are eligible for enrollment.[28] Rheumatology providers at each site record clinical data at enrollment and at follow-up visits as part of routine clinical care. Each individual site has institutional review board approval, and all study patients provided informed written consent.

Adipokine measurements. Three prehypothesized adipokines (total adiponectin, leptin, and FGF-21) were measured at enrollment on nonfasting samples stored at –80°C for up to 16 years using validated panels from Meso Scale Discovery (Rockville, MA), as previously described.[29,30] The assays were performed in accordance with previously defined protocols at the University of Nebraska Core Laboratory. Adipokine values were log-adjusted to fit a normal distribution and standardized so that a one-unit difference in the value represented a one SD difference for each analyte and the mean value was equal to zero for the population. Each individual adipokine was also dichotomized at the median value to define a high vs low serum concentration. An overall "high adipokine" score was calculated as the number of adipokines above the median value (range: 0–3).

Osteoporosis and fracture outcomes. A chart diagnosis of osteoporosis was defined by the presence of two administrative diagnosis codes over a one-year period preceding registry enrollment from the Corporate Data Warehouse (CDW) (VA administrative data) and categorized according to Healthcare Cost and Utilization Project–Clinical Classification Software (HCUP-CCS). The details of this classification can be found online.[31]

We initially identified potential fractures by querying inpatient and outpatient diagnoses within the medical record based on codes used in previously published algorithms in the VA.[32] All fractures identified using these diagnosis codes were adjudicated by a manual chart review of radiographs and medical notes from clinical providers. During chart review, we excluded fractures that were sustained during high-velocity trauma, periprosthetic fractures, fractures of face, fingers and toes, and pathologic fractures from cancer. Fractures occurring before enrollment in the registry were considered prevalent fractures. Only the first fracture occurring after enrollment in the registry was adjudicated, and follow-up time after the first fracture was censored in analyses.

Bone density assessments. Results of dual X-ray absorptiometry (DXA) testing were extracted from the registry database (entered by local investigators). Because entry of DXA results into the registry database is not uniform across all sites, we additionally queried CDW for additional results of DXA testing and extracted the corresponding radiology reports. The closest DXA result to the enrollment visit was manually reviewed for each participant, and the impression (normal, osteopenia, osteoporosis) was recorded. Bone density and T-score data were not uniformly available from these notes.

RA disease activity, serologies, markers of inflammation. Clinical joint counts (0 to 28) and patient/physician global scores were extracted from the registry. C-reactive protein (CRP) (mg/dL) levels were assessed using nephelometry (Siemens, Munich, Germany). Our primary disease activity measure was the Disease Activity Score in 28 Joints with CRP (DAS28[CRP]).[33] From banked serum, a second-generation commercial anticyclic citrullinated peptide antibodies (ACPA) were also assessed using a second-generation commercial assay (range 0–5 u/mL).[34]

Other covariables. Demographics and disease-specific characteristics were obtained from the VARA registry database. BMI was extracted from the vital sign packages available in the CDW, and the closest BMI value (within 60 days) to the enrollment date was used. BMI categories were defined as underweight (<20), normal weight (≥20–24.99), overweight (≥25–29.99), obese (≥30–34.99), and severely obese (≥35).[35]

We also defined the presence of metabolic syndrome based on the ATP III definition that requires three of the five following criteria to be present: elevated waist circumference (≥100 cm for men, ≥90 cm for women), low high-density lipoprotein, high triglycerides, evidence of insulin resistance, and high blood pressure. Because waist circumference was not available in this sample, we used previously validated algorithms to predict waist circumference from age, sex, and BMI, as previously described.[36] Insulin resistance was considered to be present if diabetes was present at enrollment or a Hemoglobin A1c value within 6 months of enrollment was greater than 6%.

Comorbidity was assessed at enrollment using the rheumatic disease comorbidity index (RDCI) with one point removed in the presence of fracture.[37] Individual comorbidities were extracted over a one-year period preceding enrollment from CDW and categorized according to HCUP-CCS.[30]

RA treatments were extracted from VA pharmacy databases. Participants were considered to be exposed to the therapy at enrollment if the enrollment visit occurred during a defined medication course.[23] Active glucocorticoid use was considered to be present at enrollment if a course of prednisone (not including dose-packs) overlapped within 30 days of the enrollment visit (positive predictive value = 85%).[29,30] Previous use and recent use (within 1 year of enrollment) of a bisphosphonate therapy was extracted from pharmacy databases.

Statistical analysis. Characteristics of the study population were described among participants who sustained or did not sustain an OP fracture in follow-up. Missing values for DAS28(CRP) were imputed by either replacing it with the DAS28 with erythrocyte sedimentation rate (N = 183) or through conditional mean imputation (N = 242, 9.6%). Missing values for baseline BMI were also imputed (5.1%) using conditional mean imputation. Logistic regression was used to evaluate the cross-sectional associations between individual adipokines (above/below the median; per 1 SD) with the presence of a history of osteoporosis (based on diagnostic codes and based on DXA impression) or OP fracture before enrollment after adjusting for age, sex, race, BMI, smoking status (current, former, never), DAS28, calendar year, comorbidity score, and the use of prednisone.

The primary analyses used Cox proportional hazards models to assess associations between adipokine measures and the time to incident fracture among all participants, clustering on study site and censoring at the first fracture, end of follow-up, or death. Hypothesized confounders that were explored in the analysis included baseline demographics, BMI, smoking status, RA disease duration, RA disease activity, ACPA status (positive vs negative), RDCI, history of osteoporosis by diagnosis codes, use of osteoporosis medications within a year of enrollment, history of previous OP fracture, and calendar year. Analytes were examined separately and then together in a single model. We explored further adjustment for DXA impression at enrollment (ie, normal, osteopenia, osteoporosis). We also explored several prehypothesized interactions with age, sex, and BMI.

The proportional hazards assumption was tested by visualizing the Shoenfeld residuals for each adipokine. There was no evidence of violation of the proportional hazards assumption. The analysis was conducted with Stata v.18.0 (College Station, TX) within the VA Informatics and Computing Infrastructure.

RESULTS

A total of 2,527 registry participants were evaluated at baseline. Enrollment characteristics of participants who did and did not experience fracture during follow-up are shown in Table 1. At enrollment, 158 (6.3%) participants had two diagnosis codes for osteoporosis and 81 (3.2%) had experienced a previous OP fracture. In the overall sample, the median (interquartile range [IQR]) values for adiponectin, leptin, and FGF-21 were 12.7 (6.1–25.3) ug/mL, 10.2 (3.8–23.9) ng/mL, and 0.61 (0.33–1.14) ng/mL, respectively.

Higher levels of leptin were observed among those with a diagnosis of osteoporosis before enrollment after adjusting for age, sex, race, BMI, comorbidity score, smoking status, calendar

Table 1. Participant characteristics at enrollment of those who did and did not develop a fracture in follow-up*

Patient characteristics	Fractured, N = 228	Did not fracture, N = 2,299	*P* value
Age, mean (SD), y	64.6 (9.7)	64.2 (11.0)	0.69
Female, n (%)	44 (19)	242 (10)	<0.001
White, n (%)	184 (81)	1,765 (91)	0.18
Black, n (%)	24 (11)	374 (16)	0.02
Smoking, n (%)			
Current	55 (24)	452 (20)	0.07
Former	108 (47)	1,229 (53)	
BMI, mean (SD)	28.6 (6.4)	28.8 (5.7)	0.61
RDCI, mean (SD)	3.70 (2.0)	3.34 (2.0)	0.01
DAS28, mean (SD)	3.78 (1.37)	3.59 (1.45)	0.07
ACPA positive, n (%)	169 (74)	1,658 (72)	0.44
Disease duration, median (IQR), y	9.0 (2.8–20.8)	7.5 (2.3–16.6)	0.07
Methotrexate, n (%)	115 (50)	1,195 (52)	0.66
Glucocorticoids, n (%)	39 (45)	828 (34)	0.04
TNFi therapy, n (%)	55 (24)	584 (25)	0.67
Enrolled <2010, n (%)	132 (58)	1,144 (50)	0.02
Osteoporosis, n (%)	15 (7)	143 (6)	0.83
Previous OP fracture, n (%)	18 (8)	63 (3)	<0.001
Previous bisphosphonate, n (%)	96 (42)	612 (27)	<0.001
Diabetes, n (%)	79 (35)	755 (33)	0.58
Hypertension, n (%)	157 (69)	1,576 (69)	0.92
Heart failure, n (%)	26 (11)	230 (10)	0.50
Osteoarthritis, n (%)	191 (84)	1,840 (80)	0.18
COPD/asthma, n (%)	88 (39)	728 (32)	0.03
DXA impression, n (%)			<0.001
Normal	31 (14)	598 (26)	
Osteopenia	91 (40)	784 (34)	
Osteoporosis	56 (25)	284 (12)	
Not done/missing	50 (22)	633 (28)	
Adipokines, median (IQR)			
Adiponectin, ug/mL	14.8 (7.0–31.6)	12.5 (6.0–25.0)	0.02
Leptin, ng/mL	12.3 (4.5–28.1)	10.0 (3.8–23.3)	0.03
FGF-21, pg/mL	721 (418–1,283)	598 (318–1,130)	0.009
Follow-up time, y	5.4 (2.6–8.7)	12.2 (6.8–15.6)	<0.001

* ACPA, anticyclic citrullinated peptide antibody; BMI, body mass index; COPD, chronic obstructive pulmonary disease; DAS28, disease activity score in 28 joints; DXA, dual energy absorptiometry; FGF-21, fibroblast growth factor 21; IQR, interquartile range; OP, osteoporotic; RDCI, rheumatic disease comorbidity index; TNFi, tumor necrosis factor inhibitor.

year, DAS28, and use of prednisone (Figure 1). In addition, all adipokine levels were higher among patients with a history of OP fracture compared with those that did not have a history of previous fracture before and after adjustment (Table 1 and Figure 1).

In cross-sectional analyses, accounting for covariables, adiponectin, and FGF-21 levels above the median were associated with higher odds of previous/prevalent OP fracture at baseline (Table 2). Likewise, participants in whom all three adipokines were above the median had greater odds of a previous fracture compared with those in whom none was elevated (odds ratio [OR] = 2.71; 95% confidence interval (CI) 1.14–6.45; $P = 0.03$). There were no significant associations between adipokines and DXA impression at baseline after adjustment (Supplementary Table 1).

In longitudinal assessments, there were 228 participants with RA who developed an OP fracture over 27,540 person-years of follow-up, which is a rate of 8.3 fractures per 1,000 person-years between 2003 and 2020. The median time to fracture was 5.4 years (IQR 2.6–8.6 years) and the range of follow-up time in the overall sample was 42 days to 17.4 years. Patients who experienced a fracture during follow-up were older and more likely to use glucocorticoids, be female, have a diagnosis of osteoporosis, and to have lower rates of normal bone density on DXA (Table 1).

Higher levels of FGF-21 and leptin (above the median) were associated with an increased risk of incident fracture in adjusted models. For example, levels of leptin above the median were associated with a 47% higher risk of fracture [hazard ratio (HR): 1.47; 95% CI 1.15–1.90; $P = 0.003$). Leptin and FGF-21 were both associated with fracture independent of each other (Table 3). Adiponectin levels above the median were not significantly associated with fracture risk (HR 1.21; 95% CI 0.94–1.55; $P = 0.13$).

Higher adipokine scores (range: 0–4) were associated with an increased fracture risk in a dose-dependent manner (Figure 2) (P for trend <0.001). A high adipokine score of three (all adipokines elevated) (N = 393) was associated with a more than two-fold increase in risk compared with a score of zero (none elevated) (N = 371) (HR 2.17; 95% CI 1.27–3.70; $P = 0.005$)

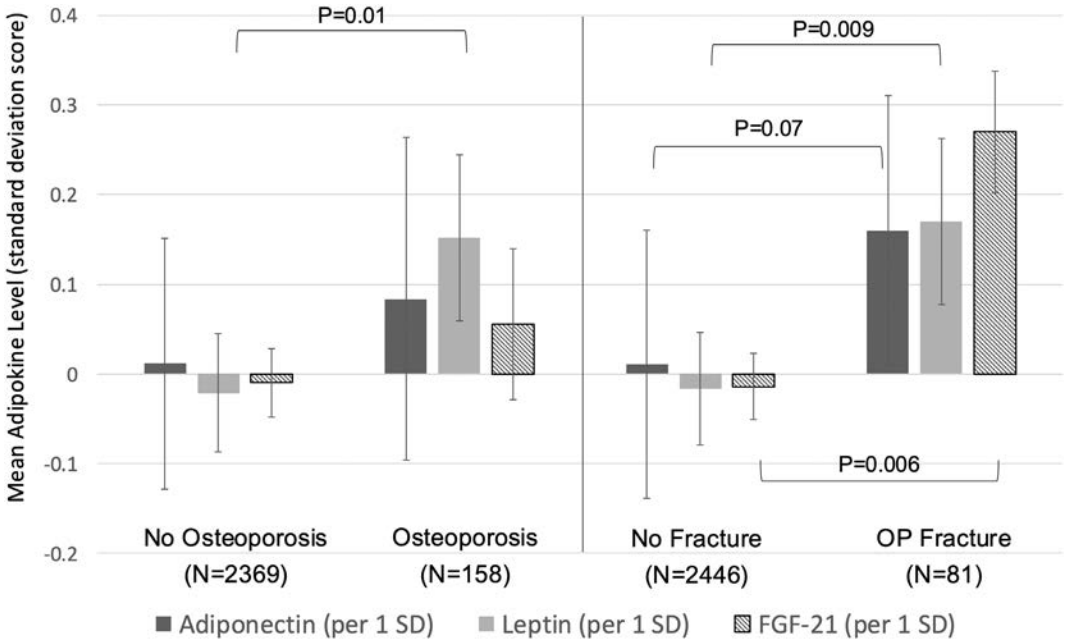

Figure 1. Adipokine levels by diagnosis of osteoporosis (two codes in last year) or history of previous osteoporotic fracture, at enrollment adjusted for age, sex, race, BMI, DAS28, smoking, RDCI, and prednisone use, and calendar date of enrollment. Adjusted for age, sex, race, BMI, DAS28, smoking status, RDCI, prednisone use, and calendar date. BMI, body mass index; DAS28, disease activity score in 28 joints; OP, osteoporotic; RDCI, rheumatic disease comorbidity index.

(Figure 2). Adding the adipokine score to the model improved the overall prediction of the model (Harrell's C-statistic 0.678 vs 0.664; $P = 0.01$).

Adipokines were not significantly associated with DXA impression at enrollment (Supplementary Table 1). Models further adjusting for DXA impression at enrollment also did not significantly affect the point estimates for adipokines (Supplementary Table 2). Associations between adipokines and incident fracture modeled as continuous variables or by quartile are shown in Supplementary Table 3. The presence of metabolic syndrome was not associated with fracture (HR 0.98; 95% CI 0.84–1.15; $P = 0.81$), and adding it to the model did not affect estimates for adipokines (not shown).

There was a significant interaction between adiponectin and BMI such that a higher level of adiponectin was strongly associated with fracture among those with a nonobese BMI (HR for <30: 1.52; 95% CI 1.09–2.12; $P < 0.001$; P for interaction <0.001) (Supplementary Table 4). There were no significant interactions between BMI and leptin or FGF-21. There were no significant interactions with age, although high adiponectin levels tended to be associated with fracture among adults over 65 years of age (HR 1.62; 95% CI 0.94–2.80; $P = 0.08$).

Table 2. Elevations in adipokines and associations with prevalent osteoporosis and fracture at enrollment adjusting for age, sex, race, BMI, prednisone use, disease activity, and smoking status*

Adipokine	Baseline osteoporosis, n = 158/2527		Previous OP fracture, n = 83/2527	
	OR (95% CI)	*P* value	OR (95% CI)	*P* value
High adiponectin	1.03 (0.74–1.42)	0.88	1.75 (1.10–2.79)	0.02
High leptin	1.36 (0.94–1.98)	0.10	1.22 (0.73–2.05)	0.44
High FGF-21	1.15 (0.83–1.60)	0.41	1.73 (1.09–2.73)	0.02
Number elevated				
None	1 (reference)	–	1 (reference)	–
1 elevated	1.39 (0.78–2.46)	0.26	1.21 (0.53–2.76)	0.64
2 elevated	1.52 (0.78–2.96)	0.22	1.66 (0.74–3.73)	0.22
All elevated	1.45 (0.69–3.03)	0.33	2.71 (1.14–6.45)	0.03

* Adjusted for age, sex, race, BMI, DAS28, smoking, RDCI, and prednisone use, and calendar date of enrollment; the presence of prevalent osteoporosis is based on diagnostic codes present before registry enrollment. The RDCI was adjusted down one point for those with prevalent fracture. BMI, body mass index; CI, confidence interval; DAS28, disease activity score in 28 joints; FGF-21, fibroblast growth factor-21; OP, osteoporotic; OR, odds ratio; RDCI, rheumatic disease comorbidity index.

Table 3. Association between adipokines and incident fracture*

	Incident osteoporotic fracture (adipokines separately), N = 2,527, P-Y = 16,572; events = 228		Incident osteoporotic fracture (adipokines together), N = 2,527, P-Y = 16,572; events = 228	
Adipokine	HR (95% CI)	*P* value	HR (95% CI)	*P* value
High adiponectin	1.21 (0.94–1.55)	0.13	1.17 (0.91–1.50)	0.29
High leptin	1.47 (1.15–1.90)	0.003	1.34 (1.11–1.63)	0.003
High FGF-21	1.39 (1.16–1.67)	<0.001	1.39 (1.07–1.82)	0.02

* In the left column, the result represents the effect for the exposure in three separate models. In the right column, adipokines were tested in the same model. All models adjusted for age, age^2, sex, race, BMI, smoking, history of osteoporosis, history of fracture, previous use of bisphosphonates, DAS28, prednisone use, comorbidity score (RDCI), and date of enrollment. The RDCI was adjusted down one point for those with prevalent fracture. BMI, body mass index; CI, confidence interval; FGF-21, fibroblast growth factor-21; HR, hazard ratio; P-Y, person-years; RDCI, rheumatic disease comorbidity index.

(Supplementary Table 5). There were no significant interactions between adipokines with sex on fracture risk.

DISCUSSION

This study demonstrated associations between circulating adipokines and the risk of incident OP fracture among patients with RA. In particular, those with elevated levels of all measured adipokines had more than twice the risk of incident OP fracture. These observations were independent of a number of other risk factors, including demographics, BMI, smoking, glucocorticoid use, RA disease activity, and comorbidity, in addition to a previous diagnosis of osteoporosis, fracture, and bone density assessments. These findings indicate that metabolic assessments beyond clinically available measures may aid in OP fracture risk stratification in RA.

Associations between adipokines and OP fracture have been observed in other populations.[7] For example, in the Osteoporotic Fractures in Men study, the investigators found that higher adiponectin levels (per SD) were associated with a 30% higher risk of fracture in adjusted models.[10] In another study by Nakamura et al, high adiponectin levels were associated with fracture among postmenopausal women.[12] Although we did not find a significant association between adiponectin and fracture overall, there was a significant association among patients who were nonobese in this study, which is a finding that is largely consistent with these previous studies. The current study may suggest that associations are stronger in older and thinner patients who may

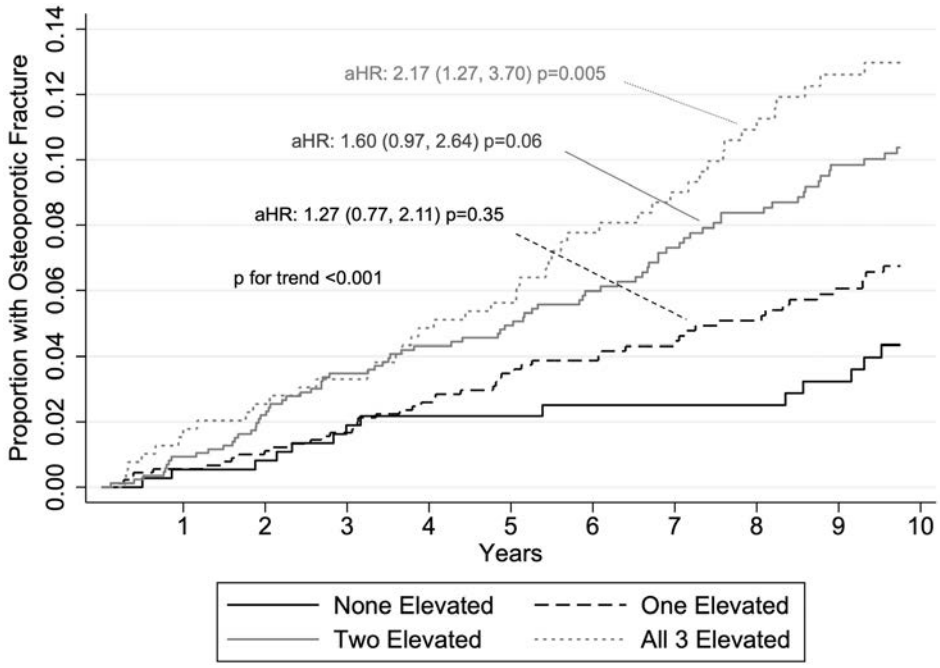

Figure 2. Kaplan-Meier curve demonstrating the proportion at risk experiencing an osteoporotic fracture over time according to the number of elevated adipokines. Models are adjusted for age, sex, race, calendar year, BMI, smoking, osteoporosis at enrollment, previous osteoporotic fracture, current prednisone use, DAS28, RDCI, and previous bisphosphonate use. aHR, adjusted hazard ratio; BMI, body mass index; DAS28, disease activity score in 28 joints; RDCI, rheumatic disease comorbidity index.

be more likely to suffer from sarcopenia or cachexia, as has previously been suggested.[38]

Nakamura et al observed that higher levels of leptin were associated with a lower rate of fracture among postmenopausal women.[12] Other studies have found no association between leptin and fracture risk.[13,39] In addition, although some studies have identified positive relationships between leptin and bone mineral density,[14] others have found inverse associations after considering adiposity.[15] We are not aware of previous studies evaluating this question in RA, which is a disease that is associated with metabolic disturbances, weight loss, sarcopenia, and heightened fracture risk. These differences across studies imply that associations between leptin and FGF-21 with fracture outcomes may be expected to vary based on the population under study and the definition for fracture used. Leptin and FGF-21 are strongly associated with obesity, and obesity has been increasingly recognized to be a risk factor for fracture, particularly among patients that are at risk for sarcopenic obesity due to functional impairment and a greater risk of falls.[40–43]

Adipokines are metabolic regulators that play a role in maintaining energy balance and may help predict outcomes by serving as a proxy for metabolic derangements. During chronic inflammatory disease states, large amounts of energy are expended by the activated immune system. Therefore chronic inflammation can disrupt normal energy homeostasis and result in fatigue, cachexia, anemia, poor physical function, insulin resistance, and bone loss, among other consequences.[4] We have previously shown that adipokines are associated with muscle loss in patients with RA, which is an important risk factor in the development of low bone density and leads to an increased risk of falls and osteoporotic fracture.[44,45] Osteoblasts and osteoclasts do express adiponectin receptors, and some studies have suggested that adiponectin may have direct effects on osteoblasts and osteoclasts in vitro.[46] However, the role of adiponectin in bone remodeling and repair is poorly understood.[11] Further study is needed to determine if adipokines directly impact bone quality, thereby serving as potential therapeutic targets in conditions in which metabolic disruptions are often observed, as in patients with inflammatory conditions such as RA.

Adipokines have been hypothesized to be associated with fracture risk through an association with metabolic obesity and metabolic syndrome. The relationship between metabolic obesity and bone health is complex and not fully understood; however, one study demonstrated higher rates of fracture among those with high waist circumference.[47] Although obesity is often observed to be protective against fracture, the metabolic complications of obesity may be deleterious.[48] In our study, we found no evidence of an association between metabolic syndrome and fracture, and the association between adipokines and fracture was independent of the presence of metabolic syndrome. Thus, we did not find evidence that adipokines increase the risk of fracture through their role as markers of metabolic obesity in this population. Further study in this area may be needed.

Despite adipokines being associated with significantly increased risks of incident OP fracture, they offered only modest prognostic value. Thus, although the current study suggests that the assessment of metabolic factors may be important in understanding risks of fracture in this population, biomarkers with greater prognostic value may be needed to add significantly to clinical care. Our findings that associations of adipokines with fracture risk were independent of baseline DXA impressions suggests that peripheral biomarkers could potentially aid in risk prediction above other imaging biomarkers in common use. However, additional study is needed to generate a potential tool from these markers, generate optimal cut-points for clinical use, and to quantify the potential benefits and the key subgroups for which its use might be most valuable. It may also be possible to combine adipokine assessments with other biomarkers to generate a more accurate and clinically-valuable tool.

Our study includes a high proportion of men with RA from a population of US veterans and may not be completely generalizable to other RA populations, particularly since women are at higher risk of osteoporosis and fracture over their lifetimes. One previous study did find that the effect of adiponectin on fracture risk, for example, was stronger for men compared with women.[7] Although there were no significant interactions by biologic sex in our study, this does not rule out clinically meaningful differences in association, and these observations should be validated in populations with a greater proportion of women. Although we adjusted for previous diagnosis of osteoporosis, DXA findings, and fracture history, there may be residual confounding that limits our ability to make direct causal inferences; although the results do suggest a prognostic value beyond other information typically available in clinic. It is also worth noting that fractures that may have been cared for outside of the VA might have been missed, particularly if that care was not billed to the VA. However, many of these fractures may also be captured by subsequent routine outpatient VA care.[49] We had to impute some covariates because of missing data (eg, disease activity) within the registry, which may have limited our ability to fully adjust for these covariates. Finally, the data available did not provide an opportunity to directly compare prediction with the FRAX calculator; thus, it remains unclear whether adipokines provide value above and beyond this risk-prediction tool. Strengths of the study include the large sample of at-risk patients, comprehensively assessed adipokine measures and defined thresholds, thorough adjustment for comorbidity and clinical disease activity, as well as the adjudicated fracture outcome.

In summary, elevated levels of the adipokines leptin and FGF-21 are independent predictors of incident OP fracture among patients with RA. Further study is needed to understand the clinical utility of the assessment of adipokines in the clinical

management of patients with RA, as well as to clarify the mechanisms underlying this association.

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 Baker 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. Wright NC, Lisse JR, Walitt BT, et al; Women's Health Initiative Investigators. Arthritis increases the risk for fractures--results from the Women's Health Initiative. J Rheumatol 2011;38(8):1680–1688.
2. van Staa TP, Geusens P, Bijlsma JW, et al. Clinical assessment of the long-term risk of fracture in patients with rheumatoid arthritis. Arthritis Rheum 2006;54(10):3104–3112.
3. Yeung SSY, Reijnierse EM, Pham VK, et al. Sarcopenia and its association with falls and fractures in older adults: a systematic review and meta-analysis. J Cachexia Sarcopenia Muscle 2019;10(3):485–500.
4. Spies CM, Straub RH, Buttgereit F. Energy metabolism and rheumatic diseases: from cell to organism. Arthritis Res Ther 2012;14(3):216.
5. Araújo JP, Lourenço P, Rocha-Gonçalves F, et al. Adiponectin is increased in cardiac cachexia irrespective of body mass index. Eur J Heart Fail 2009;11(6):567–572.
6. Behre CJ. Adiponectin: saving the starved and the overfed. Med Hypotheses 2007;69(6):1290–1292.
7. Barbour KE, Zmuda JM, Boudreau R, et al. Adipokines and the risk of fracture in older adults. J Bone Miner Res 2011;26(7):1568–1576.
8. Barbour KE, Zmuda JM, Boudreau R, et al; Health ABC Study. The effects of adiponectin and leptin on changes in bone mineral density. Osteoporos Int 2012;23(6):1699–1710.
9. Johansson H, Odén A, Karlsson MK, et al. Waning predictive value of serum adiponectin for fracture risk in elderly men: MrOS Sweden. Osteoporos Int 2014;25(7):1831–1836.
10. Johansson H, Odén A, Lerner UH, et al. High serum adiponectin predicts incident fractures in elderly men: Osteoporotic Fractures in Men (MrOS) Sweden. J Bone Miner Res 2012;27(6):1390–1396.
11. Lewis JW, Edwards JR, Naylor AJ, et al. Adiponectin signalling in bone homeostasis, with age and in disease. Bone Res 2021;9(1):1.
12. Nakamura Y, Nakano M, Suzuki T, et al. Two adipocytokines, leptin and adiponectin, independently predict osteoporotic fracture risk at different bone sites in postmenopausal women. Bone 2020;137:115404.
13. Chen LW, Chen FP, Hsieh CW, et al. Analysis of the associations among Helicobacter pylori infection, adiponectin, leptin, and 10-year fracture risk using the fracture risk assessment tool: A cross-sectional community-based study. PLoS One 2017;12(4):e0175365.
14. Wang CH, Lai YH, Lin YL, et al. Increased serum leptin level predicts bone mineral density in hemodialysis patients. Int J Endocrinol 2020;2020:8451751.
15. Di Monaco M, Vallero F, Di Monaco R, et al. Fat body mass, leptin and femur bone mineral density in hip-fractured women. J Endocrinol Invest 2003;26(12):1180–1185.
16. Fazeli PK, Faje AT, Cross EJ, et al. Serum FGF-21 levels are associated with worsened radial trabecular bone microarchitecture and decreased radial bone strength in women with anorexia nervosa. Bone 2015;77:6–11.
17. Magali Chamorro-Melo Y, Calixto OJ, Bello-Gualtero JM, et al. Evaluation of the adipokine profile (adiponectin, resistin, adipsin, vaspin, and leptin) in patients with early rheumatoid arthritis and its correlation with disease activity. Reumatologia 2022;60(3):192–199.
18. Giles JT, van der Heijde DM, Bathon JM. Association of circulating adiponectin levels with progression of radiographic joint destruction in rheumatoid arthritis. Ann Rheum Dis 2011;70(9):1562–1568.
19. Gould PW, Zemel BS, Taratuta EG, et al. Circulating fibroblast growth factor-21 levels in rheumatoid arthritis: associations with disease characteristics, body composition, and physical functioning. J Rheumatol 2021;48(4):504–512.
20. Davis LA, Cannon GW, Pointer LF, et al. Cardiovascular events are not associated with MTHFR polymorphisms, but are associated with methotrexate use and traditional risk factors in US veterans with rheumatoid arthritis. J Rheumatol 2013;40(6):809–817
21. Richards JS, Cannon GW, Hayden CL, et al. Adherence with bisphosphonate therapy in US veterans with rheumatoid arthritis. Arthritis Care Res (Hoboken) 2012;64(12):1864–1870.
22. Curtis JR, Baddley JW, Yang S, et al. Derivation and preliminary validation of an administrative claims-based algorithm for the effectiveness of medications for rheumatoid arthritis. Arthritis Res Ther 2011;13(5):R155.
23. Cannon GW, Mikuls TR, Hayden CL, et al. Merging Veterans Affairs rheumatoid arthritis registry and pharmacy data to assess methotrexate adherence and disease activity in clinical practice. Arthritis Care Res (Hoboken) 2011;63(12):1680–1690.
24. Mikuls TR, Gould KA, Bynoté KK, et al. Anticitrullinated protein antibody (ACPA) in rheumatoid arthritis: influence of an interaction between HLA-DRB1 shared epitope and a deletion polymorphism in glutathione S-transferase in a cross-sectional study. Arthritis Res Ther. 2010;12(6):R213.
25. Curtis JR, Jain A, Askling J, et al. A comparison of patient characteristics and outcomes in selected European and U.S. rheumatoid arthritis registries. Semin Arthritis Rheum 2010;40(1):2–14.e1.
26. Mikuls TR, Fay BT, Michaud K, et al. Associations of disease activity and treatments with mortality in men with rheumatoid arthritis: results from the VARA registry. Rheumatology (Oxford) 2011;50(1):101–109.
27. Mikuls TR, Padala PR, Sayles HR, et al. Prospective study of posttraumatic stress disorder and disease activity outcomes in US veterans with rheumatoid arthritis. Arthritis Care Res (Hoboken) 2013;65(2):227–234.
28. 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):315–24. doi: 10.1002/art.1780310302
29. Baker JF, England BR, George MD, et al. Adipocytokines and achievement of low disease activity in rheumatoid arthritis. Semin Arthritis Rheum 2022;55:152003.
30. Baker JF, England BR, George MD, et al. Elevations in adipocytokines and mortality in rheumatoid arthritis. Rheumatology (Oxford) 2022;61(12):4924–4934.
31. Clinical Classifications Software Refined. Healthcare Cost and Utilization Project (HCUP). Agency for Healthcare Research and Quality. Updated 2021. Accessed August 24, 2023. https://www.hcup-us.ahrq.gov/toolssoftware/ccsr/ccs_refined.jsp
32. Colón-Emeric C, Pieper CF, Grubber J, et al. Correlation of hip fracture with other fracture types: toward a rational composite hip fracture endpoint. Bone 2015;81:67–71.
33. Prevoo ML, 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.

34. Miriovsky BJ, Michaud K, Thiele GM, et al. Anti-CCP antibody and rheumatoid factor concentrations predict greater disease activity in men with rheumatoid arthritis. Ann Rheum Dis 2010;69(7):1292–1297.
35. Baker JF, Billig E, Michaud K, et al. Weight loss, the obesity paradox, and the risk of death in rheumatoid arthritis. Arthritis Rheumatol 2015;67(7):1711–1717.
36. Bozeman SR, Hoaglin DC, Burton TM, et al. Predicting waist circumference from body mass index. BMC Med Res Methodol 2012;12(1):115.
37. England BR, Sayles H, Mikuls TR, et al. Validation of the rheumatic disease comorbidity index. Arthritis Care Res (Hoboken) 2015;67(6):865–872.
38. Baker JF, Newman AB, Kanaya A, et al. The adiponectin paradox in the elderly: associations with body composition, physical functioning, and mortality. J Gerontol A Biol Sci Med Sci 2019;74(2):247–253.
39. Baker JF, Conaghan PG, Emery P, et al. Relationship of patient-reported outcomes with MRI measures in rheumatoid arthritis. Ann Rheum Dis 2017;76(3):486–490.
40. Liu HF, Meng DF, Yu P, et al. Obesity and risk of fracture in postmenopausal women: a meta-analysis of cohort studies. Ann Med 2023;55(1):2203515.
41. Halpern SD, French B, Small DS, et al. Randomized trial of four financial-incentive programs for smoking cessation. N Engl J Med 2015;372(22):2108–2117.
42. Compston JE, Watts NB, Chapurlat R, et al; Glow Investigators. Obesity is not protective against fracture in postmenopausal women: GLOW. Am J Med 2011;124(11):1043–1050.
43. Baker JF, Giles JT, Weber D, et al. Sarcopenic obesity in rheumatoid arthritis: prevalence and impact on physical functioning. Rheumatology (Oxford) 2022;61(6):2285–2294.
44. Baker JF, Davis M, Alexander R, et al. Associations between body composition and bone density and structure in men and women across the adult age spectrum. Bone 2013;53(1):34–41.
45. Baker JF, Von Feldt JM, Mostoufi-Moab S, et al. Insulin-like growth factor 1 and adiponectin and associations with muscle deficits, disease characteristics, and treatments in rheumatoid arthritis. J Rheumatol 2015;42(11):2038–2045.
46. Krumbholz G, Junker S, Meier FMP, et al. Response of human rheumatoid arthritis osteoblasts and osteoclasts to adiponectin. Clin Exp Rheumatol 2017;35(3):406–414.
47. Amouzegar A, Asgari S, Azizi F, et al. The role of metabolic syndrome and its components in incident fracture: a 15-year follow-up among the Iranian population. J Clin Endocrinol Metab 2021;106(5):e1968–e1983.
48. Chin KY, Ima-Nirwana S, Mohamed IN, et al. The association between bone health indicated by calcaneal quantitative ultrasound and metabolic syndrome in Malaysian men. J Diabetes Metab Disord 2015;14(1):9.
49. Schwab P, Sayles H, Bergman D, et al. Utilization of care outside the Veterans Affairs health care system by US veterans with rheumatoid arthritis. Arthritis Care Res (Hoboken) 2017;69(6):776–782.

Arthritis Care & Research
Vol. 78, No. 3, March 2026, pp 325–336
DOI 10.1002/acr.25627

American College of Rheumatology
Empowering Rheumatology Professionals

REVIEW ARTICLE

Economic Burden of Rheumatoid Arthritis in Low- and Middle-Income Countries: Systematic Review and Meta-Analysis

Tadesse Gebrye,[1] Chidozie E. Mbada,[1] Clara T. Fatoye,[1] Faatihah Niyi-Odumosu,[2] Ushotanefe Useh,[3] Zalmai Hakimi,[4] and Francis Fatoye[5,3]

Objective. The aim of this systematic review was to synthesize the economic impact of rheumatoid arthritis (RA) on households, health systems, and society in low- and middle-income countries (LMICs).

Methods. Electronic databases such as PubMed, Web of Science, and CINAHL were searched using keywords related to RA and cost of illness. Eligible studies were required to report RA-related costs, be conducted in LMICs, and be published in English. Quality appraisal of the included studies was conducted using the Newcastle–Ottawa Scale for cohort studies. A narrative synthesis and meta-analysis of findings was conducted.

Results. A total of 5,134 studies was initially identified for screening. After removing 1,028 duplicates, 50 studies were selected for full-text review, and 15 met the eligibility criteria and were therefore included in the review. These studies, published between 2007 and 2024, were conducted in various countries, including Turkey (n = 3), China (n = 2), and one study each from Thailand, Hungary, Mexico, Colombia, Morocco, Pakistan, India, Romania, Brazil, and Argentina. Nine studies adopted a societal perspective, whereas six used a health care perspective. The total sample size was 218,575 participants, with individual study sizes ranged from 62 to 209,292. Average annual direct costs per patient ranged from US$523 to US$2,837.90, and indirect costs ranged from US$81.80 to US$2,463.40. The pooled average annual costs for outpatients, inpatients, and medical costs were US$517.72 (95% confidence interval [CI] $3.35–$1,032.09), US$543.88 (95% CI US$499.51–US$588.24), and US$3,379.83 (95% CI US$3,137.58–US $3,622.08), respectively.

Conclusion. RA poses a significant economic challenge in LMICs, where limited health care resources and high treatment costs make care unaffordable for many. This review uniquely underscores that enhancing treatment access and optimizing resource use can reduce both medical and productivity losses, improving patient outcomes and strengthening economic resilience.

INTRODUCTION

Rheumatoid arthritis (RA) is a chronic autoimmune disease characterized by persistent inflammation primarily affecting the joints but also impacting multiple extra-articular organs such as the heart, kidneys, lungs, digestive system, and nervous system.[1] RA leads to debilitating symptoms, including pain, stiffness, fatigue, and reduced mobility, ultimately causing long-term disability and reduced quality of life.[2] RA is among the most prevalent chronic diseases worldwide, impacting around 1% of the global population. It typically affects children and young adults aged 16 to 40 years, with a higher prevalence observed in industrialized countries.[3] RA presents a significant societal burden due to its high rates of morbidity, disability, and economic costs.[4]

[1]Tadesse Gebrye, MSc, MPH, Chidozie E. Mbada, PhD, Clara T. Fatoye, MA: Manchester Metropolitan University, Manchester, United Kingdom; [2]Faatihah Niyi-Odumosu, PhD: University of the West of England, Bristol, United Kingdom; [3]Ushotanefe Useh, PhD, Francis Fatoye, PhD: North-West University, Potchefstroom, South Africa; [4]Zalmai Hakimi, PhD; Swedish Orphan Biovitrum AB, Stockholm, Sweden; [5]Francis Fatoye, PhD: Manchester Metropolitan University, Manchester, United Kingdom, and North-West University, Potchefstroom, South Africa.

Additional supplementary information cited in this article can be found online in the Supporting Information section (http://onlinelibrary.wiley.com/doi/10.1002/acr.25627).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25627.

Address correspondence via email to Tadesse Gebrye, MSc, MPH, at t.gebrye@mmu.ac.uk.

Submitted for publication April 4, 2025; accepted in revised form August 4, 2025.

SIGNIFICANCE & INNOVATIONS

- This study addresses the underexplored economic burden of rheumatoid arthritis (RA) in low- and middle-income countries (LMICs), offering a comprehensive understanding of its impact on households, health systems, and societies in resource-limited settings, in contrast to the focus of much existing research on high-income countries.
- The review adhered to Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines and conducted a rigorous quality appraisal using the Newcastle–Ottawa Scale, incorporating studies from various LMICs to provide a comprehensive perspective on RA's economic burden and highlight regional differences in costs and access to care.
- The review offers data to guide policy decisions aimed at improving access to affordable treatments, optimizing health care resources, and boosting health outcomes for individuals with RA in LMICs.

The economic impact of RA can be categorized into direct, indirect, and intangible costs.[5] Direct costs include expenses related to medical treatments, hospitalizations, medications, and transportation for health care services.[6] Indirect costs encompass productivity losses from absenteeism or early retirement due to illness.[7] Psychosocial or intangible costs, though challenging to quantify, represent the deterioration in the quality of life for patients, families, and caregivers.[8] These components collectively form the economic burden of RA, which has been extensively studied in high-income countries (HICs) but remains poorly understood in low- and middle-income countries (LMICs).

In LMICs, the economic burden of RA is significant due to high treatment costs and loss of productivity. Limited health care resources and access to specialized care further exacerbate the financial strain on individuals and health care systems.[9] Moreover, RA contributes to a larger global health challenge given that noncommunicable diseases (NCDs) increasingly affect the working-age population in these countries.[10] The rising prevalence of NCDs in LMICs, alongside the continued challenge of communicable diseases, results in a double burden that hampers national development and exacerbates poverty. For patients with RA, the inability to work due to disability can have devastating economic consequences, particularly in resource-constrained settings in which there is minimal or no social support for individuals with disability.

Given the expanding burden of RA in LMICs, it is crucial to understand the economic impact of the disease in these regions. Cost-of-illness (COI) studies provide a comprehensive analysis of the financial burden by assessing the various cost components associated with the disease. These studies offer valuable insights into where the major costs lie, which can guide policy decisions, health care planning, and the allocation of resources for RA care and treatment. However, there is a paucity of data on the economic burden of RA in LMICs, which makes it challenging to fully appreciate the extent of the problem. This paper presents a systematic review of published COI studies on RA, focusing on the economic burden in LMICs. By highlighting the economic burden of RA, this review aims to support evidence-based policymaking that can reduce the impact of the condition and improve health care delivery in LMICs, ultimately benefiting both patients with RA and the society.

METHODS

A comprehensive systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, as outlined by Moher et al.[11] This approach ensured a rigorous and transparent methodology for identifying, evaluating, and synthesizing relevant studies. Additionally, the review was formally registered with PROSPERO (PROSPERO is an international online registry where researchers prospectively record protocols for systematic reviews to promote transparency and reduce duplication), the International Prospective Register of Systematic Reviews, under Centre for Reviews and Dissemination registration number 42024566816, to ensure proper oversight and prevent duplication of efforts.

Search strategy. The literature search for this systematic review was conducted using the PubMed, Web of Science, and CINAHL databases, with studies published from inception up until March 28, 2025, being included. The search terms used encompassed various combinations such as "burden of disease," "length of stay," "cost of illness," "burden of illness," "cost of disease," "cost of sickness," "disease cost," "economic burden of disease," "sickness cost," "rheumatoid arthritis," and "arthritis." In addition to these databases, hand searches were conducted by reviewing the references of the included studies. The search terms were cross-referenced with the Medical Subject Headings terms to ensure comprehensive coverage. All references were imported into Covidence. The search process was independently conducted by two reviewers (TG and FN) to minimize bias in the study selection and exclusion. Any disagreements were resolved through discussion with a third reviewer (CM).

Inclusion and exclusion criteria. The reviewers independently selected publications for inclusion in this review based on predefined eligibility criteria. Any discrepancies in selection were resolved through consensus among the reviewers. The inclusion criteria specified that eligible studies must involve retrospective or prospective study designs conducted in primary or secondary care settings. Additionally, only studies reporting costs related to RA in LMICs, with study populations comprising patients with RA, and published in English were considered.

Exclusion criteria encompassed studies unrelated to the economic consequences of RA, as well as personal papers, conference abstracts, case reports, letters, commentaries, editorials, review articles, and studies lacking sufficient data.

Data extraction and quality assessment. The study selection process is detailed in the PRISMA flow diagram (Figure 1). Data extraction was conducted independently by two reviewers (TG and FN) using a standardized Excel-based data extraction form. To ensure accuracy, a third reviewer (CM) verified the extracted data. Information extracted from the included studies encompassed author and year of publication, study setting, cost perspective, data sources, outcome measures, sample size, year of costing, and various cost categories such as inpatient, outpatient, drug/medical, nonmedical, direct, indirect, and total costs. Supplementary Table 1 presents the definitions and classifications of health care–related costs.[12] For cost-related data, the currency, cost year, and the reported mean or median total and RA-attributable costs were also recorded. Additionally, we extracted information on the epidemiologic approach used in each study. In COI analysis, this includes two main perspectives: the incidence-based approach, which estimates the lifetime costs of newly diagnosed cases to inform potential preventive savings, and the prevalence-based approach, which assesses the current total costs associated with all existing cases over a defined period.[13]

The quality of the included studies was assessed using the Newcastle–Ottawa Scale (NOS) by two reviewers (TG and FN). The NOS evaluates studies across nine criteria, categorized into three dimensions: selection of the study population, comparability of groups, and assessment of outcomes or exposures.[14] Studies were scored on a scale with a maximum of nine points, in which a score of ≥6 indicated high quality, a score between 3 and 6 indicated moderate quality, and a score of ≤3 indicated low quality. Any disagreements in scoring were resolved through consultation with a third reviewer (CM).

Data analysis. In this study, summary descriptive statistics were employed to characterize the study background and the types of costs associated with RA. Both direct and indirect costs related to RA were further examined through quantitative analysis. A cross-country cost comparison was conducted based on the methodologies used in the included studies. To facilitate this

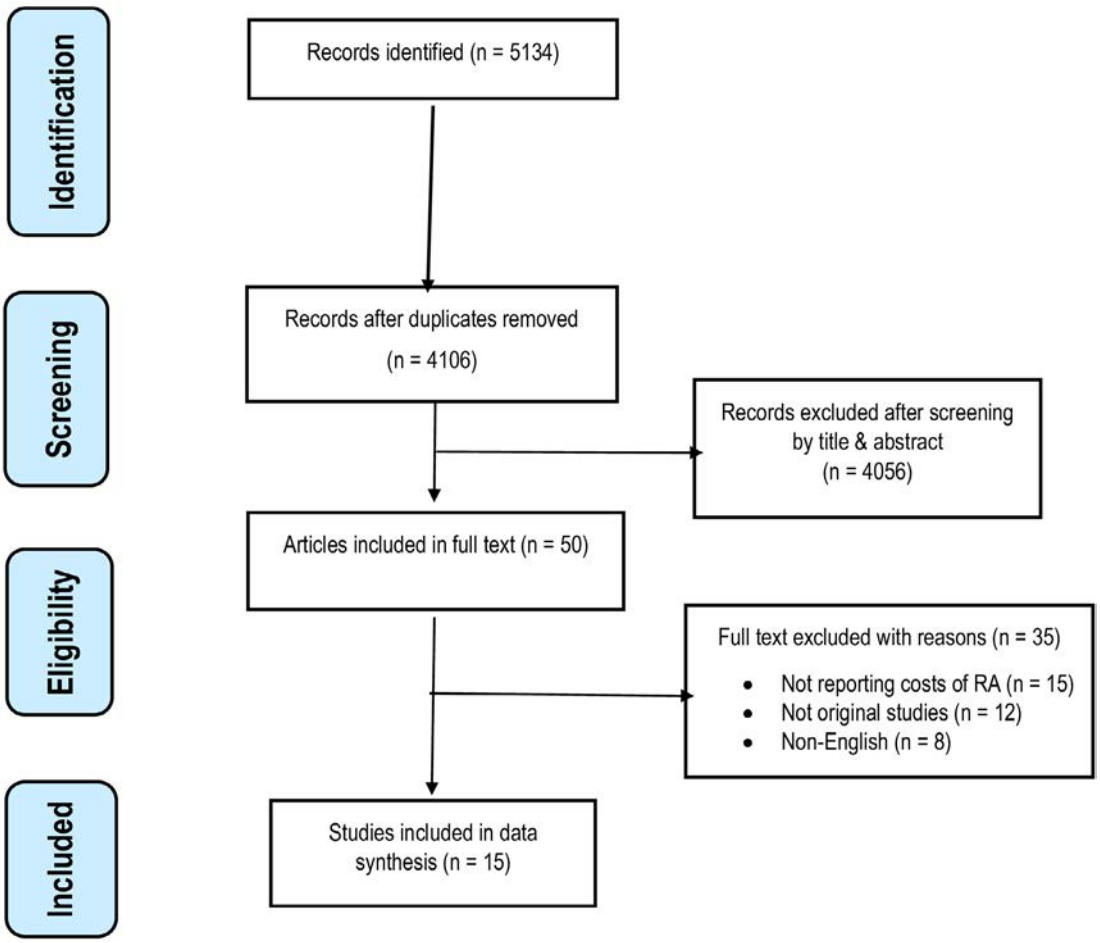

Figure 1. Flow diagram of publications included and excluded in the review. RA, rheumatoid arthritis.

comparison, costs were converted to US dollars (US$) using country-specific gross domestic product (GDP) deflators[15] and purchasing power parities (PPPs).[16] These conversions were performed as of December 2024. Estimated cost values were adjusted by multiplying them by the 2024 GDP coefficient, then dividing by the GDP of the reference year for each study and finally adjusted by the PPP conversion factor for 2024.

A meta-analysis was conducted to synthesize the cost estimates for RA from multiple studies. The analysis included studies reporting mean costs, SDs, and sample sizes. For each study, the SE was calculated using the formula $SE = SD/\sqrt{n}$, and 95% confidence intervals (CIs) were derived as mean ± 1.96 × SE. An inverse-variance weighted approach was applied to compute a pooled mean estimate, giving greater weight to studies with larger sample sizes and lower variance. Forest plots were generated to visually present individual study estimates alongside the pooled estimate, including corresponding 95% CIs.

RESULTS

A total of 5,134 studies were initially identified for screening, including 3,599 from Web of Science, 788 from CINAHL, and 747 from PubMed. After removing 1,028 duplicates, the remaining studies were screened by title and abstract, resulting in 50 studies selected for full-text review. Of these, 15 studies ultimately met the predefined eligibility criteria for inclusion. The majority of studies included in this systematic review were of moderate to high quality based on the NOS score (Supplementary Table 2).

Characteristics of the included studies. The characteristics of the included studies are presented in Table 1. The included studies estimated the COI of RA in 12 countries. These studies were published between 2007 and 2024. The total sample size across these studies was 218,575 participants; individual study sample sizes ranged from 62 to 209,292. The majority were conducted in Asia[17–21] followed by Europe,[22–26] Latin America,[27–30] and Africa.[31] Among the studies, female participants accounted for the main composition of the population, ranging from 42% to 92%. Of the 15 studies, 9 adopted a societal perspective, whereas 6 used a health care perspective.

In eight of the included studies[20,22–24,26,27,29,31] the data sources were retrospective databases including health insurance databases, disease registries, and hospital administrative records. The remaining studies described results from self-reported questionnaire surveys.[17–19,21,25,28,30] Most included studies[17–21,24–30] were conducted using a prevalence-based approach. The studies that adopted the incidence-based approach[22,23,31] focused on patients with recent onset. Cost components and measurement of direct or indirect costs also varied markedly due to the aims and data availability among studies. Nine of the included studies,[17,18,20,21,24–27] reported both direct and indirect costs, whereas the remaining studies reported only direct costs.

Medical and nonmedical costs of RA. Table 2 presents the medical and nonmedical costs of RA. In estimating the direct costs associated with RA, various cost components were incorporated into the analysis. These typically included drug-related expenses, inpatient care costs, outpatient care costs, and other health care–associated expenditures. Nine of the studies included in the analysis provided data on the annual medical costs, which ranged from US$1,728 to US$7,584 per patient per year.[16,17–19,21,24,26–29] Additionally, six studies reported the nonmedical costs of RA, which ranged from US$319 to US$7,988 per patient per year.[15,17,20–22,28,29] These figures highlight the substantial financial burden of RA, encompassing both direct medical costs and other related expenditures.

Meta-analysis. The meta-analysis of inpatient, outpatient, and medical costs for RA with 95% CI, including the pooled random-effects estimate, is presented (Figure 2). The pooled average annual costs for outpatients, inpatient, and medical costs were US$517.72 (95% CI US$3.35–US$1,032.09), US $543.88 (95% CI US$499.51–US$588.24), and US$3,379.83 (95% CI US$3,137.58–US$3,622.08), respectively.

Direct and indirect costs of RA. The direct and indirect costs associated with RA are summarized in Table 3. Ten of the studies included in the analysis reported the average annual direct costs of RA, which varied widely, ranging from US$523 to US $2,837.90.[17–24,28,31] In terms of indirect costs, seven studies provided estimates, with values spanning from US$90.59 to US $3,697 per patient per year.[25,26] Notably, one study[21] that assessed the annual direct and indirect costs of RA in China found that direct costs accounted for 90% of the total costs, whereas indirect costs made up the remaining 10%. This highlights the significant contribution of direct medical expenses to the overall economic burden of RA, though indirect costs, such as productivity loss, remain a notable cost driver.

DISCUSSION

This systematic review offers an original contribution by assessing the economic burden of RA in LMICs, a topic with limited existing synthesis. Drawing on 15 studies from 12 countries, the review reveals a substantial economic impact, with wide variation in direct and indirect costs influenced by study design, health care system differences, and data collection methods. The findings highlight critical patterns and gaps that are essential for policymakers, health care providers, and researchers working to address RA in LMICs.

Table 1. Characteristics of the included studies*

Author	Setting, year of costing	Method and sample description	Cost perspective	Source of data	Type of outcomes calculated	NOS score
Baser et al[23]	Turkey, 2013	Incidence and prevalence based, 2,613 (693 were incident cases and 1,920 prevalent cases)	Health care perspective	The Turkish National Health Insurance Database (2009–2011)	• Annual costs for incident and patients with prevalent RA after controlling for age, gender, region, comorbid conditions, and medication • Total health care costs	6
Ayan et al[24]	Turkey, 2018	Prevalence based, 62 (82.2 women)	Societal perspective	Multicenter study assessing RA-related health care costs in Turkey that was performed between May 2011 and August 2012	Direct and indirect costs	6
Hamuryudan et al[26]	Turkey, 2011	Prevalence based, 689	Societal perspective	The rheumatology outpatient clinics of 10 tertiary care hospitals in Turkey	Direct and indirect costs of RA	7
Horváth et al[22]	Hungary, 2012	Incidence based, 209,292 patients (225,623 cases)	Payer's perspective	National Health Insurance Fund Administration	Number of patients and cases, crude and weighted hospital days, and health insurance expenditure	7
Codreanu et al[25]	Romania, 2014	Prevalence-based/self-reported questionnaires, 206, mean age 54.9 ± 12.7 years	Societal perspective	The electronic database of a university tertiary rheumatology center	Indirect costs	7
Fellous et al[31]	Morocco, 2020	Incidence based, 197 (86.8% women)	Health care perspective	The Moroccan registry of biologic therapies in RA	Direct medical costs	7
Osiri et al[17]	Thailand, 2007	Prevalence based, 158 (>90% women)	Societal perspective	From 158 patients with RA who attended a major tertiary care facility in Bangkok, Thailand	Direct medical, direct nonmedical, indirect, and total costs	6
Hu et al[18]	China, 2013	Prevalence based, 133 (70.68% women)	Societal perspective	The COI questionnaire	Direct, indirect, and intangible costs	6
Xu et al[21]	China, 2009	Prevalence based, 829 (78.6% women)	Societal perspective	Data collected by trained physicians using a face-to-face interview with patients	Direct and indirect costs of RA	6
Naqvi et al[19]	Pakistan, 2019	Prevalence based, 358 (73.7% women)	Health care perspective	Rheumatology departments of two public sectors and three private sector tertiary care hospitals	Direct costs	5
Bali and Singla,[20]	India, 2022	Prevalence based/cross-sectional, 67 (89.5% women)	Societal perspective	Data collected over a two-month period from June to July 2022 from patients attending the weekly rheumatology OPD	Direct and indirect costs	5

(Continued)

Table 1. (*Cont'd*)

Author	Setting, year of costing	Method and sample description	Cost perspective	Source of data	Type of outcomes calculated	NOS score
Mendoza-Gutierrez et al[30]	Mexico, 2016 and 2017	Prevalence based, in 2016 and 2017, a total of 3,623 (2,944 women [81.3%] and 3,427 2,779 women [81.1%], respectively)	Payer's perspective	Mexican Social Security Institute	Direct economic costs	5
Santos-Moreno et al[27]	Colombia, 2019	Prevalence based, 83 patients (42% women)	Health care perspective	A rheumatology care center in Bogotá	Direct medical costs	6
Chermont et al[28]	Brazil, 2002	Prevalence based, 100 (92% women)	Societal perspective	The outpatient clinics for RA at the Division of Rheumatology of the Federal University of São Paulo	Direct and indirect costs	6
Catay et al[29]	Argentina, 2002	Prevalence based, 165 (84 women)	Societal perspective	Health management organization administrative and medical databases and questionnaire	Direct and indirect costs	7

* Incidence-based COI estimates the lifetime costs associated with new cases of a disease within a specific time period. Prevalence-based COI estimates the total costs of all existing cases of a disease within a defined population and time period. COI, cost of illness; NOS, Newcastle–Ottawa Scale; OPD, outpatient department; RA, rheumatoid arthritis.

Table 2. Medical and nonmedical costs of RA*

Author, country, year of costing	Inpatient costs, mean (±SD), inflated 2024 US$	Outpatient costs, mean (±SD), inflated 2024 US$	Medical costs, mean (±SD), inflated 2024 US$	Nonmedical costs, mean (±SD), inflated 2024 US$
Baser et al,[23] Turkey, 2013	US$335 (±US$451) for prevalent cases and US$480 (±US$646) incident cases per year	US$478.64 (±US$645) for prevalent and $547.61 (±US $737) for incident cases per year	NR	NR
Ayan et al,[24] Turkey, 2018	US$145,969 (±US$182.30), 95% CI US$122.04 (±US$152.30)–US $410,389 (±US$512.20) per patient per year	US$26,3223 (±US$32.90), 95% CI US$13.20 (±US$16.50)–US $40.68 (±US$50.80) per patient per year	US$3,539.16 (±US$4,418), 95% CI US$120.80 (±US $150.70)–US$4,361.13 (±US$5,444) per patient per year	NR
Hamuryudan et al,[26] Turkey, 2011	Mean ± SD US$195 (US$272), US $857 (US$1,194) per patient per year	Mean ± SD US$382 (US$532), US $428 (US$597) per patient per year	Mean ± SD US$2,777 (US $3,872), US$3,379 (US $4,711) per patient per year	NR
Horváth et al,[22] Hungary, 2012	NR	NR	NR	Day care: US$1.18 million (±US $1.60 million) Nursing: US$6.56 million (±US $8.86 million) Long-term care: US$30.15 million (±US$40.79 million) Rehabilitation: US$108.81 million (±US$147.08 million)
Osiri et al,[17] Thailand, 2007	NR	NR	Mean US$1,923 (±US $2,878), median US $1,036 (US$1,550) per patient per year	Mean US$213 (±US$ 319), median US$89 (US$ 133) per patient per year
Hu et al,[18] China, 2013	Mean US$593.03 (±US$798), US $1,220.35 (±US$1,641) per patient per year	Mean US$1,324.18 (±US$1,781), US$1,990.17 (±US$2,677) per patient per year	Mean US$1,283.89 (±US $1,728), US$1,898.15 (±US$2,553) per patient per year	NR
Xu et al,[21] China, 2009	NR	NR	Mean US$3,211 (±US $4,695), US$5,264 (±US $7,692) per patient per year	Mean US$232 (±US$339), US $662 (±US$967) per patient per year
Naqvi et al,[19] Pakistan, 2019	NR	NR	Medicines: US$63.25 (±US $73.79) per patient per year Medical devices costs: US $49.13 (±US$57.23) per patient per year Rheumatologist visits: US $72.05 (±US$83.97) per patient per year Physical therapy sessions: US$419.07 (±US$488.87) for per patient per year	NR
	NR	NR	NR	

(Continued)

Table 2. (*Cont'd*)

Author, country, year of costing	Inpatient costs, mean (±SD), inflated 2024 US$	Outpatient costs, mean (±SD), inflated 2024 US$	Medical costs, mean (±SD), inflated 2024 US$	Nonmedical costs, mean (±SD), inflated 2024 US$
Bali and Singla,[26] India, 2022				Out-of-pocket expenses, mean (±SD): US$230 (±US$259.67)
Mendoza-Gutierrez et al,[30] Mexico, 2016 and 2017	US$9,096,245.67 (±US $12,280,932.64) for 2016 and US$8,932,592.57 (±US $11,612,370.34) for 2017 (total cost per year)	NR	NR	NR
Santos-Moreno et al,[27] Colombia, 2019	NR	NR	Total = Mean (±SD) US $6,067 (±US$7,584), US $6,144 (±US$7,680) per patients per year Male: Mean (± SD) US $6,160 (±US$7,700), US $5,959 (±US$7,449) Female: Mean (±SD) US $5,893 (±US$7,366), US $6,545 (±US$8,181)	NR
Chermont et al,[28] Brazil, 2002	NR	NR	US$370.36 (±US$1,259.22) per patient per year	US$32.68 (±US$111.11) per patient per year
Catay et al,[29] Argentina, 2002	NR	NR	US$1,862 (±US$6,735)…US $67,032 (±US$242,460) per patient per year	US$221.90 (±US$315)…US $7,988 (±SD US$11,340) per patient per year

* CI, confidence interval; NR, not reported; RA, rheumatoid arthritis; US$, US dollars.

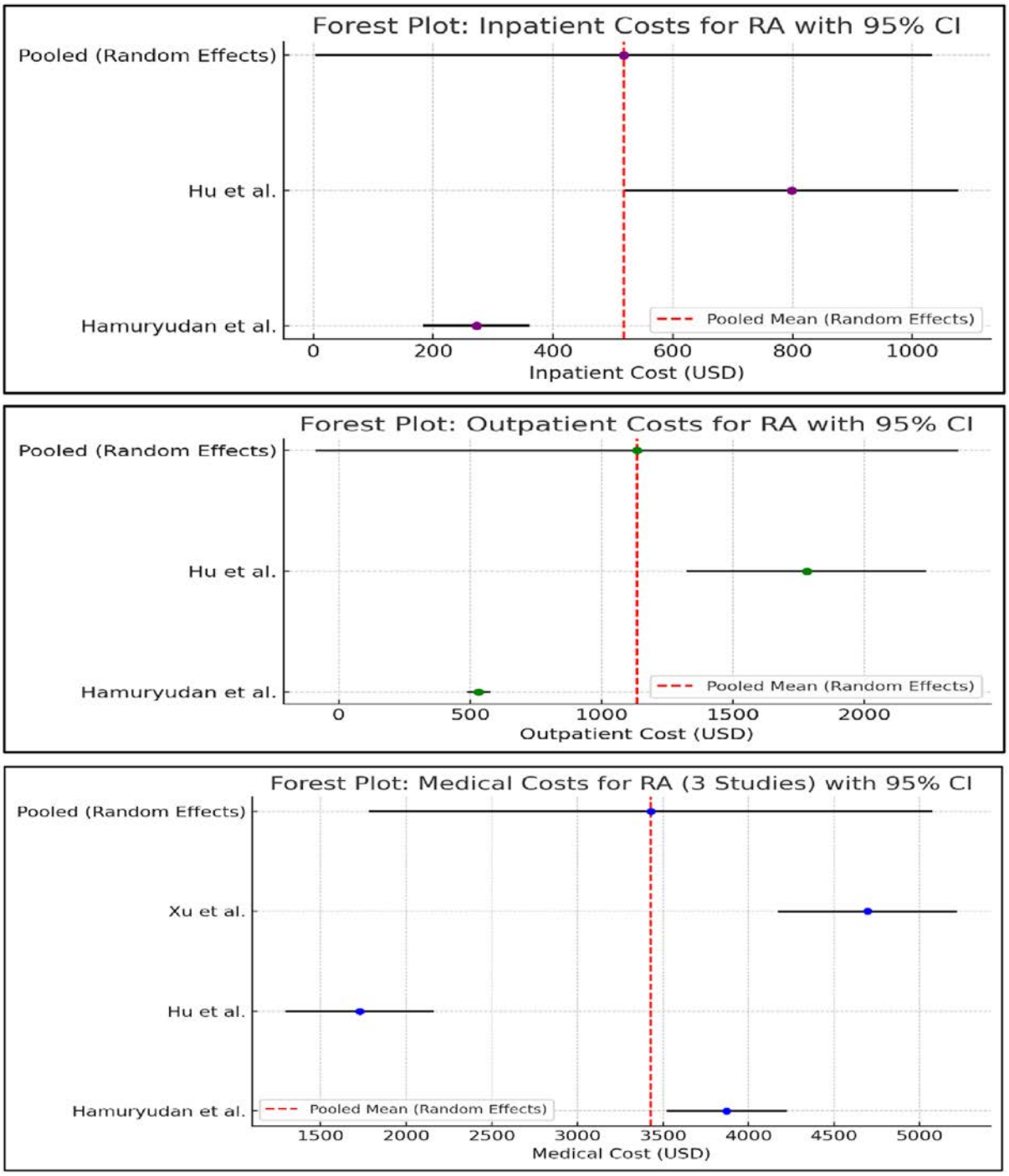

Figure 2. Forest plot of inpatient, outpatient, and medical costs for RA with 95% CI. CI, confidence interval; RA, rheumatoid arthritis. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25627/abstract.

As stated above, this review highlights significant variation in both direct and indirect costs associated with RA. Direct costs ranged from US$523 to US$2,837.90 per patient annually, whereas indirect costs varied from US$81.80 to US$2,463.40. These discrepancies are likely attributed to differences in health care systems, economic conditions, and research methodologies across countries, especially in LMICs, where patients often face long travel distances, high out-of-pocket expenses, and limited access to medications and diagnostic tools compared to high-income nations.[32] The structure of health care systems also influences these costs, with advanced health care systems typically incurring higher direct costs. Indirect costs, representing productivity losses from patients and caregivers, are harder to quantify and tend to be lower due to broader societal factors.[33]

Table 3. Direct and indirect costs of RA*

Author, country, year of costing	Direct costs	Direct costs, inflated 2024 US$	Indirect costs	Indirect costs, inflated 2024 US$	Total costs	Total costs, inflated 2024 US$
Baser et al,[23] Turkey, 2013	US$2,660 per patient, 95% CI US$2,327.50–US $2,990.51	US$3,485 per patient, 95% CI US$3,049–US$3,918	NR	NR	NR	NR
Ayan et al,[24] Turkey, 2018	Median (range) US$3,538 (US$124–US$4,353) per patient per year	Median (range) US$20,167 (US$707–US$24,812) per patient per year	Median (range) US$359 (US$124–US$1,186) per patient per year	NR	NR	NR
Horváth et al,[22] Hungary, 2012	US$148 million per year	US$198.30 million	NR	NR	NR	NR
Codreanu et al,[25] Romania, 2014	NR	NR	US$2,720 per patient per year	US$3,697 per patient per year	NR	NR
Fellous et al,[31] Morocco, 2020	Total annual cost = US $1,115,483 (drug prices, cost of infusions, cost of subcutaneous injection) Median (range) US$1,898 (US$1,678–US$11,272) per patient per year	Total annual cost ≈ US $1.25 million Median (range) US$2,132 (US$1,885–US$12,642) per patient per year	NR	NR	NR	NR
Osiri et al,[17] Thailand, 2007	Mean US$2,135 (median US$1,186) per patient per year	Mean US$2,837.90, (median US$1,576.50) per patient per year	Mean US$547 (median US $182) per patient per year	Mean US$3,564 (median US$1,183) per patient per year	Mean US$2,682 (median US$1,434) per year	Mean US$3,565 (median US$1,906) per patient per year
Hu et al,[18] China, 2013	Mean (±SD) US$1,917.21 (±US$2,559.06) per patient per year	Mean (±SD) US$2,320 (±US $3,098) per patient per year	Mean (±SD) US$492.88 (±US$1,739.74) per patient per year	Mean (±SD) US$596.39 (±US$2,106.08) per patient per year	NR	NR
Xu et al,[21] China, 2009	90.0% of the total costs	90.0% of the total costs	10.0% of the total costs	10.0% of the total costs	Mean (±SD) US$3,826 (±US$5,659) per patient per year	Mean (±SD) US $5,470.18 (±US $8,095.37) per patient per year
Naqvi et al,[19] Pakistan, 2019	Mean US$235.10 per patient per year	Mean US$523 per patient per year	NR	NR	NR	NR
Bali and Singla,[20] India, 2022	Mean (±SD) US$424 (±US $634) per patient per year	Mean (±SD) US$572 (±US $856) per patient per year	Mean (±SD) US$226 (US $273) per patient per year	Mean (±SD) US$252 (±US$302) per patient per year	Mean (±SD) US$652 (±US$746) per patient per year	Mean (±SD) US$724 (±US$828) per patient per year
Mendoza-Gutierrez et al,[30] Mexico, 2016 and 2017	NR	NR	Mean (±SD) US$1,677 (±US $4,180); Caregiver US $640 (±US$747); Workday loss US$485 (±US$1,386) per patient per year	NR	NR	NR
Chermont et al,[28] Brazil, 2002	Mean US$382.89 per patient per year	Mean US$1,723 per patient per year	Mean US$20.13 per patient per year	Mean US$90.59 per patient per year	Mean US$403.04 per patient per year	Mean US$1,813.68 per patient per year
Catay et al,[29] Argentina, 2002	NR	NR	Mean (±SD) US$1,008.80 (±US$1,913) per patient per year	Mean (±SD) US $1,712.12 (±US $3,248.54) per patient per year	US$3,093, 95% CI 1,581–4,605 per patient per year	Mean US$5,250.72, 95% CI US $2,683.56–US $7,818.06 per patient per year

* CI, confidence interval; NR, not reported; RA, rheumatoid arthritis; US$, US dollars.

Methodologic differences in studies often make direct costs more prominent because they are easier to measure than the more elusive indirect costs.[34]

The review also found substantial variation in the study designs and methods used to estimate costs. Nine studies adopted a societal perspective, whereas six used a health care perspective. The societal perspective, which includes both direct and indirect costs, provides a more holistic view of the economic burden of RA.[35] However, the health care perspective is often more feasible to implement, especially in LMICs with limited resources for large-scale societal surveys.[36] The most commonly used data sources were retrospective databases such as health insurance databases and disease registries, followed by self-reported surveys. The reliance on retrospective data and self-reported questionnaires introduces the potential for bias and inaccuracies in the cost estimates, which must be taken into account when interpreting the results.

This review shows that RA places a heavy economic burden on individuals, families, health systems, and society in LMICs, and governments need to take clear, practical steps to address it. The significance of our findings is that RA should be included in national plans for NCDs to ensure adequate resources for prevention, early diagnosis, and treatment.[37] In addition, countries also need national RA registries to track disease prevalence, treatment outcomes, and the costs involved.[38] Moreover, training more health care workers to diagnose RA early can help patients get treatment sooner, reducing disability and health care costs in the long run.[39] Expanding health insurance coverage to include RA services, as well as offering rehabilitation and programs that help patients stay in the workforce, can ease the economic burden on both individuals and society.[3]

We also need the support of HICs and global organizations to reduce the impact of RA in LMICs. HICs can help by providing funding, technical expertise, and resources to build reliable health data systems and national RA registries.[40] Additionally, international partnerships should focus on training health care workers and policymakers in early diagnosis, affordable treatment options, and how to measure health care costs effectively.[39] It is also important for global research projects to include LMICs so that studies produce evidence relevant to these settings.[41] Finally, international agencies and HIC governments should push to ensure RA and other musculoskeletal conditions are part of global NCD programs, so they get the attention and resources they deserve.[42]

This systematic review has both strengths and limitations. Its strength lies in the comprehensive approach taken to synthesize existing evidence, offering a deep understanding of the economic burden of RA. The review's originality is highlighted by its focus on LMICs, where the economic impact of RA is often underexplored, providing valuable insights for policy development in these regions. We also adhered to a standardized and widely accepted reporting framework for systematic reviews and meta-analyses to ensure transparency, consistency, and methodologic rigor throughout the review process. However, one limitation is the substantial heterogeneity across the included studies, particularly regarding data sources and methodologies, which makes it challenging to draw definitive conclusions about the overall economic burden of RA in LMICs. Only studies published in English language were included. Therefore, it is possible that relevant studies published in other languages may have been excluded. Lastly, although the studies included were of moderate to high quality, the reliance on retrospective data and self-reported questionnaire may introduce biases, potentially affecting the accuracy of the cost estimates.

In conclusion, this systematic review highlights the significant economic burden of RA on households, health systems, and society in LMICs, with both direct and indirect costs imposing considerable strain on individuals, families, and health care systems. The substantial variability in cost estimates underscores the need for more standardized, high-quality studies to better understand the full scope of RA's economic impact. Given the rising prevalence of RA and its associated costs, the significance of our findings is that there is a pressing need for targeted policies and interventions to reduce the burden of this disease in low-resource settings.

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 Gebrye 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. Chauhan K, Jandu JS, Brent LH, et al. Continuing education activity. In: Ramipril. StatPearls Publishing; 2024. Accessed January 12, 2025. https://www.ncbi.nlm.nih.gov/books/NBK537119/
2. Bullock J, Rizvi SAA, Saleh AM, et al. Rheumatoid arthritis: a brief overview of the treatment. Med Princ Pract 2018;27(6):501–507.
3. Finckh A, Gilbert B, Hodkinson B, et al. Global epidemiology of rheumatoid arthritis. Nat Rev Rheumatol 2022;18(10):591–602.
4. Zhang W, Anis AH. The economic burden of rheumatoid arthritis: beyond health care costs. Clin Rheumatol 2011;30:S25–S32.
5. Cooper NJ. Economic burden of rheumatoid arthritis: a systematic review. Rheumatology (Oxford) 2000;39(1):28–33.
6. Fortner BV, Demarco G, Irving G, et al. Description and predictors of direct and indirect costs of pain reported by cancer patients. J Pain Symptom Manage 2003;25(1):9–18.
7. Yu JS, Hansen RN, Valderrama A, et al. Indirect costs and workplace productivity loss associated with non-Hodgkin lymphoma. Leuk Lymphoma 2016;57(11):2636–2643.

8. Essue BM, Iragorri N, Fitzgerald N, et al. The psychosocial cost burden of cancer: a systematic literature review. Psychooncology 2020; 29(11):1746–1760.
9. Puchalski Ritchie LM, Khan S, Moore JE, et al. Low- and middle-income countries face many common barriers to implementation of maternal health evidence products. J Clin Epidemiol 2016;76: 229–237.
10. Rudan I, Sidhu S, Papana A, et al; Global Health Epidemiology Reference Group. Prevalence of rheumatoid arthritis in low- and middle-income countries: a systematic review and analysis. J Glob Health 2015;5(1):010409.
11. Moher D, Liberati A, Tetzlaff J, et al; PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: the PRISMA statement. Int J Surg 2010;8(5):336–341.
12. Drummond MF, Sculpher MJ, Claxton K, et al. Methods for the Economic Evaluation of Health Care Programmes. 4th ed. Oxford University Press; 2015.
13. Rice DP. Cost of illness studies: what is good about them? Inj Prev 2000;6(3):177–179.
14. Wells GA, Shea B, O'Connell D, et al. The Newcastle-Ottawa Scale (NOS) for Assessing the Quality of Nonrandomised Studies in Meta-Analyses. Ottawa Hospital Research Institute; 2015.
15. World Bank Group. Consumer price index (2010 = 100). World Bank Group. 2024. Accessed November 20, 2024. https://data.worldbank.org/indicator/FP.CPI.TOTL
16. Salary converter. Salary Converter. 2017. Accessed November 20, 2024. http://salaryconverter.nigelb.me/
17. Osiri M, Maetzel A, Tugwell P. The economic burden of rheumatoid arthritis in a developing nation: results from a one-year prospective cohort study in Thailand. J Rheumatol 2007;34(1):57–63.
18. Hu H, Luan L, Yang K, et al. Burden of rheumatoid arthritis from a societal perspective: a prevalence-based study on cost of this illness for patients in China. Int J Rheum Dis 2018;21(8):1572–1580.
19. Naqvi AA, Hassali MA, Naqvi SBS, et al. Estimation of direct cost of managing rheumatoid arthritis treatment to Pakistani patients using real-world follow-up data. Int J Rheum Dis 2020;23(3):325–333.
20. Bali O, Singla S. Cost of illness and out-of-pocket expenditure of ankylosing spondylitis and rheumatoid arthritis at a tertiary care hospital in North India using conventional disease-modifying antirheumatic drugs. Indian J Rheumatol 2024;19:170–183.
21. Xu C, Wang X, Mu R, et al. Societal costs of rheumatoid arthritis in China: a hospital-based cross-sectional study. Arthritis Care Res (Hoboken) 2014;66(4):523–531.
22. Horváth CZ, Sebestyén A, Österle A, et al. Economic burden of long-term care of rheumatoid arthritis patients in Hungary. Eur J Health Econ 2014;15:S131–S135.
23. Baser O, Burkan A, Baser E, et al. Direct medical costs associated with rheumatoid arthritis in Turkey: analysis from National Claims Database. Rheumatol Int 2013;33:2577–2584.
24. Ayan G, Esatoglu SN, Hatemi G, et al. Disease course and healthcare costs of a cohort of rheumatoid arthritis patients from Turkey. Rheumatol Int 2020;40(7):1037–1044.
25. Codreanu C, Mogoşan C, Popescu C, et al. Analysis of the indirect costs of rheumatoid arthritis in Romania. Biomed Res Int 2019; 2019(1):9343812.
26. Hamuryudan V, Direskeneli H, Ertenli I, et al. Direct and indirect healthcare costs of rheumatoid arthritis patients in Turkey. Clin Exp Rheumatol 2016;34(6):1033–1037.
27. Santos-Moreno P, Gómez-De la Rosa F, Parra-Padilla D, et al. Frequency of health care resource utilization and direct medical costs associated with psoriatic arthritis in a rheumatic care center in Colombia. Psoriasis (Auckl) 2021;11:31–39.
28. Chermont GC, Kowalski SC, Ciconelli RM, et al. Resource utilization and the cost of rheumatoid arthritis in Brazil. Clin Exp Rheumatol 2008;26(1):24–31.
29. Catay E, Del Cid CC, Narváez L, et al. Cost of rheumatoid arthritis in a selected population from Argentina in the prebiologic therapy era. Clinicoecon Outcomes Res 2012;4:219–225.
30. Mendoza-Gutierrez CF, Montiel-Ojeda D, Vargas-Chanes D, et al. Health and economic impact associated with rheumatoid arthritis discharges: a cost analysis of a two-year cohort in Mexico. BMC Health Serv Res. 2023;23(1):1320.
31. Fellous S, Rkain H, Ahid S, et al. One-year direct costs of biological therapy in rheumatoid arthritis and its predictive factors: data from the Moroccan RBSMR registry. Rheumatol Int 2021;41(4):787–793.
32. García-Mochón L, Špacírová Z, Espín J. Costing methodologies in European economic evaluation guidelines: commonalities and divergences. Eur J Health Econ 2022;23(6):979–991.
33. Fazal SA, Khan M, Nishi SE, et al. A clinical update and global economic burden of rheumatoid arthritis. Endocr Metab Immune Disord Drug Targets. 2018;18(2):98–109.
34. Cisternas MG, Blanc PD, Yen IH, et al. A comprehensive study of the direct and indirect costs of adult asthma. J Allergy Clin Immunol 2003; 111(6):1212–1218.
35. Drost RMWA, van der Putten IM, Ruwaard D, et al. Conceptualizations of the societal perspective within economic evaluations: a systematic review. Int J Technol Assess Health Care 2017;33(2): 251–260.
36. Daccache C, Rizk R, Dahham J, et al. Economic evaluation guidelines in low- and middle-income countries: a systematic review. Int J Technol Assess Health Care 2021;38(1):e1.
37. Beaglehole R, Bonita R, Horton R, et al; Lancet NCD Action Group; NCD Alliance. Priority actions for the non-communicable disease crisis. Lancet 2011;377(9775):1438–1447.
38. Cross M, Smith E, Hoy D, et al. The global burden of rheumatoid arthritis: estimates from the global burden of disease 2010 study. Ann Rheum Dis 2014;73(7):1316–1322.
39. van der Linden MP, le Cessie S, Raza K, et al. Long-term impact of delay in assessment of patients with early arthritis. Arthritis Rheum 2010;62(12):3537–3546.
40. Bloom DE, Cafiero E, Jané-Llopis E, et al. The Global Economic Burden of Non-Communicable Diseases. World Economic Forum; 2011.
41. Vallejo B, Wehn U. Capacity development evaluation: the challenge of the results agenda and measuring return on investment in the Global South. World Dev 2016;79:1–3.
42. Liu S, Wang B, Fan S, et al. Global burden of musculoskeletal disorders and attributable factors in 204 countries and territories: a secondary analysis of the Global Burden of Disease 2019 study. BMJ Open 2022;12(6):e062183.

Arthritis Care & Research
Vol. 78, No. 3, March 2026, pp 337–343
DOI 10.1002/acr.25631

Cost-Effectiveness of Low-Dose Colchicine Prophylaxis When Starting Allopurinol Using the "Start-Low Go-Slow" Approach for Gout: Evidence From a Noninferiority Randomized Double-Blind Placebo-Controlled Trial

Yana Pryymachenko,[1] Ross Wilson,[1] Nicola Dalbeth,[2] J. Haxby Abbott,[1] and Lisa Stamp[3]

Objective. The aim of this study was to investigate the cost-effectiveness of low-dose colchicine prophylaxis for preventing gout flares when starting allopurinol using the "start-low go-slow" approach.

Methods. Participants with gout, fulfilling the American College of Rheumatology recommendations for starting urate-lowering therapy and with serum urate ≥0.36 mmol/L (6 mg/dL), were randomly allocated (1:1) to either colchicine (0.5 mg daily) or placebo for six months with a further six-month follow-up. All participants received allopurinol, with monthly increase in dose to achieve target urate <0.36 mmol/L. The primary outcomes were incremental cost-effectiveness at the 6-month and 1-year follow-up from the health system perspective, measured by incremental net monetary benefit (INMB) at a willingness-to-pay threshold equivalent to gross domestic product per capita.

Results. Two hundred participants were randomized to either colchicine (n = 100) or placebo (n = 100). Mean costs were higher in the colchicine group over both six months and one year (adjusted mean difference $1,848 [95% confidence interval (CI) –$321 to $4,017] and $2,282 [95% CI –$173 to $4,737], respectively). Quality-adjusted life years were slightly higher in the colchicine group over six months (adjusted mean difference 0.008 [95% CI –0.020 to 0.035]) but lower over one year (–0.015 [95% CI –0.039 to 0.010]). Treatment with colchicine was not found to be cost-effective at either 6 months or 12 months (INMB –$1,373 [95% CI –$4,287 to $1,542] and –$3,191 [95% CI –$6,274 to –$107], probability of cost-effectiveness 17.7% and 1.5%, respectively). Similar results were obtained from a societal perspective.

Conclusion. Six months of low-dose colchicine prophylaxis when starting allopurinol using the "start-low go-slow" approach is unlikely to be cost-effective over 12 months.

INTRODUCTION

Gout is one of the most common forms of arthritis worldwide. The mainstay of long-term management is reduction of serum urate (SU) below 0.36 mmol/L (6 mg/dL). This requires the use of urate-lowering therapies (ULT), of which allopurinol is considered first line.[1] It is well recognized that the introduction of ULT is associated with a paradoxical increase in gout flares. For this reason, anti-inflammatory prophylaxis is recommended for the first six months of ULT.[1,2] Colchicine, nonsteroidal anti-inflammatories, and low-dose prednisone are all options, with low-dose colchicine (0.5–0.6 mg once or twice daily) and low-dose nonsteroidal anti-inflammatory drug (NSAID) considered first line.[1,2] The ability of colchicine or NSAID to reduce the number of gout flares when starting ULT has been demonstrated in several studies.[3,4]

In the last decade, the way in which allopurinol is commenced has altered after the recognition that the starting dose of allopurinol is related to the rare but often fatal allopurinol sensitivity syndrome.[5] It is now recommended that allopurinol is commenced at a low dose (50–100 mg daily) with gradual escalation to achieve the target urate, known as "start-low go-slow" dose

Supported by the Health Research Council of New Zealand (grant 18/151).
[1]Yana Pryymachenko, PhD, Ross Wilson, PhD, J. Haxby Abbott, PhD: Centre for Musculoskeletal Outcomes Research, University of Otago, Dunedin, New Zealand; [2]Nicola Dalbeth, MD: University of Auckland, Auckland, New Zealand; [3]Lisa Stamp, PhD: University of Otago, Christchurch, Christchurch, New Zealand.
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.25631).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25631.
Address correspondence via email to Lisa Stamp, PhD, at lisa.stamp@cdhb.health.nz.
Submitted for publication March 3, 2025; accepted in revised form August 4, 2025.

SIGNIFICANCE & INNOVATIONS

- Taking colchicine prophylaxis led to only slight improvements in health-related quality of life (HRQoL) over the six-month time horizon; however, after cessation of colchicine, HRQoL declined in the subsequent six-month period.
- Six months of low-dose colchicine prophylaxis when starting allopurinol using the "start-low go-slow" approach is unlikely to be cost-effective.

escalation strategy. More recent evidence suggests that anti-inflammatory prophylaxis with colchicine may not be required using a similar "start-low go-slow" approach with febuxostat.[6] In comparison, our recent noninferiority randomized controlled trial of people with gout starting allopurinol using the "start-low go-slow" strategy showed that colchicine reduced gout flares in the first six months.[5] An important feature of this study was that it extended for a further six months after stopping colchicine/placebo. Unexpectedly, in the six months after stopping colchicine, gout flares rose, an effect that was not observed in those who received placebo. The net effect was no difference in the mean number of gout flares per month between the colchicine and placebo groups over the entire 12-month study period.[5]

There has been little focus on cost-effectiveness of colchicine when used for anti-inflammatory prophylaxis when starting ULT. Colchicine is relatively inexpensive in most countries ($0.06 New Zealand dollars [NZD] per tablet in NZ),[7] although in the USA, it is much more expensive ($5.10 USA dollars [USD] per tab),[8] and so cost-effectiveness is an important consideration. The "Is colchicine prophylaxis required with start-low go-slow allopurinol dose escalation in gout?" noninferiority randomized controlled trial (Australian New Zealand Clinical Trials Registry Number [ACTRN] 12618001179224)[5] provided the ideal opportunity to undertake a cost-effectiveness analysis incorporating the period on anti-inflammatory prophylaxis as well as the following six-month period. Thus, the aim of this study was to determine the cost-effectiveness of six months of low-dose colchicine for anti-inflammatory prophylaxis when starting allopurinol over both 6 and 12 months.

PATIENTS AND METHODS

Study design. This was a preplanned secondary cost-effectiveness analysis of the "Is colchicine prophylaxis required with "start-low go-slow" allopurinol dose escalation in gout?" noninferiority randomized controlled trial (ACTRN 12618001179224). The study is reported in accordance with the Consolidated Health Economic Evaluation Reporting Standards Statement for reporting of economic evaluations of health interventions.[9]

The methods and primary clinical results have been reported previously.[5] Briefly, participants aged 18 years or older with gout as defined by the 2015 American College of Rheumatology (ACR) gout classification criteria[10] were recruited. The inclusion criteria were having at least one gout flare in the preceding six months, meeting the ACR recommendations for starting ULT,[11] and SU ≥0.36 mmol/L (6 mg/dL) at screening. All participants commenced allopurinol using a "start-low go-slow" approach and were randomized in a 1:1 ratio to either colchicine (0.5 mg daily) or placebo for the first six months with a further six-month follow-up after ceasing study medication. Participants, treating clinicians, and all research staff were blinded to treatment allocation. The analyses reported in this paper were done by researchers not involved in trial design or conduct. The trial was powered for clinical outcomes at six months.[5]

This study was conducted in accordance with the principles of the Declaration of Helsinki. Ethical approval was obtained from the Health and Disability Ethics Committee, New Zealand (18/STH/156) and all participants provided written informed consent. Participants gave informed consent to participate in the study before taking part.

Outcomes. The outcomes for this analysis were cost-effectiveness over 6-month and 1-year time horizons from the health system (primary) and societal perspectives, at the policy-relevant willingness-to-pay (WTP) thresholds of 0.5, 1 (primary), 2, and 3 times gross domestic product per capita[12,13,14] ($62,488 NZD in 2018).

Effectiveness was measured by incremental quality-adjusted life years (QALYs), calculated using the EQ-5D-3L health utility instrument.[15] Health utility values for the EQ-5D-3L were calculated using NZ population preference weights.[16] QALYs were calculated by applying the area under the curve method to utility values from assessments at baseline and 3-, 6-, 9-, and 12-month follow-ups.

Costs were calculated from the health system and societal perspectives. The former includes health care costs (both public and private health care costs, including patient out-of-pocket expenses), whereas the latter additionally includes productivity costs, nonhealth social services, and nonhealth out-of-pocket costs to patient and support people. These costs were calculated using participants' responses to a modified version of the Otago Costs and Consequences Questionnaire, which has been developed for use in economic evaluations alongside clinical trials in NZ.[17] Self-reported resource use and costs (for any health reason) were collected at baseline and at months 3, 6, 9 and 12 using a 3-month recall period.

Reference prices (unit costs) from a range of sources were applied to the reported quantities of NZ health services (Supplementary Table S1). Productivity losses were calculated using the friction-cost method using a 6-month friction period[18,19] and applying national median wage rates stratified by age and

sex.[20] All costs are reported in 2018 NZD ($1 NZD ≈ $0.69 USD in 2018) exclusive of goods and services tax.

Statistical analysis. Missing data were assumed to be missing at random (MAR) and imputed using the multiple imputation by chained equations method,[21] creating 50 imputed datasets.

Costs and QALYs for the treatment and control groups are reported with summary statistics. Between-group differences were estimated using linear regressions, adjusting for trial site and baseline costs and EQ-5D utility value.[22] Uncertainty was computed using nonparametric bootstrapping (drawing 100 bootstrap samples from each imputed dataset).[23]

Cost-effectiveness results are summarized by incremental cost-effectiveness ratios (ICERs) in the case that both incremental costs and QALYs are positive, and incremental net monetary benefits (INMBs) calculated at the four specified WTP thresholds. Uncertainty is reported by 95% confidence intervals (CIs) on the incremental costs and effects and INMBs, presentation of the bootstrap estimates on the cost-effectiveness plane, and cost-effectiveness acceptability curves across a range of WTP levels from $0 to $200,000 per QALY.

Multiple sensitivity analyses were performed to evaluate the robustness of our findings: (1) including only complete cases (participants completing all four follow-up assessments) to evaluate the impact of imputation of censored data; (2) excluding productivity losses from the analysis from the societal perspective; (3) increasing imputed costs or decreasing imputed health utility values by 10% to 50% to test the robustness of the estimates to deviations from the MAR assumption[21]; (4) including only the per-protocol population (excluding those participants who had sustained SU ≥0.36 mmol/L but did not have allopurinol dose increased and those who were on the study drug for less than three months); (5) including additional baseline covariates (age, sex, body mass index, and duration of gout) in the regression adjustment to account for chance differences between groups.

All analyses were done in R (version 4.3.1).[24]

RESULTS

Baseline characteristics. A total of 200 participants (100 assigned to colchicine and 100 to placebo) were recruited between February 8, 2019, and December 22, 2021, of whom 174 (87%) were retained at one-year follow-up.[5] Those lost to follow-up were younger, were less likely to be of NZ European ethnicity, and had higher baseline costs than those retained (Supplementary Table S2).

Baseline characteristics of the two groups are reported in Table 1. The colchicine group had a higher proportion of Asian and Pacific peoples' ethnicities and a lower proportion of NZ European ethnicity and a slightly higher EQ-5D utility at baseline; there were no other substantive differences between groups (Table 1).

Costs and effects. The colchicine group had higher costs than the placebo group at the 3-, 6-, and 12-month follow-up assessments (Figure 1), resulting in higher total costs over both 6-month and 1-year time horizons (Table 2). The health care resource use for each group is reported in Supplementary Table S3. EQ-5D utility values were higher in the colchicine group at the three-month follow-up and in the placebo group at six and nine months; QALYs were higher for the colchicine group over the six-month time horizon, with no difference between groups over one year (Figure 1 and Table 2).

Cost-effectiveness. After adjusting for baseline differences, over the six-month time horizon, treatment with colchicine was estimated to have incremental costs of $1,848 (95% CI –$321 to $4,017) from the health system perspective and $1,701 (95% CI –$911 to $4,312) from the societal perspective and a QALY gain of 0.008 (95% CI –0.020 to 0.035) compared to placebo (Table 3). The ICER was $242,953 from the health system perspective and $223,601 from the societal perspective. INMB at one WTP level was –$1,373 (95% CI –$4,287 to $1,542) from the health system perspective and –$3,191

Table 1. Baseline characteristics of participants*

Variable	Colchicine	Placebo	Total
N	100	100	200
Age, y	54.8 (16.0)	57.1 (15.3)	56.0 (15.7)
Male, n (%)	92	94	186
Ethnicity, n (%)			
NZ European	51	69	120
Pacific peoples	18	10	28
Asian	17	9	26
Māori	12	11	23
Other	2	1	3
Duration of gout, y	10.1 (8.8)	12.3 (11.1)	11.2 (10.1)
NZ health system costs over 3 mo, NZD	1,561 (3,604)	1,366 (3,503)	1,463 (3,546)
Societal costs over 3 mo, NZD	2,700 (5,173)	2,524 (5,031)	2,612 (5,090)
HRQoL (EQ-5D utility)	0.87 (0.18)	0.83 (0.19)	0.85 (0.19)

* Values are mean (SD) unless specified otherwise. HRQoL, health-related quality of life; NZD, New Zealand dollars.

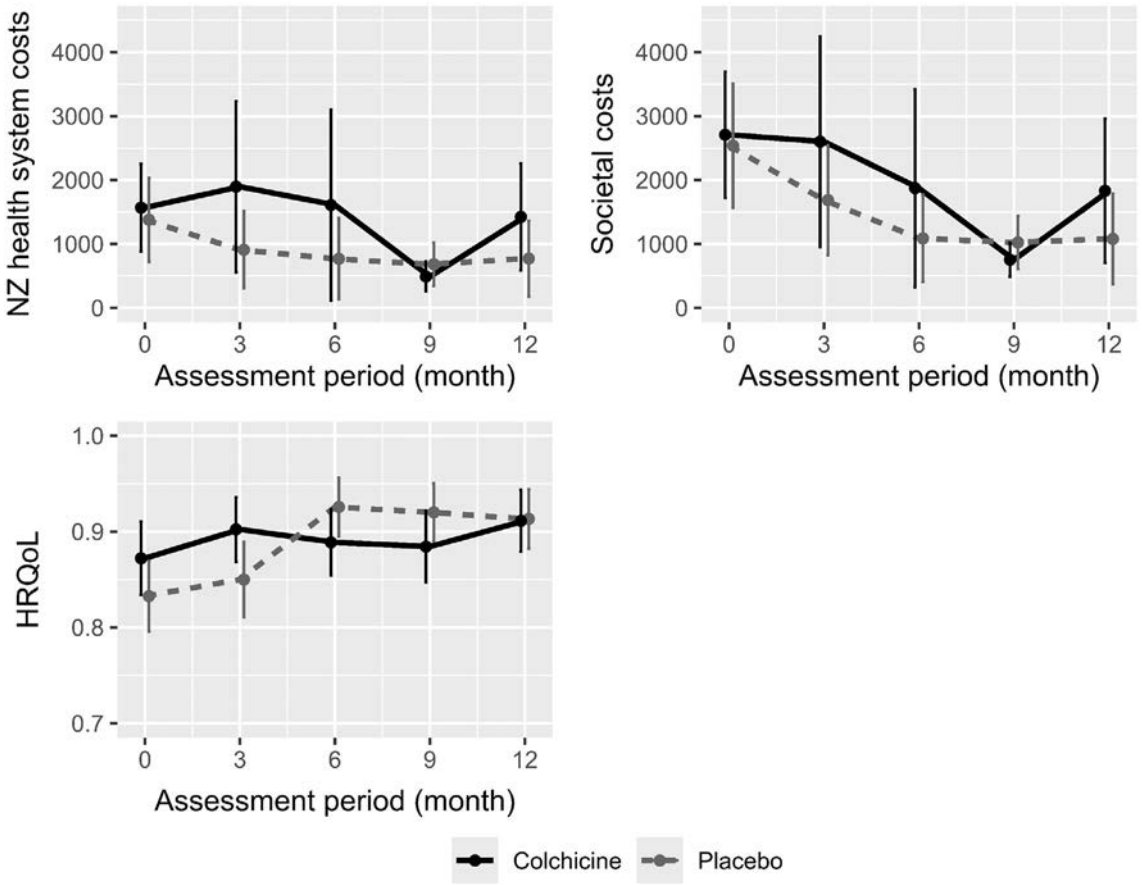

Figure 1. HRQoL (EQ-5D utility) and costs over the follow-up period. The values represent means with 95% CI. CI, confidence interval; HRQoL, health-related quality of life; NZ, New Zealand.

(95% CI –$6,274 to –$107) from the societal perspective. The probability of colchicine being cost-effective over the six-month time horizon was around 18% from the health system perspective and 24% from the societal perspective at the primary WTP level, increasing but remaining below 50% at all considered WTP thresholds (Figure 2).

Over the one-year time horizon, treatment with colchicine for six months had incremental costs of $2,282 (95% CI –$173 to $4,737) from the health system perspective and $2,167 (95% CI –$823 to $5,158) from the societal perspective and a QALY loss of –0.015 (95% CI –0.039 to 0.010) compared to placebo (Table 3). Colchicine was dominated by placebo (ie, it had higher

Table 2. Mean (SD) costs and QALYs over 6 months and 1 year follow-up*

Variable	6 mo		12 mo	
	Placebo	Colchicine	Placebo	Colchicine
Intervention (colchicine) costs	0 (0)	18 (0)	0 (0)	18 (0)
Other medication costs	192 (320)	153 (219)	392 (667)	412 (1,037)
Health professional costs	435 (518)	397 (497)	782 (867)	744 (732)
Other health care costs	1,050 (3,091)	2,956 (9,907)	1957 (4,696)	4,264 (10,500)
Investigations costs	599 (1,687)	1,810 (6,200)	1,298 (3,032)	2,469 (6,421)
Hospital costs	451 (2,341)	1,147 (4,258)	660 (2,589)	1,794 (5,000)
Total health care costs	1,677 (3,428)	3,525 (10,064)	3,130 (5,216)	5,420 (10,984)
Productivity costs	918 (2,776)	820 (3,226)	1,366 (3,169)	1,333 (3,914)
Other societal costs	179 (319)	147 (198)	380 (720)	306 (535)
Total societal costs	2,774 (4,817)	4,493 (12,044)	4,877 (6,668)	7,059 (13,258)
Patient costs	518 (865)	558 (1,594)	972 (1,378)	879 (1,882)
QALYs	0.86 (0.13)	0.89 (0.13)	0.89 (0.10)	0.89 (0.11)

* Costs are presented in 2018 NZD. NZD, New Zealand dollars; QALY, quality-adjusted life year.

Table 3. Incremental cost, effects, cost-effectiveness ratios, and net monetary benefits*

Metric	6 mo		12 mo	
	Health system perspective	Societal perspective	Health system perspective	Societal perspective
Incremental costs and effects				
Costs, NZD	1,848 (−321 to 4,017)	1,701 (−911 to 4,312)	2,282 (−173 to 4,737)	2,167 (−823 to 5,158)
QALYs	0.008 (−0.020 to 0.035)		−0.015 (−0.039 to 0.010)	
ICER, NZD	242,953	223,601	N/A[a]	N/A[a]
INMB, NZD at WTP threshold				
0.5 × GDP	−1,610 (−4,028 to 807)	−1,463 (−4,317 to 1,391)	−2,737 (−5,420 to −53)	−2,622 (−5,847 to 603)
1 × GDP	−1,373 (−4,287 to 1,542)	−1,225 (−4,539 to 2,088)	−3,191 (−6,274 to −107)	−3,076 (−6,681 to 529)
2 × GDP	−897 (−5,178 to 3,383)	−750 (−5,352 to 3,852)	−4,100 (−8,285 to 86)	−3,985 (−8,630 to 660)
3 × GDP	−422 (−6,269 to 5,425)	−275 (−6,391 to 5,842)	−5,008 (−10,491 to 475)	−4,894 (−10,784 to 996)

* GDP, gross domestic product; ICER, incremental cost-effectiveness ratio; INMB, incremental net monetary benefit; N/A, not applicable; NZD, New Zealand dollars; QALY, quality-adjusted life year; WTP, willingness-to-pay.
[a] Colchicine prophylaxis dominated by placebo.

costs and lower QALYs), and the probability of its being cost-effective was below 5% at all WTP levels. INMB at one WTP level was –$1,225 (95% CI –$4,539 to $2,088) from the health system perspective and –$3,076 (95% CI –$6,681 to $529) from the societal perspective (Table 3 and Figure 2). The bootstrap replications of costs and effects are shown on the cost-effectiveness plane in Supplementary Figure S1.

Sensitivity analyses. The results were robust to restricting the analysis sample to complete cases, excluding productivity costs from the societal perspective, using only the per-protocol population, and including additional covariates (Supplementary Tables S4–7). They were also robust to increasing the imputed costs or reducing the imputed QALYs (Supplementary Figure S2), suggesting plausibility of the MAR assumption.

DISCUSSION

The results of this study suggest that low-dose colchicine anti-inflammatory prophylaxis when starting allopurinol using the "start-low go-slow" approach for gout is unlikely to be cost-effective. Compared to placebo, colchicine use resulted in higher health care costs and negligible differences in QALYs.

The cost-effectiveness of ULT with allopurinol used in a treat-to target SU manner for the long-term management of gout is well established.[25–27] Adherence with ULT remains a core issue in the management of gout with high rates of nonadherence and nonpersistence with ULT in gout.[28,29] Gout flares are central to the patients' experience of gout, and the paradoxical increase in flares after starting ULT can contribute to poor persistence with therapy. The use of anti-inflammatory prophylaxis to prevent gout flares has been shown to reduce gout flares when starting ULT and therefore thought to contribute to persistence with ULT.

To our knowledge, there is only one other study examining the cost-effectiveness of anti-inflammatory prophylaxis during the early phase of ULT. Robinson et al examined the cost-effectiveness of colchicine prophylaxis when commencing allopurinol for gout, employing simulation modeling for the USA and Australia and reported colchicine to be very cost-effective at six months.[30] It is important to note there are some key differences between the current analysis and that from Robinson et al. Firstly,

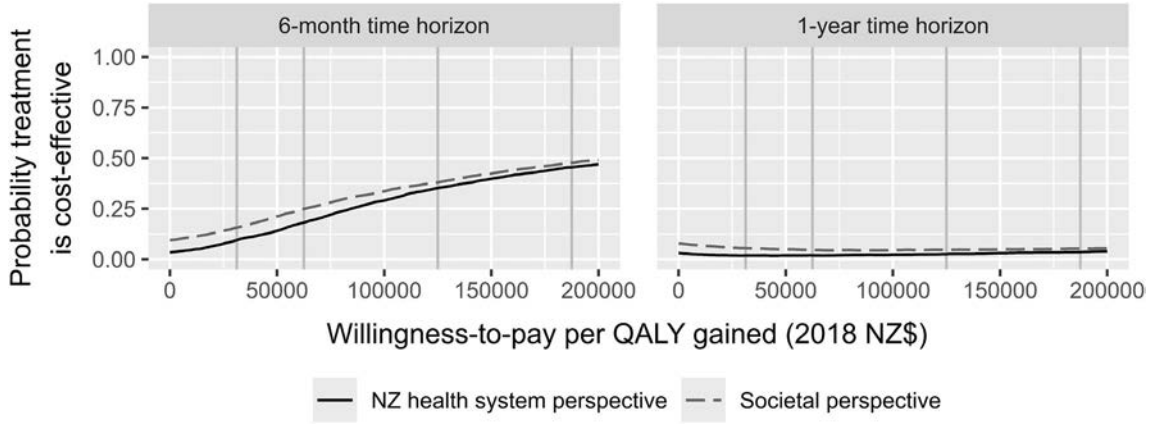

Figure 2. Cost-effectiveness acceptability curves. The vertical lines represent willingness-to-pay thresholds of 0.5, 1, 2, and 3 × GDP/capita. GDP, gross domestic product; NZ, New Zealand; QALY, quality-adjusted life year.

our cost-effectiveness analysis is a prespecified secondary analysis of a randomized controlled trial that included 200 participants using contemporary allopurinol dosing strategies. In comparison, Robinson et al developed a decision tree model based on a study[4] of 43 people commencing allopurinol at 100 mg daily, with dose escalation every two to three weeks by 50 to 100 mg, which is faster than currently recommended.

Of particular importance is that the randomized controlled trial upon which our cost-effectiveness analysis was based included the six-month period after stopping colchicine or placebo. Although flares were less frequent during the period of prophylaxis for those receiving colchicine compared to placebo, there was an unexpected rise in flares after ceasing colchicine, but not placebo. Consistent with this pattern of flares over the entire 12-month study period, QALYs increased in those who received colchicine over the 6-month time horizon compared to placebo, with a decline in QALYs after cessation of colchicine, resulting in no difference in QALYs over the entire 12-month period between those who received colchicine and those who received placebo. It is also important to note that there was no difference in the reduction in SU or the mean (SD) allopurinol dose between those who received colchicine and those who received the placebo, indicating no difference in adherence with allopurinol between arms.

An interesting observation is that the colchicine group had much higher health care costs associated with investigations and hospital stays. This could be driven by higher number of serious adverse events in this group.[5]

Strengths of this study include the high-quality source trial with long follow-up period and comprehensive cost data collection. Sensitivity analyses confirmed that our results were robust to modeling assumptions and several potential sources of bias, including imputation of missing outcomes data, measurement error and uncertainty in productivity costs, and violations of the study protocol. There are several limitations of this study. The trial was designed and powered to a clinical noninferiority research question, not for the economic evaluation. As is often seen in trial-based cost-effectiveness analyses, uncertainty in incremental costs was wide; nevertheless, our results provide moderate certainty at six months and high certainty at one year that colchicine is not cost-effective when used as anti-inflammatory prophylaxis during the initiation of allopurinol. Participants were required to have at least one gout flare in the past six months, so our results likely overstate the effectiveness of colchicine in the general gout population who are having less frequent flares. Finally, this study was conducted in NZ where the cost of colchicine is low; however if the cost of colchicine were higher, as in the USA, our results would show colchicine to be even less cost-effective.

The rise in flares after stopping colchicine, an effect not seen in those who stopped placebo, was an unexpected finding in the primary trial. This effect has since been observed in other clinical trials.[31] Whether a longer period of anti-inflammatory prophylaxis would continue to reduce flares and therefore improve the cost-effectiveness is unclear. Future research might focus on reduction of rebound gout flares after ceasing prophylaxis or a more prolonged period of prophylaxis when starting ULT.

Six months of treatment with colchicine for anti-inflammatory prophylaxis against gout flares when starting allopurinol is unlikely to be cost-effective, either during the period of treatment (due to higher costs and minimal QALY gains) or at one-year follow-up (due to higher costs and QALY losses, largely driven by gout flares following colchicine withdrawal).

ACKNOWLEDGMENTS

Open access publishing facilitated by University of Otago, as part of the Wiley - University of Otago agreement via the Council of Australian University Librarians.

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 Stamp 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. FitzGerald JD, Dalbeth N, Mikuls T, et al. 2020 American College of Rheumatology guideline for the management of gout. Arthritis Care Res (Hoboken) 2020;72(6):744–760.
2. Richette P, Doherty M, Pascual E, et al. 2016 updated EULAR evidence-based recommendations for the management of gout. Ann Rheum Dis 2017;76(1):29–42.
3. Wortmann RL, Macdonald PA, Hunt B, et al. Effect of prophylaxis on gout flares after the initiation of urate-lowering therapy: analysis of data from three phase III trials. Clin Ther 2010;32(14):2386–2397.
4. Borstad GC, Bryant LR, Abel MP, et al. Colchicine for prophylaxis of acute flares when initiating allopurinol for chronic gouty arthritis. J Rheumatol 2004;31(12):2429–2432.
5. Stamp L, Horne A, Mihov B, et al. Is colchicine prophylaxis required with start-low go-slow allopurinol dose escalation in gout? A non-inferiority randomised double-blind placebo-controlled trial. Ann Rheum Dis 2023;82(12):1626–1634.
6. Yamanaka H, Tamaki S, Ide Y, et al. Stepwise dose increase of febuxostat is comparable with colchicine prophylaxis for the prevention of gout flares during the initial phase of urate-lowering therapy: results from FORTUNE-1, a prospective, multicentre randomised study. Ann Rheum Dis 2018;77(2):270–276.
7. Wilson K, Chong, D, eds. Pharmaceutical schedule. Vol 25. The Pharmaceutical Management Agency; 2018.
8. Ly DP, Giuriato MA, Song Z. Changes in prescription drug and health care use over 9 years after the large drug price increase for colchicine. JAMA Intern Med 2023;183(7):670–676.
9. Husereau D, Drummond M, Augustovski F, et al; CHEERS 2022 ISPOR Good Research Practices Task Force. Consolidated Health

Economic Evaluation Reporting Standards 2022 (CHEERS 2022) statement: updated reporting guidance for health economic evaluations. BMC Med 2022;20(1):23.

10. 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.
11. Khanna D, Fitzgerald JD, Khanna PP, et al; American College of Rheumatology. 2012 American College of Rheumatology guidelines for management of gout. Part 1: systematic nonpharmacologic and pharmacologic therapeutic approaches to hyperuricemia. Arthritis Care Res (Hoboken) 2012;64(10):1431–1446.
12. Robinson LA, Hammitt JK, Chang AY, et al. Understanding and improving the one and three times GDP per capita cost-effectiveness thresholds. Health Policy Plan 2017;32(1):141–145.
13. Thokala P, Ochalek J, Leech AA, et al. Cost-effectiveness thresholds: the past, the present and the future. PharmacoEconomics 2018; 36(5):509–522.
14. Edney LC, Haji Ali Afzali H, Cheng TC, et al. Estimating the reference incremental cost-effectiveness ratio for the Australian health system. PharmacoEconomics 2018;36(2):239–252.
15. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Ann Med 2001;33(5):337–343.
16. Devlin NJ, Hansen P, Kind P, et al. Logical inconsistencies in survey respondents' health state valuations – a methodological challenge for estimating social tariffs. Health Econ 2003;12(7):529–544.
17. Pinto D, Robertson MC, Hansen P, et al. Good agreement between questionnaire and administrative databases for health care use and costs in patients with osteoarthritis. BMC Med Res Methodol 2011; 11(11):45.
18. Brouwer WB, Koopmanschap MA. The friction-cost method: replacement for nothing and leisure for free? PharmacoEconomics 2005; 23(2):105–111.
19. van den Hout WB. The value of productivity: human-capital versus friction-cost method. Ann Rheum Dis 2010;69(suppl 1):i89–i91.
20. Statistics NZ. Earnings for people in paid employment by region, sex, age groups and ethnic groups. 2021. Accessed June 9, 2025. https://explore.data.stats.govt.nz/vis?fs[0]=Economy%2C0%7CIncome%23CAT_INCOME%23&pg=0&fc=Economy&bp=true&snb=16&vw=tb&df[ds]=ds-nsiws-disseminate&df[id]=INC_INC_001&df[ag]=STATSNZ&df[vs]=1.0&dq=.98.98.98.98.&to[TIME]=false
21. Faria R, Gomes M, Epstein D, et al. A guide to handling missing data in cost-effectiveness analysis conducted within randomised controlled trials. PharmacoEconomics 2014;32(12):1157–1170.
22. Manca A, Hawkins N, Sculpher MJ. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. Health Econ 2005;14(5):487–496.
23. Schomaker M, Heumann C. Bootstrap inference when using multiple imputation. Stat Med 2018;37(14):2252–2266.
24. *R*. Version 4.3.1. R Project for Statistical Computing; 2023. http://www.R-project.org
25. Jutkowitz E, Choi HK, Pizzi LT, et al. Cost-effectiveness of allopurinol and febuxostat for the management of gout. Ann Intern Med 2014; 161(9):617–626.
26. Doherty M, Jenkins W, Richardson H, et al. Efficacy and cost-effectiveness of nurse-led care involving education and engagement of patients and a treat-to-target urate-lowering strategy versus usual care for gout: a randomised controlled trial. Lancet 2018; 392(10156):1403–1412.
27. van de Laar CJ, Janssen CA, Janssen M, et al. Model-based cost-effectiveness analyses comparing combinations of urate lowering therapy and anti-inflammatory treatment in gout patients. PLoS One 2022;17(1):e0261940.
28. Dehlin M, Ekström EH, Petzold M, et al. Factors associated with initiation and persistence of urate-lowering therapy. Arthritis Res Ther 2017;19(1):6. https://doi.org/10.1186/s13075-016-1211-y
29. Briesacher BA, Andrade SE, Fouayzi H, et al. Comparison of drug adherence rates among patients with seven different medical conditions. Pharmacotherapy 2008;28(4):437–443.
30. Robinson PC, Dalbeth N, Donovan P, et al. Cost-effectiveness of colchicine prophylaxis for gout flares when commencing allopurinol. Arthritis Care Res (Hoboken) 2021;73(10):1537–1543.
31. Stamp LK, Frampton C, Newcomb JA, et al. Gout flares after stopping anti-inflammatory prophylaxis: a rapid literature review and meta-analysis. Arthritis Care Res (Hoboken) 2025;77(6):720–726.

Arthritis Care & Research
Vol. 78, No. 3, March 2026, pp 344–351
DOI 10.1002/acr.25622

American College of Rheumatology
Empowering Rheumatology Professionals

Long-Term Opioids in Gout: A Matched Cohort Study From the Veterans Health Administration

Lindsay N. Helget,[1] Bryant R. England,[1] Punyasha Roul,[1] Harlan Sayles,[2] Tuhina Neogi,[3] James R. O'Dell,[1] Joshua F. Baker,[4] and Ted R. Mikuls[1]

Objective. Though used frequently to treat flare, risk of long-term opioid exposure in gout has not been well defined. In this study, we examined the hypothesis that people with gout are more likely than individuals without gout to be prescribed long-term opioids over time.

Methods. In this matched cohort study using national Veterans Health Administration (VHA) data, multivariable Cox regression was used to examine the association of gout with long-term opioid receipt (defined using a validated administrative algorithm). Patients with gout were identified using diagnostic codes and matched with up to 10 controls without gout by age, sex, and VHA enrollment year. In analyses limited to gout, factors associated with long-term opioid exposure were identified.

Results. Over a mean follow-up of 4.52 years, patients with gout were more likely to receive long-term opioids than controls (6.9% vs 3.8%). This risk remained following covariate adjustment (adjusted hazard ratio 1.30; 95% confidence interval 1.28–1.32). Among patients with gout, factors independently associated with an increased likelihood of long-term opioid exposure included more recent index year, younger age, female sex, non-Hispanic Black race and ethnicity, rural residence, being underweight or obese, former or current smoking, greater comorbidity, urate-lowering therapy receipt, and requirement of rheumatology consultation.

Conclusion. Patients with gout are more likely to receive long-term opioids than counterparts without gout, independent of other factors. This risk is greater in underrepresented gout populations, those with greater comorbidity, patients requiring rheumatology consultations, and individuals prescribed urate-lowering treatments. Additional investigation is needed to elucidate whether deployment of optimal gout management minimizes long-term opioid exposure in patients with gout.

INTRODUCTION

Characterized by recurrent and often severely painful articular flares, gout represents the most common form of inflammatory arthritis worldwide.[1] National prevalence estimates approach 4%, which translates to >10 million individuals with gout in the United States alone.[2–4] The burden posed by gout is even greater among select populations, such as those receiving care in the US Veterans Health Administration (VHA), where its prevalence approaches 6%.[2] The articular deposition of monosodium urate leading to arthritis flares, though potentially preventable with appropriate management, frequently results in progressive joint damage, disability, and marked pain. Consequently, gout leads to substantial health care use, including >200,000 attributable emergency department encounters annually in the United States and rates of inpatient hospitalizations for gout doubling between

The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs or the US government.

Supported by a Scientist Development Award from the University of Nebraska Medical Center Department of Internal Medicine (to Dr Helget). Dr England's work was supported by the Department of Veterans Affairs (VA) Clinical Science Research & Development (CSR&D) (IK2CX002203) and the Rheumatology Research Foundation. Dr Neogi's work was supported by grants from the NIH (K24-AR-070892 and P30-AR-072571). Dr Baker's work was supported by a CSR&D Merit Award (CX001703) and Rehabilitation Research & Development Merit Award (RX003644). Dr Mikuls's work was supported by grants from the VA (BX004600 and BX003635), the NIH (U54-GM-115458), and the US Department of Defense (PR200793).

[1]Lindsay N. Helget, MD, Bryant R. England, MD, PhD, Punyasha Roul, MS, James R. O'Dell, MD, Ted R. Mikuls, MD, MSPH: Veterans Affairs Nebraska-Western Iowa Health Care System and University of Nebraska Medical Center, Omaha; [2]Harlan Sayles, MS: University of Nebraska Medical Center, Omaha; [3]Tuhina Neogi, MD, PhD: Boston University School of Medicine, Boston, Massachusetts; [4]Joshua F. Baker, MD, MS: Corporal Michael J. Crescenz Veterans Affairs Medical Center 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.25622).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25622.

Address correspondence via email to Lindsay N. Helget, MD, at lindsay.helget@unmc.edu.

Submitted for publication January 4, 2025; accepted in revised form July 24, 2025.

SIGNIFICANCE & INNOVATIONS

- This is among the first studies to evaluate long-term opioid use in a US gout population.
- In the Veterans Health Administration (VHA) from 2000 to 2020, patients with gout were 30% more likely to receive long-term opioid therapy, independent of other factors.
- Factors most strongly associated with an increased likelihood of long-term opioid exposure in patients with gout included female sex, underweight body mass index (BMI), current smoking, a higher Rheumatic Disease Comorbidity Index score, urate-lowering therapy receipt, and requirement of rheumatology consultation.
- Factors associated with a lower risk of long-term opioid exposure in patients with gout included overweight BMI, the presence of chronic kidney disease, and evidence of adequate serum urate control.

1993 and 2011.[5] Gout-related health care encounters are frequently accompanied by prescriptions for anti-inflammatory medications and other analgesics, including opioids.[6] Previous surveys have shown that between 20% and 30% of patients with gout receive an opioid prescription at the time of discharge from an emergency department encounter.[7,8]

The United States continues amid an opioid epidemic that has accounted for >100,000 deaths in 2021 alone, twice the number of opioid-related US deaths in 2015.[9] With similar trends in other parts of the world, approximately one-third of deaths from opioid overdose are ascribed to misuse of prescription medications.[10] Examining claims data, Brummett and colleagues observed that among patients receiving postprocedure opioid analgesics, up to 6% transitioned to long-term use.[11] Indeed, opioid prescriptions for acute-pain management are well recognized to serve as a "gateway" to long-term use.[12] With reports of frequent administration and dispensing in short-term management settings, it remains unknown whether gout might similarly serve as a risk factor for long-term opioid exposure.

To date, there have been no large-scale investigations examining the frequency of long-term opioid use in gout. In this study, we sought to examine the long-term risk of prolonged opioid exposure in gout, testing the hypothesis that gout would predispose patients to a higher risk of exposure to long-term prescription opioids compared to patients without gout. Among patients with gout, we also sought to identify factors associated with long-term opioid exposure over time.

METHODS

Study design and participants. We performed a matched cohort study within the VHA. We examined national Veterans Affairs (VA) administrative and electronic health record data from January 2000 through July 2020, accessed through the VA Corporate Data Warehouse (CDW) located within the VA Informatics and Computing Infrastructure.[13] The study was approved by the VA Nebraska-Western Iowa Health Care System Institutional Review Board with a waiver of informed consent.

We identified patients with gout using an administrative algorithm that required patients to have an *International Classification of Diseases, Ninth Revision* (ICD-9) code of 274.XX from at least two separate encounters at least 30 days apart.[2] We limited the period for cohort creation up to September 2015 to avoid misclassification from ICD-10 implementation, which started in October 2015. Each patient with gout was matched with up to 10 patients without gout by birth year, sex, and year of VHA enrollment. Patients without gout (controls) were defined as having no prior gout diagnosis and no previous pharmacy dispensing of urate-lowering therapy (ULT; allopurinol, febuxostat, probenecid, or pegloticase). Individuals with any opioid prescription in the 365 days before the index date were excluded. A Consolidated Standards of Reporting Trials diagram is provided in Figure 1. Notably, a higher proportion of potential patients with gout (16.9%) than controls (6.8%) were excluded from the analysis due to opioid exposure during the one-year preindex period.

The index date for patients with gout was defined as the date corresponding to the second diagnostic code, whereas controls were assigned an index date that corresponded to the same date as their matched counterpart with gout. All patients with gout and controls had one year or more of preindex VA follow-up. Patients were observed from the index date until the earliest of incident opioid exposure leading to long-term administration, until death, or up to five years following the index date (through July 2020). Follow-up time was limited to five years because long-term opioid prescriptions occurring later were more likely to be attributable to causes other than gout. Patients without gout who eventually met the gout algorithm were also censored at the time they fulfilled the algorithm, after which they were matched to controls without gout and contributed to follow-up time and events as a case. Vital status was determined by linkage with the National Death Index.

Long-term opioid exposure. Using a previously validated approach, long-term opioid exposure was defined as ≥90 cumulative days' supply from two or more pharmacy dispensing episodes occurring in a six-month window in the absence of a gap ≥32 days.[14] Medications considered as opioids in this study included acetaminophen/hydrocodone, buprenorphine, codeine, fentanyl, hydromorphone, methadone, morphine, oxycodone, and tramadol.

Patient demographics and comorbidities. Covariates were defined using CDW data and a one-year lookback period. Covariates included age, sex, race and ethnicity, urban or rural residence, body mass index (BMI), smoking status (current,

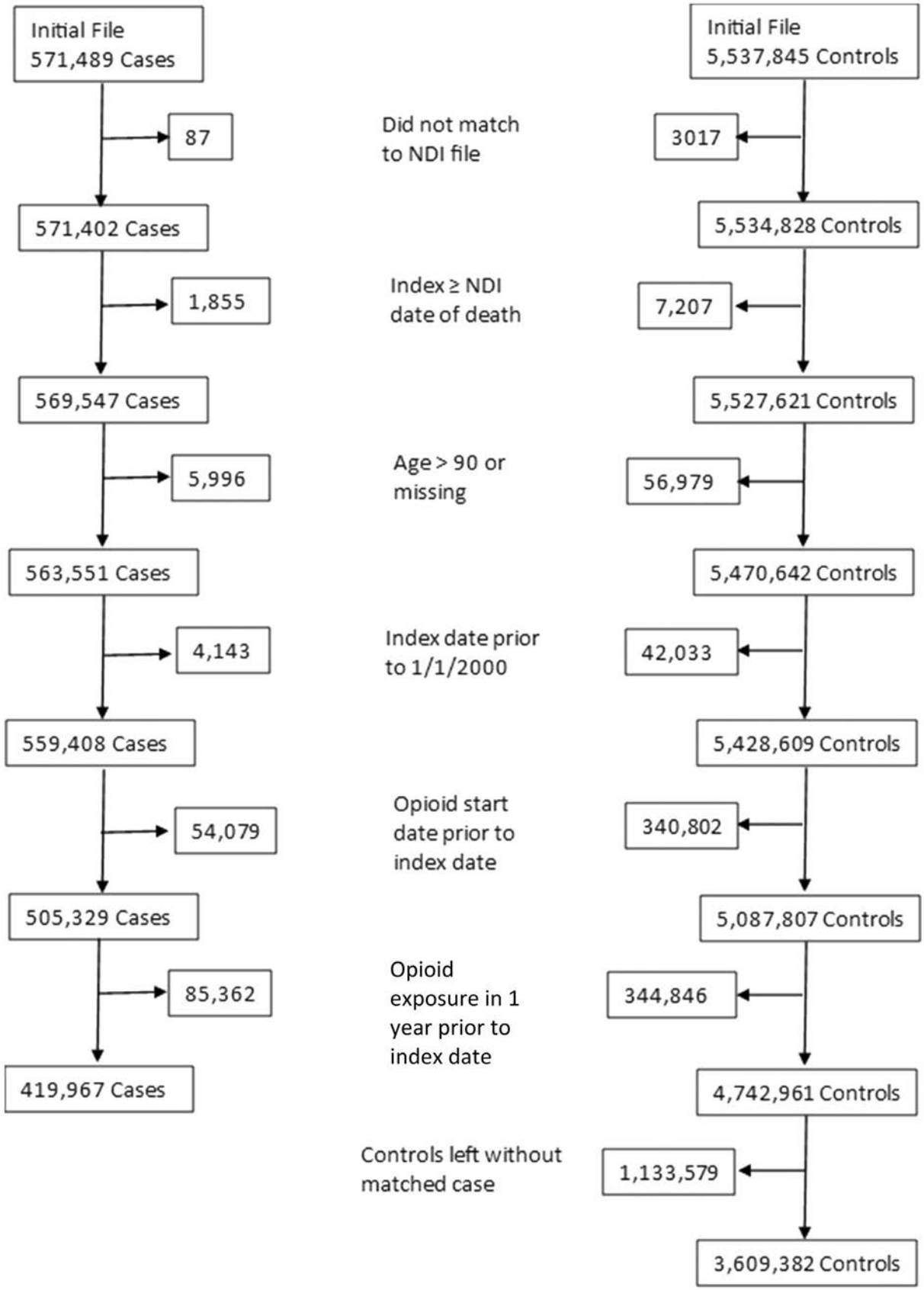

Figure 1. Consolidated Standards of Reporting Trials diagram demonstrating steps and exclusions involved in the development of a matched cohort of patients with gout and controls without gout. NDI, National Death Index.

former, never), and comorbidities. As a social determinant of health, race and ethnicity (non-Hispanic White, non-Hispanic Black, or other) was included as a covariate given differences in gout burden reported across populations.[15,16] "Other" race and ethnicity included those reporting American Indian or Alaska Native, Asian, Black Hispanic, multiple races, Native Hawaiian or Pacific Islander, or White Hispanic status. These groups were combined given low numbers of individuals in each category, which resulted in limited power to conduct more granular analyses based on self-reported race and ethnicity. BMI was calculated from vital signs data using the weight most proximate to the index date and modal height and was categorized as <20, 20 to <25 (referent), 25 to <30, and ≥30. The Rheumatic Disease Comorbidity Index (RDCI) score was calculated as a measure of comorbidity burden.[17] Comorbid conditions composing the RDCI and those included in the sensitivity analyses were defined by the presence of two or more diagnostic codes[18] from inpatient or outpatient encounters at any time before or on the index date. Additionally, chronic kidney disease (CKD) was defined using the estimated glomerular filtration rate available from laboratory data and values most proximate to the index date. CKD was classified as stage 1 (>90 mL/min; referent), 2 (60–89 mL/min), 3 (30–59 mL/min), or 4 to 5 (<30 mL/min). Residence was defined using the VA urban–rural indicator variable available in the CDW, with comparisons of individuals defined as urban versus rural or highly rural.[19]

Gout-related care factors. To examine associations between measures of gout care and long-term opioid exposure, serum urate (sU) control and ULT receipt were assessed during each year of follow-up post index. Briefly, patients with gout with a mean sU level of >7 mg/dL during a one-year period were considered to have inadequate sU control, and those with a mean sU level of ≤7 mg/dL were considered to have adequate sU control. Patients were considered to have received ULT with at least one episode of pharmacy dispensing of allopurinol, febuxostat, or probenecid during each one-year block. Missing sU values for each one-year period were imputed using available values for up to two years prior. Thus, sU control and ULT receipt variables, measured before any long-term opiate receipt, could change year to year within a patient. As informative measures of gout severity (ie, tophi; flare frequency, intensity, and duration) were not reliably available in the CDW, we modeled rheumatology care as a surrogate of severity.[20] The requirement of rheumatology consultation was defined as a rheumatology visit occurring during follow-up with an associated diagnostic code (ICD-9 274.X or ICD-10 M10.X). In contrast to other covariates, which were ascertained at baseline or during the at-risk follow-up period (ie, sU control/ULT receipt), we considered patients to have satisfied the requirement of rheumatology consultation with any rheumatology encounter occurring during the maximum five-year follow-up. This broader time allowance was based on the assumption that a need for subspecialty consultation would serve as a surrogate of gout severity independent of whether it occurred before opioid exposure or in the immediate postexposure period.

Statistical analyses. Crude incidence rates of long-term opioid exposure were calculated for those with and without gout. Associations of gout (vs non-gout) with long-term prescription opioid exposure were quantified using univariable and then multivariable Cox proportional hazards regression, with the cumulative incidence plotted over time. In addition to accounting for age, sex, and enrollment year through matching, covariates examined in multivariable regression included race and ethnicity, BMI, smoking, urban or rural residence, CKD stage, and RDCI score. To further examine the impact of obesity-related comorbidities occurring more frequently in gout, the multivariable analysis was repeated after including the presence of baseline diabetes, hypertension, myocardial infarction, and cardiovascular disease (individual components of the RDCI). Univariable and multivariable associations between individual characteristics and long-term opioid exposure were also examined in separate analyses limited to gout. In addition to the covariates noted, additional covariates in this model included age, sex, enrollment year, a requirement of rheumatology care, and time-varying measures of sU control and ULT treatment. Allowing for the measurement of sU control and ULT treatment, follow-up time for this analysis began one-year post index. To account for diagnoses potentially associated with increased risk of opioid exposure, we performed three separate sensitivity analyses that included the aforementioned covariates in addition to the following baseline conditions: (1) mental health conditions (anxiety, depression, or posttraumatic stress disorder), (2) other chronic pain conditions (chronic headaches, chronic back pain, fibromyalgia, or osteoarthritis), or (3) all of the conditions included in 2 or 3. All analyses were completed using Stata v18 (StataCorp).

RESULTS

Participant characteristics. We identified 419,967 patients with gout and 3,609,382 matched controls (Figure 1). The cohort was overwhelmingly male (99%), 59% reported non-Hispanic White race and ethnicity, and patients had a mean age of approximately 68 years (Table 1). Race and ethnicity, smoking, and CKD status were more likely to be missing for patients without gout compared to those with gout. Patients with gout had higher BMI values and a greater frequency of comorbidity compared to those without gout. Among patients with gout, 8.9% required a rheumatology visit during the up to five-year follow-up period. Of the one-year follow-up periods assessed in patients with gout, 21% were characterized by inadequate sU control and 46% were characterized by ULT receipt.

Table 1. Characteristics of patients with and without gout*

	Without gout (n = 3,609,382)	With gout (n = 419,967)
Age, mean (SD), y	67.6 (11.7)	67.6 (11.6)
Sex, n (%)		
Male	3,579,838 (99)	416,548 (99)
Female	29,544 (1)	3,419 (1)
Race and ethnicity, n (%)		
White, non-Hispanic	2,131,557 (59)	264,077 (63)
Black, non-Hispanic	326,407 (9)	59,558 (14)
Other	227,673 (6)	29,659 (7)
Missing	923,745 (26)	66,673 (16)
Residence, n (%)		
Urban	2,336,825 (65)	263,488 (63)
Rural	1,264,075 (35)	155,041 (37)
Missing	8,482 (0)	1,438 (0)
Smoking status, n (%)		
Current	1,102,968 (31)	136,089 (32)
Former	1,218,500 (34)	174,483 (42)
Never	635,606 (18)	83,877 (20)
Missing	652,308 (18)	25,518 (6)
Body mass index, n (%)		
<20	51,072 (2)	3,052 (1)
20 to <25	439,279 (14)	26,852 (6)
25 to <30	1,211,827 (38)	124,227 (30)
30+	1,448,317 (46)	261,198 (63)
RDCI score, mean (SD)	1.15 (1.37)	1.83 (1.35)
RDCI components, n (%)		
Lung disease	264,830 (7)	40,267 (10)
Myocardial infarction,	65,708 (2)	13,444 (3)
Cardiovascular disease	710,233 (20)	141,446 (34)
Stroke	65,444 (2)	11,513 (3)
Hypertension	1,405,801 (39)	296,625 (71)
Diabetes	598,140 (17)	113,448 (27)
Fracture	10,998 (0.3)	1,295 (0.3)
Depression	310,863 (9)	45,496 (11)
Ulcer	81,811 (2)	13,781 (3)
Cancer	268,216 (7)	41,017 (10)
Chronic kidney disease, n (%)		
Stage 1 (>90 mL/min)	633,381 (18)	71,236 (17)
Stage 2 (60–89 mL/min)	1,312,655 (36)	177,803 (42)
Stage 3 (30–59 mL/min)	481,743 (13)	121,557 (29)
Stage 4–5 (<30 mL/min)	47,006 (1)	18,954 (5)
Missing	1,134,597 (31)	30,417 (7)

* All without gout vs with gout comparisons were significant ($P < 0.001$), with the exception of age, sex, and fracture prevalence. RDCI, Rheumatic Disease Comorbidity Index.

Long-term opioid exposure in patients with and without gout. Over a follow-up period of 17.3 million patient-years (mean follow-up 4.52 years), 6.9% of patients with gout received long-term prescription opioids compared to 3.8% of patients without gout (Table 2). The crude incidence (per 1,000 patient-years) of long-term opioid exposure was 16.0 (95% confidence interval [CI] 15.8–16.2) for gout versus 8.9 (95% CI 8.9–8.9) in patients without gout, corresponding to an unadjusted hazard ratio of 1.79 (95% CI 1.77–1.82) (Table 2). After accounting for covariates in a multivariable regression model, patients with gout were 30% more likely to receive long-term prescription opioids during follow-up (adjusted hazard ratio [aHR] 1.30; 95% CI 1.28–1.32) (Figure 2). The association of gout with long-term opioid exposure was not significantly changed following the addition of baseline diabetes, hypertension, myocardial infarction, and cardiovascular disease as covariates (aHR 1.32; 95% CI 1.30–1.34; data not shown). The aforementioned estimates (crude incidence in controls and aHR in patients with gout) translate to the occurrence of one additional case of long-term prescription opioid exposure for every 74 patients with gout seen over a five-year period. Sensitivity analyses that included baseline mental health conditions and/or other conditions associated with chronic pain as additional covariates did not meaningfully change these results (Supplemental Table 1).

Determinants of long-term opioid exposure in patients with gout. Univariable and multivariable associations between individual characteristics and gout-related long-term opioid exposure were examined (Table 3). In a multivariable regression model limited to gout, factors most strongly associated with an increased likelihood of long-term opioid exposure included female sex (aHR 1.29; 95% CI 1.14–1.46), underweight BMI (aHR 1.32; 95% CI 1.13–1.54), current smoking (aHR 1.62; 95% CI 1.55–1.68), a higher RDCI score (aHR 1.26; 95% CI 1.25–1.27), ULT receipt (aHR 1.30; 95% CI 1.27–1.34), and requirement of rheumatology consultation (aHR 1.60; 95% CI 1.54–1.66). Other factors less strongly associated with an increased risk of long-term opioid exposure included more recent index year, younger age, non-Hispanic Black race and ethnicity, rural residence, obese BMI, and former or missing smoking status. Factors associated with a lower risk of long-term opioid exposure included overweight BMI, the presence of CKD, and evidence of adequate sU control.

DISCUSSION

In this matched cohort study, we found that patients with gout were 30% more likely to be prescribed long-term opioid therapy when compared to patients without gout, after accounting for other factors influencing receipt. Among patients with gout, a number of factors were associated with a higher likelihood of opioid exposure, including sex, race, being underweight, obesity, comorbidity, lack of sU control, and requiring a rheumatology consultation.

To our knowledge, this is the first large-scale study to compare rates of long-term opioid exposure between patients with gout and those without gout, though prior reports have detailed the high prevalence of opioid treatment in the acute setting of gout flare. In a study examining US emergency department records between 2015 and 2017, Dalal and colleagues found that 28% of patients with gout received an opioid at discharge.[7] In another single-center study of consecutive patients with gout seen in the emergency department, >50% received an opioid prescription at discharge.[8] Although we cannot say for certain based on the results from our study that the use of opioids to manage

Table 2. The frequency, crude incidence rates, and unadjusted/adjusted HRs of long-term opioid exposure in patients with and without gout*

	Patients initiating long-term opioids, n (%)	Patient-years of follow-up	Crude incidence (95% CI) per 1,000 patient-years	Unadjusted HR (95% CI)	Adjusted HR (95% CI)[a]
Without gout	137,506 (3.8)	15,453,930	8.9 (8.9–8.9)	Reference	Reference
With gout	28,962 (6.9)	1,809,599	16.0 (15.8–16.2)	1.79 (1.77–1.82)	1.30 (1.28–1.32)

* Patients with gout (n = 419,967) and without gout (n = 3,609,382) were matched on birth year, sex, and Veterans Health Administration enrollment year. CI, confidence interval; HR, hazard ratio.
[a] Covariates included gout status, race and ethnicity, urban or rural residence, body mass index category, smoking status, Rheumatic Disease Comorbidity Index score, and chronic kidney disease stage (patients and controls matched for age, sex, and year of Veterans Health Administration enrollment).

acute pain from gout flares serves as a gateway to long-term opioid exposure, it is a concern that warrants further investigation. Based on national estimates and the crude incidence observed in this population, our results suggest that as many as 165,000 patients with gout in the United States transition to long-term opioid use each year.

In addition to acute symptoms from flare, emerging data suggest that gout may be associated with the development of generalized pain hypersensitivity, which could also contribute to higher rates of opioid exposure. One of every five patients with crystal-proven gout surveyed in a single outpatient rheumatology clinic demonstrated evidence of generalized pain hypersensitivity.[21] Although opioids offer robust analgesia in the setting of acute pain, evidence now suggests the possibility of opioids promoting paradoxical effects in gout, with these agents demonstrating the capacity to activate the NLRP3 inflammasome, the primary pathologic driver of gout-related inflammation.

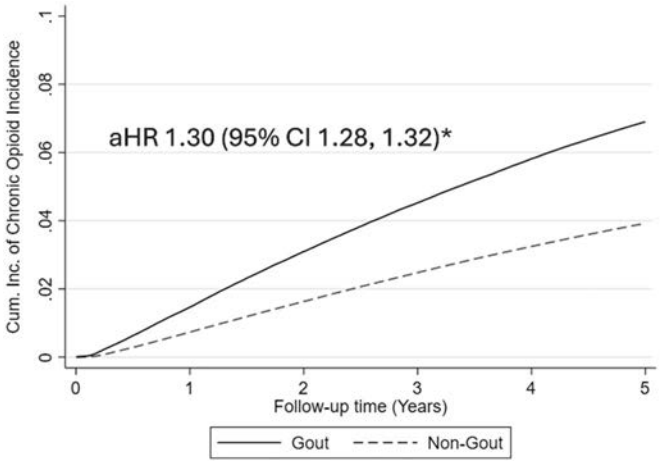

Figure 2. aHR and 95% CI estimated using multivariable Cox proportional hazards regression. Patients with gout and those without gout were matched for age, sex, and year of Veterans Health Administration enrollment. The model includes gout status, race and ethnicity, urban or rural residence, body mass index category, smoking status, Rheumatic Disease Comorbidity Index score, and chronic kidney disease stage. *aHR, adjust hazard ratio; CI, confidence interval; Cum. Inc., cumulative incidence.

NLRP3 activation leads to the activation and release of interleukin-1β (IL-1β) and IL-18, potentially worsening or prolonging the inflammation and pain associated with gouty arthritis.[22,23] Currently, international guidelines, including those from the American College of Rheumatology and EULAR, do not endorse opioids as a recommended therapy in gout, nor do these explicitly recommend against their use.[24,25] The findings of our study suggest future guidelines should recommend against opioid use in acute gouty arthritis given a higher risk of long-term opioid use in this population.

Beyond demonstrating a higher risk in gout, our analyses identified factors among those with gout that appear to influence the receipt of long-term opioids. Notably, some of the risk factors identified parallel those identified in populations without gout. For example, the temporal increase of long-term opioid exposure observed between the index years of 2000 and 2015 reflects trends reported from the National Health and Nutrition Survey over a similar time frame.[26] Likewise, patient characteristics of underweight, obesity, rural residence, and tobacco smoking have all been associated with a higher risk of receiving prescriptions for opioid analgesics in populations without gout.[11,27,28]

Reasons for the higher risk of long-term opioid exposure among those with greater cumulative comorbidity are unknown. Results from a secondary analysis that included obesity-related conditions suggest that these comorbidities are unlikely to serve as a meaningful source of confounding. It is possible that the use of prescription opioids in these patients reflects relative contraindications of other analgesics or anti-inflammatory agents used in gout, a question that warrants additional investigation. Underweight patients were also at higher risk for long-term opioid exposure for unclear reasons, though these findings are consistent with multiple studies, which have found lower BMI to be associated with higher rates of chronic pain.[29,30] In contrast to the higher risk associated with overall comorbidity burden, the incidence of long-term opioid exposure was lower in patients with CKD (a condition not included as a component of the RDCI). Lower use in these patients may reflect conventional guidance calling for marked dose reductions or avoidance of opioids in the context of significant renal dysfunction.[31]

Other risk factors, including female sex and being a person of color, align with disparities previously reported in gout. Compared

Table 3. Univariable and multivariable associations of patient characteristics with the receipt of long-term opioid prescriptions among patients with gout*

Characteristic	Unadjusted HR (95% CI)	Adjusted HR (95% CI)
Demographics		
Age (per 10 y)	0.79 (0.78–0.80)	0.80 (0.79–0.81)
Female sex	1.42 (1.27–1.58)	1.29 (1.14–1.46)
Calendar year (index)	1.00 (1.00–1.00)	1.02 (1.01–1.02)
Race and ethnicity		
Non-Hispanic White	Referent	Referent
Non-Hispanic Black	1.25 (1.21–1.29)	1.08 (1.05–1.12)
Other	1.03 (0.99–1.08)	1.00 (0.95–1.06)
Missing	0.69 (0.67–0.72)	0.82 (0.78–0.85)
Residence		
Urban	Referent	Referent
Rural	1.10 (1.08–1.13)	1.12 (1.09–1.15)
Missing	0.94 (0.77–1.15)	1.05 (0.84–1.32)
General health measures		
Body mass index		
Underweight (<20)	1.73 (1.52–1.98)	1.32 (1.13–1.54)
Normal (20–24.9)	Referent	Referent
Overweight (25–29.9)	0.89 (0.84–0.95)	0.88 (0.82–0.94)
Obese (≥30)	1.38 (1.31–1.46)	1.09 (1.02–1.16)
Smoking		
Never	Referent	Referent
Former	1.15 (1.10–1.19)	1.12 (1.08–1.17)
Current	2.04 (1.97–2.11)	1.62 (1.55–1.68)
Missing	1.06 (0.99–1.14)	1.11 (1.02–1.21)
RDCI score	1.23 (1.22–1.24)	1.26 (1.25–1.27)
Chronic kidney disease		
Stage 1	Referent	Referent
Stage 2	0.73 (0.71–0.75)	0.85 (0.82–0.88)
Stage 3	0.75 (0.73–0.78)	0.92 (0.88–0.96)
Stage 4 or 5	0.84 (0.79–0.90)	0.90 (0.83–0.97)
Missing	0.37 (0.35–0.40)	0.61 (0.57–0.66)
Gout-related factors		
Adequate sU control[a]	0.78 (0.76–0.81)	0.95 (0.93–0.98)
ULT receipt[a]	1.34 (1.31–1.38)	1.30 (1.27–1.34)
Rheumatology encounter	1.97 (1.91–2.04)	1.60 (1.54–1.66)

* All variables listed were included in the model. CI, confidence interval; HR, hazard ratio; RDCI, Rheumatic Disease Comorbidity Index; sU, serum urate; ULT, urate-lowering therapy.
[a] Time-varying covariates.

to men with gout, women with gout are less likely to receive ULT and are more likely to have comorbid CKD and hypertension.[3,32] Moreover, women with gout are more likely to present with atypical joint involvement and, as a possible result, experience longer diagnostic delays. Both veterans and nonveterans reporting non-Hispanic Black race and ethnicity are substantially more likely than individuals reporting non-Hispanic White race to have gout.[33,34] In addition to a higher prevalence, non-White patients with gout are also less likely to receive ULT than White patients and are less likely to achieve target sU level goals with treatment.[3,35] Differences in treatment outcome appear to persist even in the context of protocolized treat-to-target ULT and after accounting for other sociodemographic measures and medication adherence.[35] Given less favorable outcomes reported in these populations, findings herein suggest that previous disparities in gout may be compounded by disproportionate rates of long-term opioid exposure.

In the absence of validated measures of disease severity in administrative data, we operationalized the requirement for rheumatology consultation as a surrogate of gout severity. In previous work, Edwards and colleagues noted that rheumatologists are far more likely to provide care for patients with tophaceous or uncontrolled gout.[20] In this previous report, patients with tophaceous or uncontrolled gout were in turn more likely to receive ULT, more frequently accrued emergency department visits, and had higher rates of CKD and heart failure. In addition to observing a higher risk of long-term opioid exposure with ULT dispensing, perhaps owing to greater gout severity, we identified a modest (albeit statistically significant) association between better sU control and lower risk of long-term opioid exposure. These results may be impacted by imprecision in defining sU control given the low frequency of regular laboratory assessments and infrequent implementation of treat-to-target ULT in real-world settings such as the VA.[36] It is possible that the potential for sU control to mitigate long-term opioid exposure is underestimated in our study, as patients satisfying this definition may have experienced less severe gout and thus garnered less benefit from gout management.

Given restrictions in administrative data available, we were not able to reliably identify either the indication or specialty origin for index prescriptions, results that would inform future efforts to reduce long-term opioid exposures in this population. Additional limitations include the overwhelmingly male VA population examined, which could limit generalizability. As we were reliant on algorithms leveraging administrative data, misclassification of gout status and long-term opioid exposure is possible, although we would anticipate such nondifferential misclassification to bias these results toward the null. Notably, the definition of long-term opioid exposure deployed in our study differs from that recently used in a study of veterans following ambulatory surgery, which required more than 180 days of exposure over a one-year follow-up period.[37] Finally, inherent to observational research, unmeasured confounding is possible.

In conclusion, we found in the VHA that patients with gout were more likely than those without gout to be prescribed opioids, leading to long-term use, after accounting for other possible determinants of receipt. This study also identified several factors that appear to identify patients with gout at highest risk of receipt, factors that might ultimately be used in targeted efforts to reduce long-term prescription opioid use in this population.

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 Helget 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. Juraschek SP, Miller ER III, Gelber AC. Body mass index, obesity, and prevalent gout in the United States in 1988-1994 and 2007-2010. Arthritis Care Res (Hoboken) 2013;65(1):127–132.
2. Helget LN, England BR, Roul P, et al. Incidence, prevalence, and burden of gout in the Veterans Health Administration. Arthritis Care Res (Hoboken) 2021;73(9):1363–1371.
3. Chen-Xu M, Yokose C, Rai SK, et al. Contemporary prevalence of gout and hyperuricemia in the United States and decadal trends: The National Health and Nutrition Examination Survey, 2007-2016. Arthritis Rheumatol 2019;71(6):991–999.
4. Yokose C, McCormick N, Lu N, et al. Trends in prevalence of gout among US Asian adults, 2011-2018. JAMA Netw Open 2023;6(4): e239501.
5. Lim SY, Lu N, Oza A, et al. Trends in gout and rheumatoid arthritis hospitalizations in the United States, 1993-2011. JAMA 2016; 315(21):2345–2347.
6. Garg R, Sayles HR, Yu F, et al. Gout-related health care utilization in US emergency departments, 2006 through 2008. Arthritis Care Res (Hoboken) 2013;65(4):571–577.
7. Dalal DS, Mbuyi N, Shah I, et al. Prescription opioid use among patients with acute gout discharged from the emergency department. Arthritis Care Res (Hoboken) 2020;72(8):1163–1168.
8. Brunetti L, Vekaria J, Lipsky PE, et al. Treatment of acute gout flares in the emergency department: prescribing patterns and revisit rates. Ann Pharmacother 2022;56(4):422–429.
9. National Institute on Drug Abuse. Drug overdose deaths: facts and figures. Updated June 30, 2023. Accessed September 20, 2023. https://nida.nih.gov/research-topics/trends-statistics/overdose-death-rates
10. National Centers for Drug Abuse Statistics. Opioid epidemic: addiction statistics. Accessed September 20, 2023. https://drugabusestatistics.org/opioid-epidemic/
11. Brummett CM, Waljee JF, Goesling J, et al. New persistent opioid use after minor and major surgical procedures in US adults. JAMA Surg 2017;152(6):e170504.
12. Shah A, Hayes CJ, Martin BC. Characteristics of initial prescription episodes and likelihood of long-term opioid use - United States, 2006-2015. MMWR Morb Mortal Wkly Rep 2017;66(10):265–269.
13. US Department of Veterans Affairs. National Center for Veterans Analysis and Statistics. Updated April 14, 2021. Accessed November 3, 2021. https://www.va.gov/vetdata/Veteran_Population.asp
14. Quinn PD, Hur K, Chang Z, et al. Association of mental health conditions and treatments with long-term opioid analgesic receipt among adolescents. JAMA Pediatr 2018;172(5):423–430.
15. Dalbeth N, Dowell T, Gerard C, et al. Gout in Aotearoa New Zealand: the equity crisis continues in plain sight. N Z Med J 2018;131(1485): 8–12.
16. Thompson MD, Wu YY, Cooney RV, et al. Modifiable factors and incident gout across ethnicity within a large multiethnic cohort of older adults. J Rheumatol 2022;49(5):504–512.
17. England BR, Sayles H, Mikuls TR, et al. Validation of the Rheumatic Disease Comorbidity Index. Arthritis Care Res (Hoboken) 2015; 67(6):865–872.
18. Dolomisiewicz A, Ali H, Roul P, et al. Updating and validating the Rheumatic Disease Comorbidity Index to incorporate ICD-10-CM diagnostic codes. Arthritis Care Res (Hoboken) 2023;75(10):2199–2206.
19. West AN, Lee RE, Shambaugh-Miller MD, et al. Defining "rural" for veterans' health care planning. J Rural Health 2010;26(4):301–309.
20. Edwards NL, Schlesinger N, Clark S, et al. Management of gout in the United States: a claims-based analysis. ACR Open Rheumatol 2020; 2(3):180–187.
21. Ten Klooster PM, Kraiss JT, Munters R, et al. Generalized pain hypersensitivity and associated factors in gout. Rheumatology (Oxford) 2022;61(9):3640–3646.
22. Zare N, Sharafeddin F, Montazerolghaem A, et al. NLRs and inflammasome signaling in opioid-induced hyperalgesia and tolerance. Inflammopharmacology 2024;32(1):127–148.
23. Liu YR, Wang JQ, Li J. Role of NLRP3 in the pathogenesis and treatment of gout arthritis. Front Immunol 2023;14:1137822.
24. FitzGerald JD, Dalbeth N, Mikuls T, et al. 2020 American College of Rheumatology guideline for the management of gout. Arthritis Care Res (Hoboken) 2020;72(6):744–760.
25. Richette P, Doherty M, Pascual E, et al. 2016 updated EULAR evidence-based recommendations for the management of gout. Ann Rheum Dis 2017;76(1):29–42.
26. Mojtabai R. National trends in long-term use of prescription opioids. Pharmacoepidemiol Drug Saf 2018;27(5):526–534.
27. Prunuske JP, St Hill CA, Hager KD, et al. Opioid prescribing patterns for non-malignant chronic pain for rural versus non-rural US adults: a population-based study using 2010 NAMCS data. BMC Health Serv Res 2014;14:563.
28. Stokes A, Lundberg DJ, Sheridan B, et al. Association of obesity with prescription opioids for painful conditions in patients seeking primary care in the US. JAMA Netw Open 2020;3(4):e202012.
29. Yamada K, Kubota Y, Iso H, et al. Association of body mass index with chronic pain prevalence: a large population-based cross-sectional study in Japan. J Anesth 2018;32(3):360–367.
30. Chen C, Winterstein AG, Fillingim RB, et al. Body weight, frailty, and chronic pain in older adults: a cross-sectional study. BMC Geriatr 2019;19(1):143.
31. Dean M. Opioids in renal failure and dialysis patients. J Pain Symptom Manage 2004;28(5):497–504.
32. Dirken-Heukensfeldt KJ, Teunissen TA, van de Lisdonk H, et al. "Clinical features of women with gout arthritis." A systematic review. Clin Rheumatol 2010;29(6):575–582.
33. Helget LN, England BR, Roul P, et al. Incidence, prevalence, and burden of gout in the Veterans Health Administration. Arthritis Care Res (Hoboken) 2021;73(9):1363–1371.
34. McCormick N, Lu N, Yokose C, et al. Racial and sex disparities in gout prevalence among US adults. JAMA Netw Open 2022;5(8): e2226804.
35. Helget LN, O'Dell JR, Newcomb JA, et al. Determinants of achieving serum urate goal with treat-to-target urate-lowering therapy in gout. Arthritis Rheumatol 2024;76(4):638–646.
36. Coburn BW, Michaud K, Bergman DA, et al. Allopurinol dose escalation and mortality among patients with gout: a national propensity-matched cohort study. Arthritis Rheumatol 2018;70(8):1298–1307.
37. Liu JY, Soybel DI. Persistent and chronic opioid use after ambulatory surgery in US veterans (2011-2018). Surgery 2024;176(6):1688–1696.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 78, No. 3, March 2026, pp 352–361
DOI 10.1002/acr.25623

International Investigation of the Gut-Lung Axis in Systemic Sclerosis–Interstitial Lung Disease

Kristofer Andréasson,[1] Arissa Young,[2] Swapna Joshi,[2] Jennifer S. Labus,[2] Andrea Hsiu Ling Low,[3] Vanessa Smith,[4] Zsuzsanna McMahan,[5] Susanna Proudman,[6] Antonia Valenzuela,[7] Phoebe Hunter,[6] Grace Hyun Kim,[2] Gracijela Bozovic,[1] Jonathan Goldin,[2] Ezinne Aja,[2] Jonathan P. Jacobs,[2] and Elizabeth R. Volkmann[2]

Objective. Mounting evidence supports an association between the intestinal microbiota and diverse pulmonary pathologies (ie, gut-lung axis). Although intestinal dysbiosis is a feature of systemic sclerosis (SSc), no prior studies have investigated the relationship between intestinal microbiota and SSc-associated interstitial lung disease (ILD) in a multinational cohort. This study aimed to characterize the intestinal microbiota of SSc-ILD and determine whether specific bacterial species and functional pathways are associated with ILD severity.

Methods. Patients with SSc with and without ILD from seven SSc Centers across five continents provided a stool sample. Shotgun metagenomic sequencing was performed using the Illumina NovaSeq 6000 to characterize microbial composition at the species level. Quantitative image analysis of high-resolution computed tomography scans of the chest was used to measure radiologic extent of ILD (QILD). Multivariate sparse partial least squares analyses were employed to identify a species signature of ILD and to determine whether specific species and functional pathways are associated with QILD.

Results. Among 285 participants (mean disease duration of 9.8 years), 62.5% had ILD. In a multivariate analysis of all participants, patients with ILD had a unique microbial signature compared to those without ILD characterized by increased abundance of candidate pathobiont species. In a subgroup of participants with SSc-ILD (n = 103), specific bacterial species and functional pathways were associated with QILD.

Conclusion. This multicenter study demonstrates that distinct intestinal bacterial species are linked to the presence and radiologic extent of ILD in SSc. These species and/or their metabolic products may influence ILD pathogenesis and represent novel treatment targets.

INTRODUCTION

Systemic sclerosis (SSc) is a complex autoimmune disease with multiple clinical phenotypes and diverging molecular and pathologic characteristics.[1] The incidence and prevalence of autoimmune diseases, including SSc, have increased over the past decades with notable socioeconomic, seasonal, and regional variations,[2] underscoring the role of the environment in disease pathogenesis. Moreover, the rise in autoimmune diseases has occurred in a time frame too short to be solely attributed to genetic drift or natural selection,[3] further supporting environmental factors as drivers of disease.

This was an independent, investigator-initiated study supported by Boehringer Ingelheim Pharmaceuticals, Inc. Dr Andréasson's work was supported by Ulla och Roland Gustafssons Donationsfond, Alfred Österlunds Stiftelse, and Anna-Greta Crafoords Stiftelse. Drs Andréasson and Volkmann's work was supported by Boehringer Ingelheim Pharmaceuticals, Inc. Dr Volkmann's work was supported by the National Heart, Lung, and Blood Institute, NIH (grant K23-HL-150237). Dr Volkmann also received anonymous donation from a patient.

[1]Kristofer Andréasson, MD, PhD, Gracijela Bozovic, MD, PhD: Lund University, Lund, Sweden; [2]Arissa Young, MD, Swapna Joshi, PhD, Jennifer S. Labus, PhD, Grace Hyun Kim, PhD, Jonathan Goldin, MD, PhD, Ezinne Aja, PhD, Jonathan P. Jacobs, MD, PhD, Elizabeth R. Volkmann, MD, MS: University of California, Los Angeles; [3]Andrea Hsiu Ling Low, BMedSci, BMBS, MCI: Singapore General Hospital, SingHealth Duke-National University of Singapore Academic Medical Centre, Singapore; [4]Vanessa Smith, MD: Ghent University, Ghent, Belgium; [5]Zsuzsanna McMahan, MD, MS: The University of Texas Health Science Center at Houston; [6]Susanna Proudman, MBBS, Phoebe Hunter: Royal Adelaide Hospital and University of Adelaide, Adelaide, South Australia, Australia; [7]Antonia Valenzuela, MD, MS: Pontificia Universidad Católica de Chile, Santiago, Chile.

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.25623).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25623.

Address correspondence via email to Elizabeth R. Volkmann, MD, MS, at evolkmann@mednet.ucla.edu.

Submitted for publication January 29, 2025; accepted in revised form July 24, 2025.

SIGNIFICANCE & INNOVATIONS

- Numerous studies (although none within systemic sclerosis [SSc]) have demonstrated that the gastrointestinal (GI) microbiome plays an important role in immune system homeostasis, both within the GI tract and in distant organs, including the lungs.
- Dysbiosis (ie, alterations in the GI microbiome) is a feature of diverse autoimmune diseases, including SSc; however, prior studies in SSc have not evaluated the relationship between specific bacterial species and SSc features, such as interstitial lung disease (ILD).
- This is the first study to illustrate that patients with SSc-ILD have a unique GI microbial profile by applying shotgun metagenomics. Using data from patients with SSc from five continents, the present study demonstrated that certain bacterial species are associated with the severity of ILD, even after adjusting known confounders.
- Understanding the connection between the gut-lung axis in SSc-ILD may improve our knowledge of the pathogenesis of this condition and reveal novel treatment strategies aimed at improving dysbiosis.

The gastrointestinal (GI) microbiome represents an acquired genome, which is dynamic and shaped by the environment (eg, lifestyle, diet, geography, medications, etc). The GI microbiome plays a central role in both immune homeostasis and food digestion. It can also affect immune function in distant organs, including the lungs.[4,5] For example, several studies have identified bidirectional relationships between GI microbiota and pulmonary diseases, including asthma[6] and chronic obstructive pulmonary disease.[7] The purported gut-lung axis arises from complex interactions among microbes resident in the GI tract, their metabolic products, circulating immune cells, and lung parenchyma.[5]

Alterations in the GI microbiome are a feature of SSc,[8,9] and a recent study demonstrated microbial community differences in patients with SSc with versus without interstitial lung disease (ILD) in patients from two geographic regions (United States and Sweden).[10] However, this study did not evaluate whether microbial composition was associated with ILD severity. Moreover, prior GI microbiome research in SSc was based on the 16S ribosomal RNA (rRNA) amplicon sequencing method, which has poor species-level resolution compared with whole genome shotgun sequencing (WGS).[11] In addition, these early SSc microbiome studies included relatively few samples, limiting the ability to control for confounders.[8]

The present study aimed to investigate the role of GI microbiome in SSc-ILD in an international, multicenter cohort using WGS. We hypothesized that specific bacterial species are associated with the presence or absence of ILD across cohorts and that certain species and functional pathways are associated with the severity of ILD as measured by the quantitative radiologic extent of ILD, even after adjusting for known confounders. We specifically sought to include patients with SSc from different geographic regions with varying environmental exposures and dietary habits, to discover a shared set of species associated with ILD, independent of these external factors. Improving our understanding of the role of the GI microbiome in ILD may reveal novel therapeutic strategies to modify the GI microbiome through diet, personalized prebiotics or probiotics, selective antibiotics, and/or specific bacterial transfers.

PATIENTS AND METHODS

Study participants. Adult patients fulfilling the 2013 American College of Rheumatology/EULAR Classification Criteria for SSc[12] were consecutively recruited from the rheumatology clinics of the University of California, Los Angeles (UCLA), United States; Lund University (LU), Sweden; Singapore General Hospital, Singapore; Johns Hopkins University, United States; Ghent University, Belgium; University of Adelaide, Australia; and Pontificia Universidad Católica de Chile, Chile. Study exclusion criteria included inflammatory bowel disease, current/prior GI malignancy, use of any antibiotics within four weeks of the stool collection, chronic use of antibiotics (defined as the use of antibiotics for any indication more than three times in the preceding year), any prior GI surgery (except for a remote history of appendectomy), and a history of fecal microbial transplantation. Patients were allowed to remain on proton pump inhibitors (PPIs) to minimize the risk of unnecessary morbidity during the study. The institutional review boards of all study sites approved the study protocol, and written and informed consent was obtained from each participant.

Clinical assessments. Clinical features of the participants with SSc were evaluated at the time of the stool collection (Table 1). Disease duration was defined based on the onset of the first non-Raynaud symptom attributable to SSc. Immunomodulatory therapy use was based on any use of immunomodulatory medications up until the date of the stool collection. High-resolution computed tomography (HRCT) of the chest was used to detect the presence of ILD. The presence of all other disease features was defined based on a physician's clinical diagnosis identified through chart review and/or confirmed by the site principal investigator based on objective testing (eg, a diagnosis of small intestinal bacterial overgrowth [SIBO] was based primarily on lactulose breath testing).

Radiologic evaluation. HRCT scans of the chest performed within four months of stool collection for the UCLA and LU study participants underwent quantitative image analysis (QIA) to calculate the radiologic extent of ILD (QILD) whole lung[13]

Table 1. Baseline characteristics of participants with systemic sclerosis with and without ILD*

	No ILD (n = 107)	ILD present (n = 178)	All patients (n = 285)
Lund University, Sweden	18/107 (16.82)	64/178 (35.96)	82/285 (28.77)
University of California, Los Angeles, United States	37/107 (34.58)	55/178 (30.9)	92/285 (32.28)
Singapore General Hospital, Singapore	19/107 (17.76)	34/178 (19.1)	53/285 (18.6)
Johns Hopkins University, United States	14/107 (13.08)	10/178 (5.62)	24/285 (8.42)
Ghent University, Belgium	10/107 (9.35)	5/178 (2.81)	15/285 (5.26)
Pontificia Universidad Católica de Chile	2/107 (1.87)	8/178 (4.49)	10/285 (3.51)
Royal Adelaide Hospital, Australia	7/107 (6.54)	2/178 (1.12)	9/285 (3.16)
Age	53.47 ± 13.55	55.46 ± 12.44	54.71 ± 12.88
BMI	24 ± 4.96	25.54 ± 4.5	24.95 ± 4.73
Female sex	92/107 (85.98)	146/178 (82.02)	238/285 (83.51)
SIBO	16/106 (15.09)	22/173 (12.72)	38/279 (13.62)
PPI use	68/99 (68.69)	140/159 (88.05)	208/258 (80.62)
Probiotic use	20/107 (18.69)	27/178 (15.17)	47/285 (16.49)
Immunomodulatory therapy[a]	49/98 (50)	129/177 (72.88)	178/275 (64.73)
Anti–Scl-70 positive	21/100 (21)	65/168 (38.69)	86/268 (32.09)
Anticentromere positive	41/101 (40.59)	12/155 (7.74)	53/256 (20.70)
Anti-RNA polymerase III positive	10/72 (13.89)	15/119 (12.61)	25/191 (13.09)
Disease duration based on first non-Raynaud symptom	10.17 ± 9.09	9.61 ± 7.87	9.81 ± 8.33
Diffuse cutaneous disease	32/103 (31.07)	71/176 (40.34)	103/279 (36.92)

* Values represent n/N (%) or mean ± SD. BMI, body mass index; ILD, interstitial lung disease; PPI, proton pump inhibitor; SIBO, small intestinal bacterial overgrowth.
[a] Present or past immunomodulatory therapy (eg, mycophenolate, azathioprine, methotrexate, cyclophosphamide, rituximab).

(details provided in Supplementary Material). The QILD score represents the summation of all abnormal patterns of ILD, including fibrosis (ie, parenchymal replacement with connective tissue, architectural distortion, and volume loss), ground glass opacity (ie, areas with pure ground glass opacification defined as increased opacity still transparent for underlying bronchial and vascular structures with preserved architecture), and honeycomb cysts.[14] Each pixel on the HRCT image can only have one label, representing the dominant feature assigned to that pixel. Therefore, the scores for fibrosis, ground glass opacity, and honeycomb cysts do not overlap. A higher QILD score indicates greater radiologic extent of ILD.

Fecal microbiome analysis. Participants at all sites provided stool specimens, which were immediately frozen at –80°F, as previously described.[15] Non-UCLA samples were shipped overnight on dry ice to UCLA for further processing.[15] Shotgun metagenomics were performed using the Illumina NovaSeq 6000. Shotgun reads were input into MetaPhlAn4 for taxonomic identification of species for compositional analysis[16] and subsequently underwent center log-ratio transformation. Samples were filtered to retain species with at least 10% nonzero counts in both the UCLA and LU cohorts as these two sites accounted for more than half of the total study participants. This resulted in 379 and 341 species identified in the UCLA and LU samples, respectively. Of these, 257 species were present in both sites and were included in the present analyses. Functional analyses were also performed by inputting shotgun reads into HUMAnN 3.0 for annotation of gene families and identification of the microbial source of these genes using Kyoto Encyclopedia of genes and genome (KEGG) orthology.[17] Functional profiling from metagenomic sequencing provides information on the functional activities of microbial communities and can link these activities to clinical features.[18] All samples from all sites were analyzed simultaneously at UCLA to avoid any batch effects.

Statistical analysis. *Primary outcome: taxonomic profiling of SSc-ILD.* General linear models were applied to identify differentially abundant species based on ILD presence, controlling for factors or covariates known to affect intestinal microbial composition in prior studies. These factors were determined a priori and included body mass index (BMI),[19] current PPI use (yes/no),[20] current supplemental probiotic use (yes/no),[21] current or prior immunomodulatory therapy (yes/no),[22] presence of SIBO (yes/no),[23] and study site.[10]

Because species do not function in isolation, but instead interact with each other, we investigated whether linear combinations of species abundance differentiated ILD presence and absence, as well as ILD severity, using sparse partial least squares analyses with Mixomics v 6.30.4 in R v4.42. Sparse partial least squares–discriminate analysis (sPLS-DA) was used to identify a species signature of SSc-ILD across all sites, adjusting for the aforementioned covariates.

Given that the clinical phenotype of ILD varies among patients with SSc,[24] we also investigated the relationship between microbial species and QILD scores. Multivariate, general

linear models were applied to identify differentially abundant species based on QILD scores, adjusting for the aforementioned covariates. Sparse partial least squares regression (sPLS-R) was used to determine whether a species signature was associated with the severity of ILD based on the QILD score. For additional details on this approach, please see the Supplementary Material.

Secondary outcome: functional profiling of SSc-ILD. To characterize the metabolic potential of microbial communities and its members, multivariate, general linear models were applied to identify KEGG pathways[25] associated with ILD presence, as well as ILD severity, controlling for BMI, PPI use, current probiotic use, current or prior immunomodulatory therapy use, presence of SIBO, and study site.

For all statistical analyses, we considered $P < 0.05$ as the threshold for reporting. Given that this was the first study to examine the relationship between GI microbiome and ILD in SSc, we consider our analyses as exploratory and hypothesis generating and, therefore, did not correct for multiple inference testing. We also computed effect size estimates, Cohen's *d* for mean differences, and standardized beta for association analyses for the primary purpose of informing power analyses for future studies in this area. All tests were two-sided and performed in R.

RESULTS

Participant characteristics. Among 285 participants with SSc, 178 (62.5%) had any ILD on HRCT. The mean age ± SD was 54.7 ± 13 years, and the mean disease duration was 9.8 ± 8.3 years. There were no significant differences in disease duration or in percentages of female patients (82% vs 86%) or patients with SIBO (13% vs 15%) in patients with versus without ILD; however, patients with ILD were more likely to use PPIs (88% vs 69%) or have received current or past immunomodulatory therapy (73% vs 50%) than patients without ILD (Table 1). Among the subgroup of patients with ILD who were positive for anticentromere antibody, the mean ± SD for the forced vital capacity percentage predicted and diffusing capacity for carbon monoxide percentage predicted was 77.8 ± 24.7 and 66.21 ± 22.5, respectively. The mean ± SD for the QILD score for the whole lung was 15.5% ± 12.9% (range of 3%–51.1%).

Intestinal microbiota signature of ILD. Among 257 species, the relative abundance of 10 distinct species was associated with the presence of ILD among all study sites, after adjusting for BMI, current PPI use, current probiotic use, current or prior immunomodulatory therapy, presence of SIBO, and study site (Supplementary Table S1). The abundance of certain species was higher in patients with ILD, whereas the abundance of other species was lower in patients with ILD. Absolute effect sizes for individual species were small to moderate, with Cohen's *d* ranging from 0.26 to 0.33.

The sPLS-DA produced one component solution composed of 10 species (Figure 1). Five species were less abundant in patients with SSc with ILD, and five species were more abundant in patients with SSc with ILD, including species deemed pathobiont in other autoimmune disease states (ie, *Collinsella*

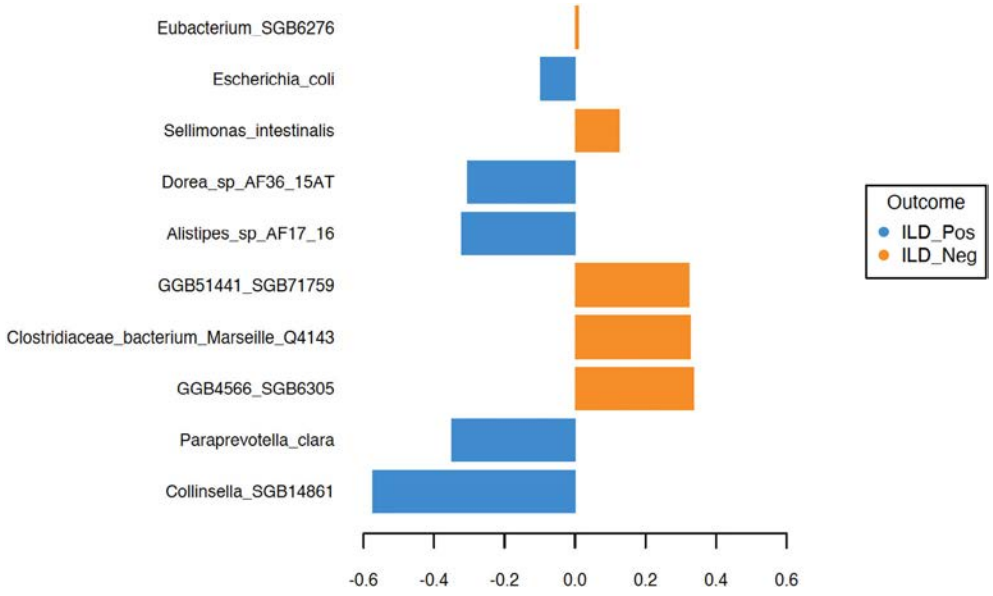

Figure 1. Feature loadings on the species signature differentiating ILD presence versus absence based on sparse partial least squares–discriminate analysis in all participants (N = 285). Species with higher loadings (ie, weights) contribute the most to the signature score and are ranked from highest (bottom) to lowest (top). For each species, orange bars signify that the species is more abundant in patients without ILD, and the blue bars signify that the species is more abundant in patients with ILD. Unknown bacterial species are indicated as species genome bins (SGB) and genus genome bins (GGB). ILD, interstitial lung disease; Neg, negative; Pos, positive.

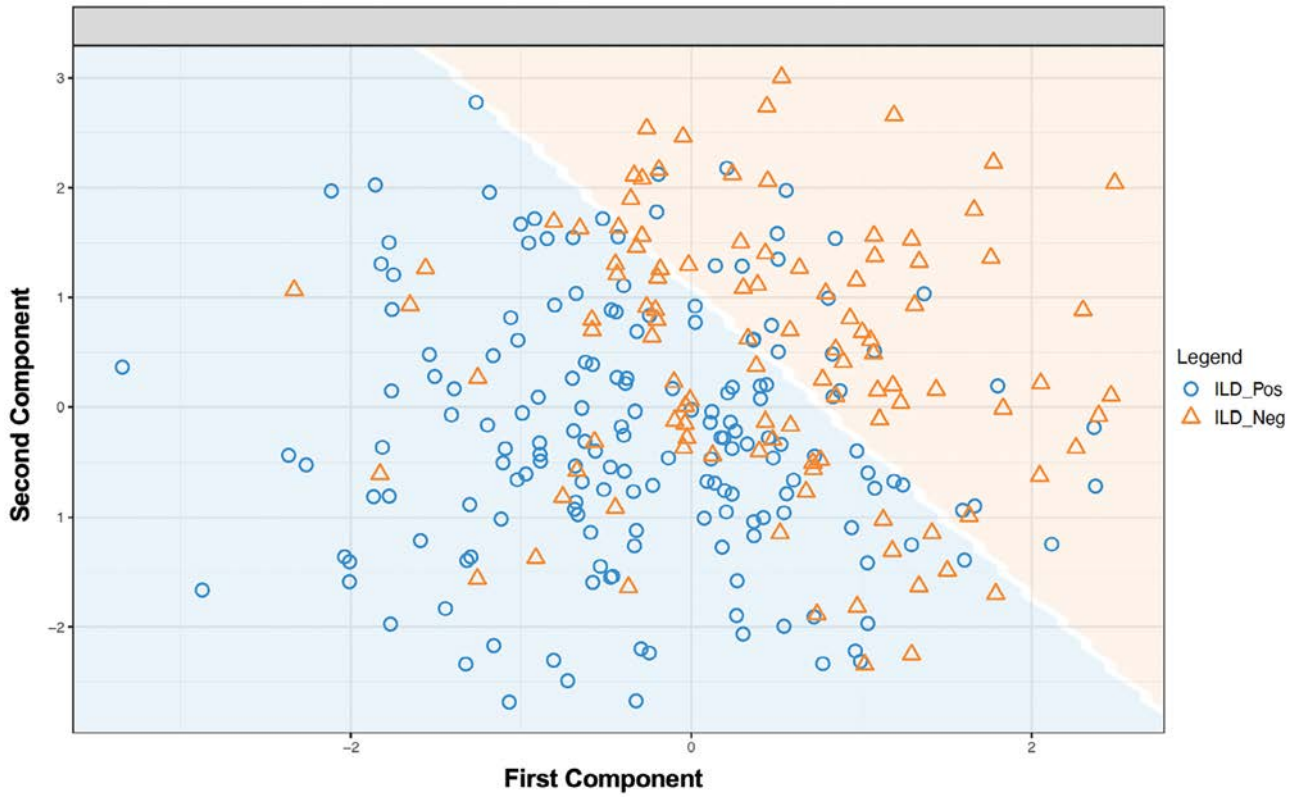

Figure 2. Visualization of discrimination of samples by species signature in the entire cohort based on sparse partial least squares–discriminate analysis (N = 285). The first component (x-axis) is the first signature; the second signature (y-axis) is provided for visualization purpose only. Each symbol represents a patient sample. Orange triangles present patients without ILD, and blue circles represent patients with ILD. ILD, interstitial lung disease; Neg, negative; Pos, positive. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25623/abstract.

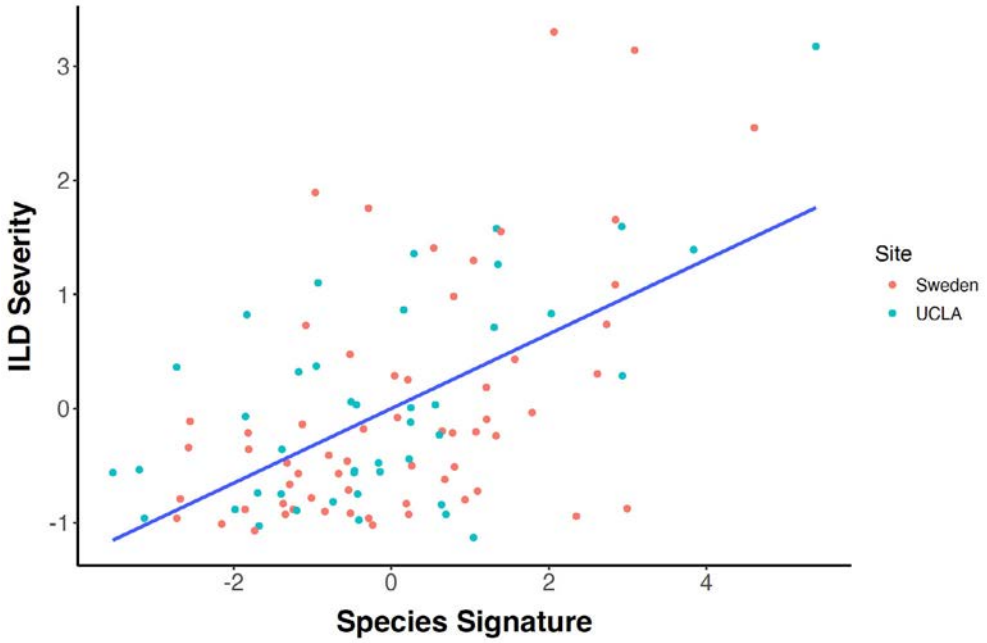

Figure 3. Scatterplot of individual scores for the species signature and ILD severity based on QILD scores from sparse partial least squares regression analysis (n = 103). Blue line reflects the regression line showing the significant association, even after adjusting for body mass index, proton pump inhibitor use, probiotic use, immunomodulatory use, small intestinal bacterial overgrowth, and study site (r = 0.56; 95% confidence interval 0.41–0.68). ILD, interstitial lung disease; QILD, quantitative image analysis to calculate the radiologic extent of ILD; UCLS, University of California, Los Angeles. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25623/abstract.

SGB14861[26] and *Escherichia coli*[27]). Many of these species overlapped with the 10 nominally differentially abundant species mentioned above. Figure 2 illustrates the discrimination of patient samples by species signature. An independent *t*-test demonstrated significant group differences based on the ILD derived species signature ($t_{228} = -6.05$; $P < 0.0001$) with a robust effect size (Cohen's $d = 0.74$), suggesting that the interaction between species accounts for more variance and larger effect size differences than individual species.

Species associated with radiologic features of ILD. Among 103 participants with SSc-ILD from UCLA (n = 42) and LU (n = 61) with HRCT scans available and amenable to QIA, the abundance of 18 bacterial species correlated with QILD scores, after adjusting for BMI, current PPI use, current probiotic use, current or prior immunomodulatory therapy, presence of SIBO, and study site (Supplementary Table S2). The absolute standardized beta for these species ranged from 0.20 to 0.31 for the QILD models.

The sPLS-R analysis revealed one linear combination of species showing large effect size association with QILD scores. Specifically, higher scores on this species signature were associated with higher QILD scores ($r = 0.56$; 95% CI 0.41–0.68; $t_{101} = 6.94$; $P < 0.0001$) (Figure 3). As shown in Figure 4, this signature comprised 25 species with the greatest contribution from *Dysosmobacter welbionis*, *GGB3005_SGB3996*, and *Anaeromassilibacillus sp An250* (each positively associated with QILD scores, ie, more severe ILD) and *Anaerobutyricum hallii*, *Bifidobacterium adolescentis*, and *Streptococcus parasanguinis* (each negatively associated with QILD scores, ie, less severe ILD). Interestingly, species deemed commensal in other disease states (eg, *B adolescentis*[28]; Figure 5A) were less abundant in patients with more severe ILD, whereas species deemed pathobiont (eg, *Anaeromassilibacillus sp An250*; Figure 5B) were more abundant in patients with more severe ILD. All but two of the 18 species associated with QILD scores in the individual multivariate analyses were present in the species signature for ILD severity. Figure 5 and Supplementary Figures S1–S14 show the predicted abundances of species associated with QILD scores in both the multivariate analyses and the sPLS-R analysis. These estimates were derived from general linear models incorporating QILD scores, BMI, PPI use, probiotic use, immunomodulatory therapy use, SIBO, and study site.

Exploratory analysis: functional pathways. None of the functional pathways were associated with ILD presence versus absence in the entire multicenter cohort. However, 15 distinct functional pathways were associated with QILD scores, 6 of which were positively associated with QILD scores and 9 of which were negatively associated with QILD scores (Supplementary Table S3). We subsequently identified the taxonomic source of these pathways (ie, bacterial species associated with these pathways), and we discovered that a number of the bacterial species associated with these functional pathways were also associated with QILD scores in the present study (Supplementary Table S4).

DISCUSSION

ILD affects the majority of patients with SSc and represents a leading cause of morbidity and mortality.[29] Existing therapies for SSc-ILD fail to attenuate the course of ILD in many patients,[30] and an improved understanding of the environmental factors that drive disease pathogenesis is paramount to the discovery of new treatments for this often-fatal condition. The present study examined the relationship between the GI microbiota and SSc-ILD in an international, multicenter cohort to further our understanding of the gut-lung axis in SSc, two organ systems that dynamically and continuously interact with the environment.

First, we observed that patients with SSc-ILD from five continents had a unique GI microbial signature compared with patients with SSc without ILD, suggesting that specific bacterial species and/or their metabolic products participate in ILD pathogenesis. Second, we observed that specific bacterial species were associated with the severity of ILD as measured by the quantitative radiologic extent of ILD. Finally, this study identified functional pathways, by which these bacteria may exert their role in the pathogenesis of SSc-ILD.

A notable finding of this study is that patients with SSc from geographically distinct areas of the world had a common repertoire of species that were specifically associated with ILD. The abundance of these species was increased across all patients with SSc-ILD, regardless of the study site. Although patients from each country had differences in overall GI microbial composition (ie, beta diversity [Supplementary Figure S15]), specific bacterial species were associated with ILD presence, even after adjusting for potential confounders. To our knowledge, this is the first study to link GI bacterial species with SSc-ILD using WGS.

Prior studies on the GI microbiome in various ILDs employed 16S rRNA sequencing,[31] which has inherent taxonomic classification challenges due to limited capacity to differentiate closely related species.[32] These prior studies demonstrated intestinal microbial community differences in patients with SSc with and without ILD.[10,33] Another study found an association between GI inflammation and SSc-ILD.[34] In the latter study, patients with SSc-ILD had higher fecal calprotectin levels compared with patients with SSc without ILD.[34] Fecal calprotectin levels are strongly associated with SIBO in patients with SSc.[35,36] Collectively, these studies support the hypothesis that dysbiosis in the GI tract is associated with ILD, and specific gut bacteria, bacterial components, and/or their metabolites may modulate immune function in the lung. Although the cross-sectional nature of the present study precludes our ability to determine causation, the finding that specific bacterial species were more abundant in

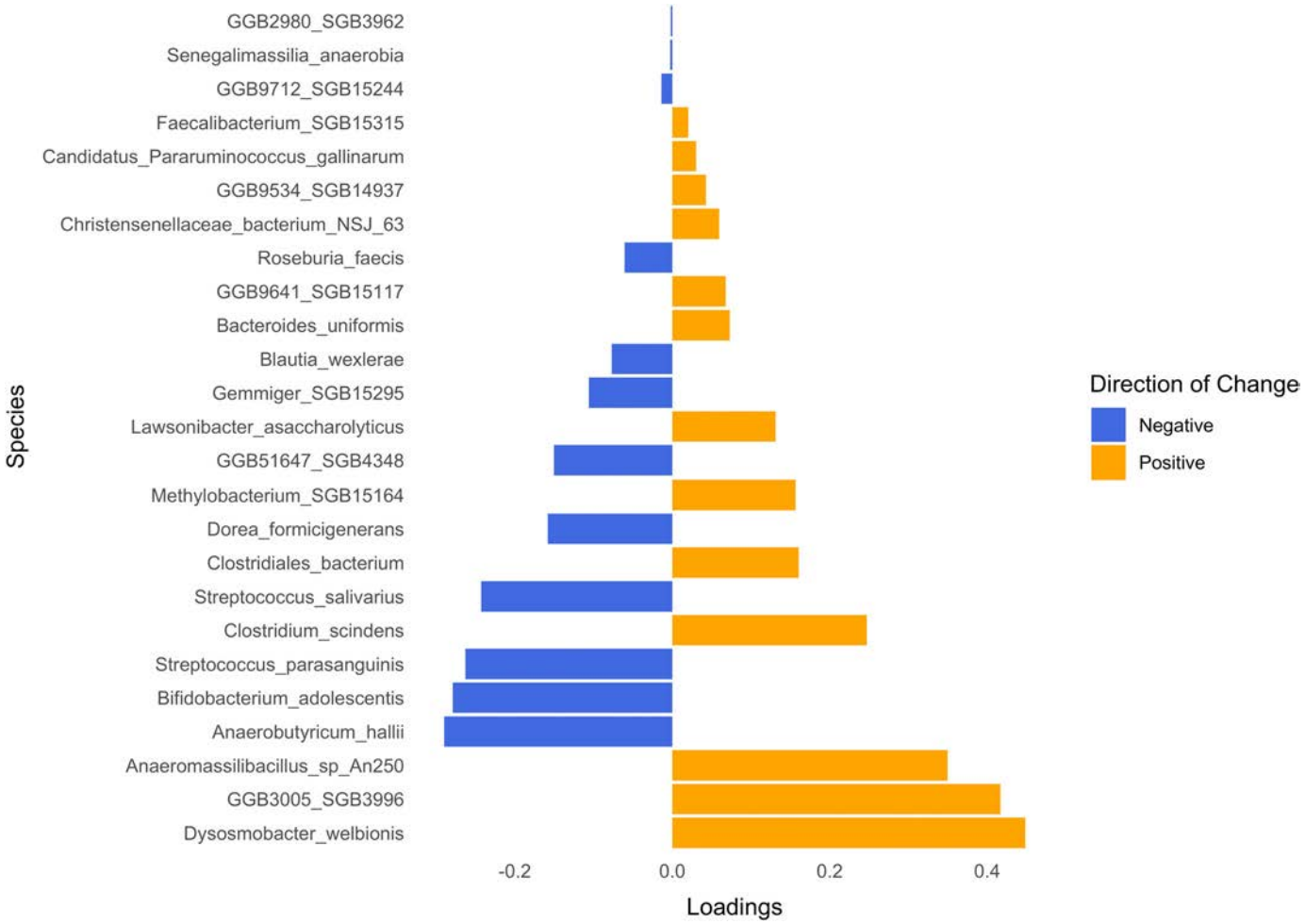

Figure 4. Feature loadings on the species-derived ILD severity signature from sparse partial least squares regression analysis (n = 103). Features are ranked from highest (bottom) to lowest (top); species with higher loadings (ie, weights) contribute the most to the signature score. For each species, positive and negative loadings reflect positive (orange bars) and negative association (blue bars) between that species and QILD scores (ie, radiologic extent of ILD). ILD, interstitial lung disease; QILD, quantitative image analysis to calculate the radiologic extent of ILD. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25623/abstract.

patients with ILD and other species were less abundant in patients with ILD, across all study sites, provides evidence to support a link between these two organ systems. Although caution is needed when labeling specific bacterial species as pathobiont or commensal, we did find it intriguing that certain species deemed pathobiont in other disease states were more abundant in patients with SSc with ILD, as well as in patients with more severe ILD.

Another notable finding of the present study was that in a subgroup of patients with SSc-ILD who underwent HRCT imaging of the chest at the time of stool collection, the abundance of specific bacterial species correlated with the radiologic extent of ILD. These findings suggest that unique intestinal bacterial species may associate with distinct clinical phenotypes of SSc-ILD. Given the clinical heterogeneity of ILD in SSc, it is not surprising that increased abundance of some bacterial species was associated with more severe ILD, while increased abundance of other bacterial species was associated with less severe ILD. For example, increased abundance of *B adolescentis* was associated with lower QILD score (Figure 5A). A prior study demonstrated that prophylactic oral administration of *B adolescentis* reduced protein expression of interleukin-1β, tumor necrosis factor α (TNF-α), and phosphorylated NF-κB p65 in mice exposed to chronic restraint stress.[37] Higher levels of TNF-α and receptor activator of NF-κB in bronchoalveolar lavage fluid were independently associated with higher quantitative lung fibrosis scores in a prior study of patients with SSc-ILD.[38] Although a myriad of factors may affect TNF-α and NF-κB levels in SSc-ILD, specific bacterial species and/or their metabolic products may play a role in modulating pulmonary inflammation.

In contrast, increased abundance of *Clostridium scindens* was associated with increased QILD scores (Supplementary Figure S12). This species produces secondary bile acids. Reflux of stomach content, including bile acids, is common in patients with SSc with esophageal dysmotility and may induce lung injury and accelerate ILD progression.[39] Prior studies have

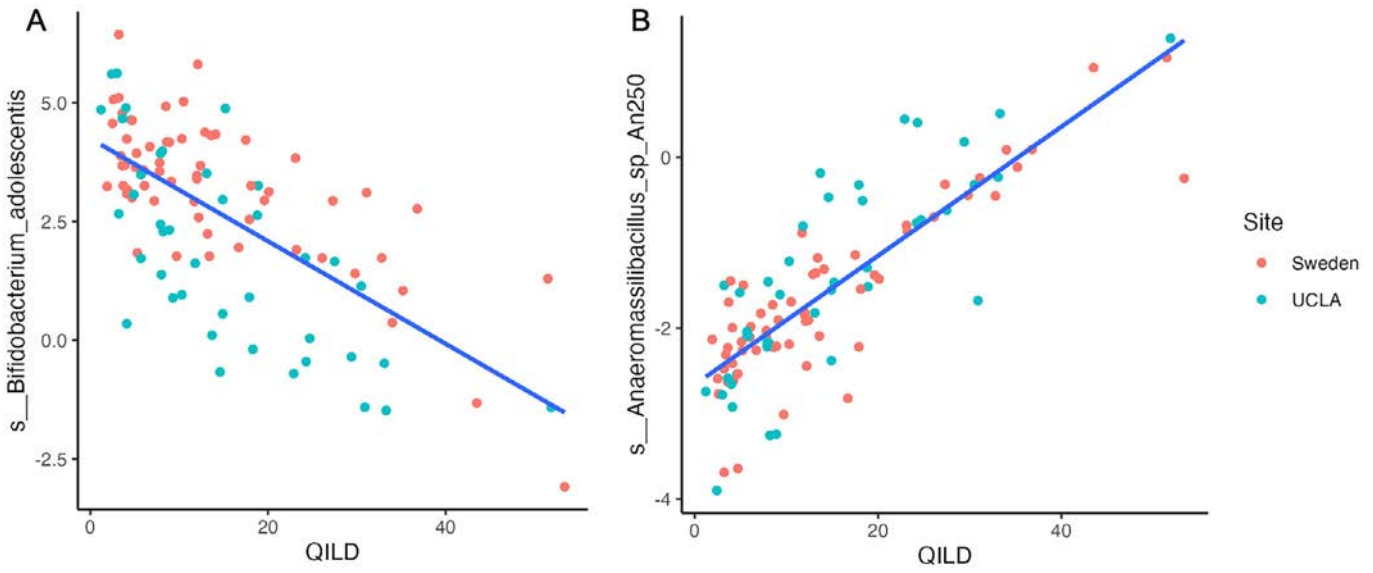

Figure 5. Relationship between *Bifidobacterium adolescentis* abundance (A) and *Anaeromassilibacillus sp An250* (B) and QILD scores based on general linear model estimates (Red = Lund University; Blue = UCLA; n = 103). The models adjusted for body mass index, proton pump inhibitor use, probiotic use, immunomodulatory use, small intestinal bacterial overgrowth, and study site. The y-axis represents the center log-ratio transformation of relative abundances. Positive fold change scores in this ratio indicate increased relative abundance of a species and negative fold change scores indicated decreased relative abundance of a species. QILD, quantitative image analysis to calculate the radiologic extent of interstitial lung disease; UCLA, University of California, Los Angeles. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25623/abstract.

demonstrated that increased self-reported reflux is associated with radiologic progression of ILD in patients with SSc-ILD.[40]

The results of the functional analyses provide initial insight into the metabolic pathways that may contribute to SSc-ILD pathogenesis. For instance, increased glycogen I degradation was associated with increased ILD scores. Although no prior studies have examined fecal metabolic pathways associated with SSc-ILD, alteration of the glycolysis pathway was also observed in patients with idiopathic pulmonary fibrosis compared with controls in a study, which performed gas chromatography–mass spectrometry-based metabolic profiling on lung tissue.[41] Increased urea cycle activity was also positively associated with ILD scores in the present study. Prior studies have demonstrated an association between the production of urea and GI microbiota.[42] One study found that overabundant intestinal urea alters macrophage function and may contribute to colorectal carcinogenesis.[43] The pathway analysis results from the present study are hypothesis generating and may help guide future metabolomics research in SSc-ILD.

This study has some important limitations. First, we did not have an equal number of participants from all countries, increasing the risk of selection bias. Reassuringly, the 257 bacterial species under investigation were found in 98% of patient samples from each site. Second, the cross-sectional design limits our ability to understand the direction of the association between GI microbiota and ILD. Based on other studies investigating the relationship between the gut microbiome and other organ systems (eg, brain-gut axis[44]), we suspect that a bidirectional relationship exists, but this requires confirmation with future longitudinal studies. Third, we were unable to identify a dietary history questionnaire validated across all of the international sites, and therefore, we did not collect information on diet in all participants. However, by including patients from diverse regions throughout the world, we aimed to discover a common species signature of ILD and ILD severity regardless of environmental differences, including diet. A fourth limitation is that not all patients in this study had an HRCT that was amenable to undergoing QIA, and the imaging was performed within four months of the stool collection. Although one would not expect significant changes on HRCT within a four-month period in a patient with SSc-ILD, it is possible that changes in fecal microbial composition could have occurred between the HRCT date and the stool collection date. Finally, the magnitude of the reported effect sizes for individual species were relatively small, indicating the need for larger studies in the future to avoid Type 2 errors. However, even though small to moderate effect sizes were observed for individual species, multivariate analyses revealed a signature capturing interactions between species, which suggested a larger effect size difference (eg, Cohen's d = 0.74). This signature score has biologic validity given that microbial species do not act in isolation and instead interact though various mechanisms to exert their effects on the host.

Strengths of this study include the fact that we studied an international cohort of patients with SSc from different geographic regions with varying environmental exposures. Another strength was our use of WGS to perform taxonomic classification, which is an advancement over the prior SSc studies in this area,

which could not identify bacterial species. In addition, we used QIA to objectively measure the radiologic extent of ILD. Finally, this is the first study to adjust for several known confounders associated with GI microbial composition. In taking this important step to minimize the influence of these confounders, we hope to inspire other researchers in this field to consider adopting a similar approach to maximize the likelihood of discovering true microbiota-disease state associations in SSc.

In summary, the present study demonstrated that patients with SSc-ILD from around the world have a unique GI microbial signature compared with patients with SSc without ILD. Notably, specific bacterial species were associated with the radiologic extent of ILD, suggesting that these bacteria and/or their metabolites may participate in different immunologic pathways that participate in SSc-ILD pathogenesis. Although these findings require replication in other cohorts, the results reported herein provide compelling evidence to support further research into the gut-lung axis in SSc-ILD.

ACKNOWLEDGMENTS

We thank the patients who participated in this study, along with the study coordinators and staff who helped to facilitate the collection and international shipping of stool samples during the COVID-19 pandemic. We would also like to acknowledge the UCLA Goodman-Luskin Microbiome Center Microbiome Core and Integrative Biostatistics and Bioinformatics Core.

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 Volkmann 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

Boehringer Ingelheim Pharmaceuticals, Inc (BIPI) had no role in the study design or in the collection, analysis, or interpretation of the data, the writing of the manuscript, or the decision to submit the manuscript for publication. BIPI was given the opportunity to review the publication for medical and scientific accuracy as it relates to BIPI substances, as well as intellectual property considerations. Publication of this article was not contingent upon approval by BIPI.

REFERENCES

1. Volkmann ER, Andréasson K, Smith V. Systemic sclerosis. Lancet 2023;401(10373):304–318.
2. Conrad N, Misra S, Verbakel JY, et al. Incidence, prevalence, and co-occurrence of autoimmune disorders over time and by age, sex, and socioeconomic status: a population-based cohort study of 22 million individuals in the UK. Lancet 2023;401(10391):1878–1890.
3. Ng SC, Shi HY, Hamidi N, et al. Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies. Lancet 2017;390(10114):2769–2778.
4. Al Bander Z, Nitert MD, Mousa A, et al. The gut microbiota and inflammation: an overview. Int J Environ Res Public Health 2020;17(20):7618.
5. Budden KF, Gellatly SL, Wood DL, et al. Emerging pathogenic links between microbiota and the gut-lung axis. Nat Rev Microbiol 2017;15(1):55–63.
6. Frati F, Salvatori C, Incorvaia C, et al. The role of the microbiome in asthma: the gut-lung axis. Int J Mol Sci 2018;20(1):123.
7. Qu L, Cheng Q, Wang Y, et al. COPD and gut-lung axis: how microbiota and host inflammasome influence COPD and related therapeutics. Front Microbiol 2022;13:868086.
8. Volkmann ER. Is there a role for the microbiome in systemic sclerosis? Expert Rev Clin Immunol 2023;19(3):237–240.
9. Bellocchi C, Volkmann ER. Update on the gastrointestinal microbiome in systemic sclerosis. Curr Rheumatol Rep 2018;20(8):49.
10. Andréasson K, Lee SM, Lagishetty V, et al. Disease features and gastrointestinal microbial composition in patients with systemic sclerosis from two independent cohorts. ACR Open Rheumatol 2022;4(5):417–425.
11. Ranjan R, Rani A, Metwally A, et al. Analysis of the microbiome: advantages of whole genome shotgun versus 16S amplicon sequencing. Biochem Biophys Res Commun 2016;469(4):967–977.
12. van den Hoogen F, Khanna D, Fransen J, et al. 2013 classification criteria for systemic sclerosis: an American College of Rheumatology/-European League Against Rheumatism collaborative initiative. Arthritis Rheum 2013;65(11):2737–2747.
13. Kim HG, Tashkin DP, Clements PJ, et al. A computer-aided diagnosis system for quantitative scoring of extent of lung fibrosis in scleroderma patients. Clin Exp Rheumatol 2010;28(5 suppl 62):S26–S35.
14. Bankier AA, MacMahon H, Colby T, et al. Fleischner Society: glossary of terms for thoracic imaging. Radiology 2024;310(2):e232558.
15. Volkmann ER, Chang YL, Barroso N, et al. Association of systemic sclerosis with a unique colonic microbial consortium. Arthritis Rheumatol 2016;68(6):1483–1492.
16. Blanco-Míguez A, Beghini F, Cumbo F, et al. Extending and improving metagenomic taxonomic profiling with uncharacterized species using MetaPhlAn 4. Nat Biotechnol 2023;41(11):1633–1644.
17. Beghini F, McIver LJ, Blanco-Míguez A, et al. Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. Elife 2021;10:e65088.
18. Franzosa EA, McIver LJ, Rahnavard G, et al. Species-level functional profiling of metagenomes and metatranscriptomes. Nat Methods 2018;15(11):962–968.
19. Puljiz Z, Kumric M, Vrdoljak J, et al. Obesity, gut microbiota, and metabolome: from pathophysiology to nutritional interventions. Nutrients 2023;15(10):2236.
20. Bruno G, Zaccari P, Rocco G, et al. Proton pump inhibitors and dysbiosis: current knowledge and aspects to be clarified. World J Gastroenterol 2019;25(22):2706–2719.
21. Hemarajata P, Versalovic J. Effects of probiotics on gut microbiota: mechanisms of intestinal immunomodulation and neuromodulation. Therap Adv Gastroenterol 2013;6(1):39–51.
22. Gabarre P, Loens C, Tamzali Y, et al. Immunosuppressive therapy after solid organ transplantation and the gut microbiota: bidirectional interactions with clinical consequences. Am J Transplant 2022;22(4):1014–1030.
23. Banaszak M, Górna I, Woźniak D, et al. Association between gut dysbiosis and the occurrence of SIBO, LIBO, SIFO and IMO. Microorganisms 2023;11(3):573.

24. Volkmann ER. Natural history of systemic sclerosis-related interstitial lung disease: how to identify a progressive fibrosing phenotype. J Scleroderma Relat Disord 2020;5(2 suppl):31–40.
25. Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res 2000;28(1):27–30.
26. Chen J, Wright K, Davis JM, et al. An expansion of rare lineage intestinal microbes characterizes rheumatoid arthritis. Genome Med 2016;8(1):43.
27. Gilliland A, Chan JJ, De Wolfe TJ, et al. Pathobionts in inflammatory bowel disease: Origins, underlying mechanisms and implications for clinical care. Gastroenterology 2024;166(1):44–58.
28. Qian X, Si Q, Lin G, et al. *Bifidobacterium adolescentis* is effective in relieving type 2 diabetes and may be related to its dominant core genome and gut microbiota modulation capacity. Nutrients 2022; 14(12):2479.
29. Volkmann ER, Fischer A. Update on morbidity and mortality in systemic sclerosis-related interstitial lung disease. J Scleroderma Relat Disord 2021;6(1):11–20.
30. Raghu G, Montesi SB, Silver RM, et al. Treatment of systemic sclerosis-associated interstitial lung disease: evidence-based recommendations. an official American Thoracic Society clinical practice guideline. Am J Respir Crit Care Med 2024;209(2):137–152.
31. Chioma OS, Hesse LE, Chapman A, et al. Role of the microbiome in interstitial lung diseases. Front Med (Lausanne) 2021;8:595522.
32. Muhamad Rizal NS, Neoh HM, Ramli R, et al. Advantages and limitations of 16S rRNA next-generation sequencing for pathogen identification in the diagnostic microbiology laboratory: perspectives from a middle-income country. Diagnostics (Basel) 2020;10(10):816.
33. Andréasson K, Alrawi Z, Persson A, et al. Intestinal dysbiosis is common in systemic sclerosis and associated with gastrointestinal and extraintestinal features of disease. Arthritis Res Ther 2016;18(1):278.
34. Caimmi C, Bertoldo E, Venturini A, et al. Relationship between increased fecal calprotectin levels and interstitial lung disease in systemic sclerosis. J Rheumatol 2019;46(3):274–278.
35. Marie I, Leroi AM, Menard JF, et al. Fecal calprotectin in systemic sclerosis and review of the literature. Autoimmun Rev 2015;14(6): 547–554.
36. Polkowska-Pruszyńska B, Gerkowicz A, Rawicz-Pruszyński K, et al. The role of fecal calprotectin in patients with systemic sclerosis and small intestinal bacterial overgrowth (SIBO). Diagnostics (Basel) 2020;10(8):587.
37. Guo Y, Xie JP, Deng K, et al. Prophylactic effects of *Bifidobacterium adolescentis* on anxiety and depression-like phenotypes after chronic stress: a role of the gut microbiota-inflammation axis. Front Behav Neurosci 2019;13:126.
38. Volkmann ER, Tashkin DP, Leng M, et al. Biological correlates of radiological features of systemic sclerosis interstitial lung disease. ERJ Open Res 2025;11(1):00596–2024.
39. De Luca D, Alonso A, Autilio C. Bile acid-induced lung injury: update of reverse translational biology. Am J Physiol Lung Cell Mol Physiol 2022;323(1):L93–L106.
40. Volkmann ER, Tashkin DP, Leng M, et al. Association of symptoms of gastroesophageal reflux, esophageal dilation, and progression of systemic sclerosis-related interstitial lung disease. Arthritis Care Res (Hoboken) 2023;75(8):1690–1697.
41. Kang YP, Lee SB, Lee JM, et al. Metabolic profiling regarding pathogenesis of idiopathic pulmonary fibrosis. J Proteome Res 2016; 15(5):1717–1724.
42. Zarei I, Koistinen VM, Kokla M, et al. Tissue-wide metabolomics reveals wide impact of gut microbiota on mice metabolite composition. Sci Rep 2022;12(1):15018.
43. Chen H, Tong T, Lu SY, et al. Urea cycle activation triggered by host-microbiota maladaptation driving colorectal tumorigenesis. Cell Metab 2023;35(4):651–666.e7.
44. Zhao L, Xiong Q, Stary CM, et al. Bidirectional gut-brain-microbiota axis as a potential link between inflammatory bowel disease and ischemic stroke. J Neuroinflammation 2018;15(1):339.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 78, No. 3, March 2026, pp 362–370
DOI 10.1002/acr.25625

Identification of Adult Patients With Dermatomyositis Using Real-World Data Sources

Benjamin Martin,[1] Will Kelly,[1] Hannah Morgan-Cooper,[2] Thomas Falconer,[3] Elizabeth Park,[2] Priya Desai,[2] David Fiorentino,[2] Lorinda Chung,[2] Sean Yen,[1] Zachary Wang,[1] Didem Saygin,[4] Michael George,[5] Gowtham A. Rao,[6] Joel Swerdel,[7] Azza Shoaibi,[7] and Christopher A. Mecoli[1]

Objective. Studying rare diseases like dermatomyositis (DM) in single-center cohorts is challenging due to small sample sizes and limited generalizability. This study develops and evaluates case identification algorithms for DM to enable coordinated analysis across multiple data sources.

Methods. Case identification algorithms were developed to identify adult patients with DM within 11 independent electronic health record or claims databases, totaling over 800 million patients, using the Observational Medical Outcomes Partnership Common Data Model. Algorithm performance was assessed through manual chart review and using Observational Health Data Sciences and Informatics open-source tools (CohortDiagnostics and PheValuator), which quantify incidence rates and performance metrics such as sensitivity and positive predictive value (PPV).

Results. Eight DM case identification algorithms were evaluated across 11 databases, revealing significant variability in performance, with sensitivity and PPV differing by more than 30% between some databases. Overall, we identified one incidence algorithm and one prevalence algorithm with good performance, demonstrated by sensitivity rates of 42% and 49% and PPV values of 83% and 84%, respectively. PheValuator quantified algorithm performance within each database, allowing for direct comparison of different criteria. Additionally, CohortDiagnostics generated incidence rates stratified by age decile and sex, aligning with previous epidemiologic data.

Conclusion. We developed and validated multiple DM case identification algorithms across diverse databases, demonstrating their accuracy through multiple evaluation methods. This approach enables more generalizable, reproducible research using real-world data and can be applied to other rheumatic diseases.

INTRODUCTION

Dermatomyositis (DM) is a rare chronic autoimmune disease characterized by inflammation and weakness of the muscles and skin rash. It can affect multiple organ systems, leading to substantial morbidity and mortality.[1] The incidence of DM varies, with estimates ranging from 2 to 20 cases per 1,000,000 person-years.[1–2] Due to its rarity, clinical research studies on DM often have inadequate power for comparative effectiveness or long-term outcome studies, particularly when conducted within single academic centers. One potential solution is to leverage real-world data sources such as electronic health record (EHR) data, insurance claims data, and national health care system registries to study DM. These data sources provide longitudinal

This manuscript is the result of funding in whole or in part by the NIH. It is subject to the NIH Public Access Policy. Through acceptance of this federal funding, the NIH has been given a right to make this manuscript publicly available in PubMed Central upon the official date of publication, as defined by the NIH.

Supported in part by the Jerome L. Greene Foundation and the NIH (grants 5U24-HD-113136 and 5T15-LM-013979 to Dr Martin). Dr Saygin's work was supported by the Rheumatology Research Foundation Scientist Development Award, Rukel Family Foundation, Chicago Center on Musculoskeletal Pain Pilot Grant, and RUSH to Progress Pilot Award.

[1]Benjamin Martin, PhD, Will Kelly, BS, Sean Yen, MD, MS, Zachary Wang, MS, Christopher A. Mecoli, MD, MHS: Johns Hopkins University School of Medicine, Baltimore, Maryland; [2]Hannah Morgan-Cooper, MS, Elizabeth Park, MD, Priya Desai MS, David Fiorentino, MD, PhD, Lorinda Chung, MD, MS: Stanford University School of Medicine, Stanford, and Stanford Health Care, Redwood City, California; [3]Thomas Falconer, MS: Columbia University Irving Medical Center/New York Presbyterian Hospital, New York; [4]Didem Saygin, MD: Rush University Medical Center, Chicago, Illinois; [5]Michael George, MD, MSCE: University of Pennsylvania Perelman School of Medicine, Philadelphia; [6]Gowtham A. Rao, MD: Observational Health Data Sciences and Informatics (OHDSI) Collaborators, OHDSI, New York, New York; [7]Joel Swerdel, PhD, MS, MPH, Azza Shoaibi, PhD: OHDSI Collaborators, OHDSI, New York, New York, and Global Epidemiology, Office of the Chief Medical Officer, Johnson & Johnson, New Brunswick, New Jersey.

Drs Martin and Kelly are co-first authors and contributed equally to this work.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25625).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25625.

Address correspondence via email to Christopher A. Mecoli, MD, MHS, at cmecoli1@jhmi.edu.

Submitted for publication April 3, 2025; accepted in revised form July 22, 2025.

SIGNIFICANCE AND INNOVATIONS

- We developed and tested eight reusable DM case-identification algorithms across 11 OMOP-mapped EHR/claims databases (⟩800M patients), enabling multi-site, generalizable real-world studies.
- Identified one incident (#2) and one prevalent (#4) algorithm with strong precision (PPV ∼83–84%) and moderate sensitivity (∼42–49%), giving practical options depending on study goals.
- Validated performance using two complementary methods—manual chart review (ACR/EULAR 2017 as reference) and PheValuator—showing substantial agreement (Cohen's κ = 0.70).
- Quantified substantial cross-database variability (⟩30% swings in sensitivity/PPV) and surfaced mapping gaps (e.g., specialty provider, autoantibody results), providing concrete guidance on phenotype portability and data-quality priorities.

datasets with large numbers of cases, offering great potential for studying the disease.

Historically, the study of DM leveraging real-world data has been restricted to a single data source, typically the one the researchers are affiliated with or have access to. This focus on a single data source diminishes the generalizability of the study results. Furthermore, different studies use different DM algorithms, limiting reproducibility and comparability of results. Perhaps most important, there exist large gaps on how currently used DM case identification algorithms perform across different data sources. To address these challenges, our study has developed several DM case identification algorithms–that is, reusable computerized search queries–and reports the performance characteristics for each across a range of different real-world data sources. The focus on creating a set of accurate and generalizable case identification algorithms for DM across a variety of databases using a common data model (CDM) is a critical preliminary step in the multistage process of generating clinical evidence from a distributed network of data.

To validate algorithms to identify DM across a range of datasets, we leveraged the Observational Health Data Sciences and Informatics (OHDSI) network. OHDSI uses the Observational Medical Outcomes Partnership (OMOP) CDM, a standardized data model that facilitates the harmonization of disparate health care data sources. By using the OMOP CDM, the OHDSI community can integrate and analyze data from various sources, such as EHRs, insurance claims, and registries, to generate real-world evidence. Currently, 12% of the world's population health data has been converted to the OMOP CDM, and it is used by multiple US academic medical centers, The European Medicines Agency Data Analysis and Real World Interrogation Network, the US National COVID Cohort Collaborative, the National Institutes of Health All Of Us Program, and many others. Furthermore, the OHDSI community has created multiple open-source tools to create and evaluate case identification algorithms and analytical tools to conduct multisite studies. Our objectives were (1) to develop several case identification algorithms for DM, (2) use novel methods to evaluate their performance, and (3) compare and contrast this performance across several different data sources.

MATERIALS AND METHODS

Creation of algorithms. We aimed to create algorithms to identify adult patients with a clinical diagnosis of DM using real-world data (EHRs, claims, and registries) converted to the OMOP CDM. We began by conducting a literature review on prior studies examining the performance characteristics of algorithms for DM[3–8] (Supplemental Table 1). Drawing from the literature, we created eight DM algorithms using OHDSI methodology through the open-source tool ATLAS, a web-browser interface that allows transparent creation of disease algorithms (https://atlas.ohdsi.org/).[9] We designed algorithms varying the inclusion criteria such as requirement for laboratory values (eg, myositis-specific autoantibodies), prescriptions for immunosuppressive or immunomodulatory medications, the number of diagnosis codes, or requirement to have received a DM diagnosis from a specialty provider to test and compare across a spectrum of algorithms.

Evaluation of algorithms. The evaluation of each algorithm was performed in several ways. First, for a subgroup of the patients at Johns Hopkins University, we conducted a gold standard manual chart review to confirm or exclude DM diagnosis based on validated 2017 American College of Rheumatology (ACR)/ European League Against Rheumatism (EULAR) Idiopathic Inflammatory Myositis Classification Criteria for DM.[10] This subgroup included patients with DM as well as polymyositis, inclusion body myositis, immune-mediated necrotizing myopathy, antisynthetase syndrome, and other genetic and connective tissue disorders and was not chosen at random but consisted of patients referred to our Johns Hopkins Myositis Center from 2004 to 2024 for suspected muscle disease.

Next, each DM case identification algorithm was applied to all 11 data sources, resulting in the generation of 11 cohorts of adult patients with DM. The accuracy of the DM case identification algorithms in each of these 11 DM cohorts was evaluated by two open-source programs: CohortDiagnostics[11] and PheValuator.[12–13] CohortDiagnostics generates descriptive summaries of the conditions, medications, procedures, and other clinical data for patients included by each algorithm. These summaries allow researchers to assess whether the characteristics of the identified cohorts are consistent with what is known about the target condition. Such data, in aggregate and paired with existing knowledge and clinical experience, can

provide insight into how well each case identification algorithm performs.

Finally, we used the open-source program PheValuator to evaluate each case identification algorithm. The PheValuator tool generates a probabilistic estimate of disease likelihood (adult DM) as a surrogate for manual chart review. Unlike case-finding methods that depend only on a limited list of diagnosis codes and inclusion rules, the PheValuator model uses all available data (eg, diagnoses, medications, procedures, and laboratory results) structured in the OMOP CDM to arrive at a probabilistic estimate. This probabilistic estimate is derived from a Least Absolute Shrinkage and Selection Operator regression model, trained with cross-validation on a highly specific (xSpec) cohort, and produces probability scores for patients in a large test set. These scores are then compared with the binary outputs of the case identification algorithms to calculate performance metrics, including sensitivity and positive predictive value (PPV). The process involves the following steps:

1. Identifying a cohort of patients with a very high probability of having the disease of interest, in this case, adult DM. This is referred to as the extra specific, or "xSpec" cohort.
2. A cohort of patients with a very high sensitivity for identifying patients with the disease of interest. This is called the extra sensitive, or "xSens" cohort. Its primary purpose is to identify reliable negatives; the inverse of this cohort (ie, individuals not in xSens) are treated as reliable negatives in model training. These patients are highly likely not to have adult DM.
3. Identifying a cohort of patients to estimate the prevalence of the disease of interest, called the "prevalence" cohort. The estimated prevalence from this cohort informs how many positive (xSpec) and negative (non-xSens) subjects should be sampled when building the diagnostic model.

With these three cohorts identified, PheValuator can develop a predictive model for DM and generate a probabilistic gold standard in any data source used. Each DM algorithm can then be tested against this probabilistic gold standard derived from PheValuator. This approach allows researchers to understand how the algorithms perform in any data source without needing access to individual patient charts. Previous studies have demonstrated that results from PheValuator show reliable agreement with manual chart review.[12–13] The cohort definitions and inclusion criteria for the xSpec, xSens, and prevalence cohorts used to perform the PheValuator analysis on each of the 11 databases included in this study were the same. These cohort definitions are provided in Supplemental Material 1 and the concept sets used for each criterion are detailed in Supplemental Material 2.

To assess the validity of this approach, we compared manual chart review results from the Johns Hopkins University data to PheValuator's performance on the same patient subgroup within the Johns Hopkins OMOP CDM. We evaluated sensitivity, PPV, and overall agreement, using Cohen's kappa statistic to measure concordance between PheValuator and the manual chart review (considered the gold standard).

Real-world data sources. Each DM algorithm was tested in 11 databases. The databases used include Johns Hopkins University OMOP CDM, Stanford University Research Repository (STARR), Columbia University OMOP CDM, the Japan Medical Data Center (JMDC), Merative Medicare (MDCR), Merative Medicaid (MDCD), Merative Commercial Claims and Encounters, Optum's longitudinal EHR repository (Optum EHR), Optum's Clinformatics Data Mart (Optum DOD), IQVIA Pharmetrics Adjudicated Health Plan Claims Data, and HealthVerity. The characteristics of each database are described in Supplemental Table 2. These databases were chosen due to their availability through the OHDSI community, as well as a desire to have both EHR and claims data available for analysis. The Johns Hopkins University, Stanford University, and Columbia University databases comprise data from each institution's entire EHR mapped to OMOP. Altogether, these databases comprise three US tertiary referral academic centers, four EHR data sources (three academic centers plus Optum EHR), claims data of Medicaid, Medicare, and privately insured individuals in the United States, as well as claims data from Japan. The presented analysis was reviewed separately by each organization's institutional review board and/or data governance committee and individually granted approval by each entity for execution on the above listed databases under their purview.

RESULTS

Creation of case identification algorithms. Eight case identification algorithms were developed using the ATLAS open-source cohort discovery tool with inclusion criteria listed in Table 1. Both the computable and human readable case identification algorithms have been shared in the Supplemental Material as well as the study's GitHub repository (https://github.com/ohdsi-studies/MyositisNetworkStudy). Two case identification algorithms returned no patients for all databases: algorithm 8, which included physician specialty criteria, and algorithm 7, which required a myositis-specific autoantibody test with a positive result. This indicates that these data, physician specialty and laboratory results, were either not available in the data sources used (such as the claims data sources) or were not mapped to the particular OMOP databases queried in this study. Algorithm 6, which only used an idiopathic inflammatory myopathy (IIM) autoantibody test order as a criterion and did not require a positive result, returned results for three out of four EHR databases and two EHR-enriched claims databases.

Table 1. Calculated sensitivity and PPV using PheValuator for each case identification algorithm across all databases*

Case identification algorithm	Criteria	Database	Cohort count, n	Sensitivity (95% CI)	PPV (95% CI)
1	Age >18 years at index date	Johns Hopkins University	740	0.525 (0.462–0.588)	0.951 (0.901–0.980)
	Second diagnostic code of DM within 30–365 days of index date	STARR	603	0.624 (0.574–0.673)	0.886 (0.842–0.921)
		Columbia University	444	NR	NR
	Incident: 365-day minimum observation period before index	Optum EHR	7,442	0.561 (0.479–0.640)	0.599 (0.515–0.679)
		Optum DOD	8,006	NR	NR
		HealthVerity	18,575	0.614 (0.529–0.693)	0.574 (0.492–0.653)
		IQVIA Pharmetrics	9,849	0.571 (0.467–0.671)	0.709 (0.596–0.806)
		JMDC	4,494	0.662 (0.616–0.706)	0.818 (0.774–0.856)
		Truven CCAE	9,421	0.542 (0.448–0.634)	0.780 (0.675–0.864)
		Truven MDCD	1,534	0.488 (0.398–0.579)	0.635 (0.531–0.731)
		Truven MDCR	1,897	0.375 (0.321–0.432)	0.768 (0.693–0.833)
2	Age >18 years at index date	Johns Hopkins University	485	0.389 (0.329–0.452)	0.962 (0.904–0.989)
	Second diagnostic code of DM within 30–365 days of index date	STARR	434	0.484 (0.434–0.536)	0.917 (0.870–0.951)
		Columbia University	248	NR	NR
	At least one immunosuppressive medication prescribed	Optum EHR	4,999	0.503 (0.422–0.584)	0.725 (0.631–0.806)
		Optum DOD	5,729	NR	NR
	Incident: 365-day minimum observation period before index	HealthVerity	13,567	0.534 (0.450–0.617)	0.703 (0.609–0.786)
		IQVIA Pharmetrics	5,853	0.378 (0.282–0.481)	0.881 (0.744–0.960)
		JMDC	2,070	0.340 (0.296–0.386)	0.944 (0.897–0.974)
		Truven CCAE	7,098	0.458 (0.366–0.552)	0.857 (0.746–0.933)
		Truven MDCD	975	0.336 (0.254–0.426)	0.724 (0.591–0.833)
		Truven MDCR	1,398	0.307 (0.256–0.362)	0.812 (0.729–0.878)
3	Age >18 years at index date	Johns Hopkins University	697	0.638 (0.576–0.697)	0.911 (0.860–0.948)
	At least one immunosuppressive medication prescribed	STARR	621	0.687 (0.638–0.733)	0.858 (0.814–0.895)
		Columbia University	427	NR	NR
	Incident: 365-day minimum observation period before index	Optum EHR	8,920	0.777 (0.704–0.840)	0.646 (0.573–0.714)
		Optum DOD	9,435	NR	NR
		HealthVerity	21,109	0.651 (0.567–0.728)	0.534 (0.458–0.609)
		IQVIA Pharmetrics	8,290	0.643 (0.540–0.737)	0.750 (0.644–0.838)
		JMDC	2,443	0.423 (0.377–0.470)	0.922 (0.876–0.955)
		Truven CCAE	11,457	0.695 (0.603–0.776)	0.752 (0.660–0.830)
		Truven MDCD	1,666	0.544 (0.453–0.633)	0.567 (0.473–0.657)
		Truven MDCR	2,363	0.612 (0.555–0.666)	0.756 (0.698–0.808)
4	Age >18 years at index date	Johns Hopkins University	637	0.440 (0.412–0.467)	0.982 (0.968–0.992)
	Second diagnostic code of DM within 30–365 days of index date	STARR	476	0.502 (0.459–0.544)	0.927 (0.891–0.954)
		Columbia University	301	NR	NR
	At least one immunosuppressive medication prescribed	Optum EHR	5,402	0.536 (0.466–0.606)	0.750 (0.672–0.817)
		Optum DOD	7,320	NR	NR
	Prevalent: 0 days minimum observation period before index	HealthVerity	16,274	0.586 (0.524–0.645)	0.701 (0.636–0.760)
		IQVIA Pharmetrics	7,874	0.489 (0.429–0.550)	0.824 (0.757–0.879)
		JMDC	2,308	0.363 (0.323–0.403)	0.950 (0.912–0.975)
		Truven CCAE	9,275	0.544 (0.482–0.606)	0.850 (0.787–0.901)
		Truven MDCD	1,175	0.411 (0.348–0.475)	0.783 (0.702–0.851)
		Truven MDCR	1,779	0.398 (0.360–0.437)	0.837 (0.790–0.876)

(Continued)

Table 1. (*Cont'd*)

Case identification algorithm	Criteria	Database	Cohort count, n	Sensitivity (95% CI)	PPV (95% CI)
5	Age >18 years at index date Second diagnostic code of DM within 30–365 days of index date Prevalent: 0 days minimum observation period before index	Johns Hopkins University	1,043	0.712 (0.686–0.737)	0.971 (0.958–0.981)
		STARR	656	0.628 (0.586–0.668)	0.899 (0.865–0.928)
		Columbia University	519	NR	NR
		Optum EHR	8,143	0.594 (0.524–0.662)	0.609 (0.538–0.677)
		Optum DOD	10,649	NR	NR
		HealthVerity	23,036	0.664 (0.604–0.720)	0.571 (0.514–0.626)
		IQVIA Pharmetrics	13,519	0.716 (0.659–0.768)	0.703 (0.646–0.756)
		JMDC	5,261	0.669 (0.629–0.708)	0.829 (0.791–0.862)
		Truven CCAE	12,910	0.636 (0.574–0.694)	0.769 (0.706–0.823)
		Truven MDCD	1,976	0.573 (0.509–0.636)	0.681 (0.613–0.744)
		Truven MDCR	2,473	0.515 (0.475–0.554)	0.783 (0.740–0.821)
6	Age >18 years at index date Myositis-specific autoantibody test ordered Prevalent: 0 days minimum observation period before index	Johns Hopkins University	308	0.193 (0.171–0.215)	0.968 (0.938–0.986)
		STARR	138	0.119 (0.093–0.149)	0.835 (0.735–0.909)
		Columbia University	138	NR	NR
		Truven CCAE	383	0.023 (0.008–0.049)	0.600 (0.262–0.878)
		Truven MDCR	43	0.005 (0.001–0.014)	0.750 (0.194–0.994)
7	Age >18 years at index date Myositis-specific autoantibody test positive Prevalent: 0 days minimum observation period before index	All databases	0		
8	Age >18 years at index date Second diagnostic code of DM within 30–365 days of index date Seen by rheumatologist, neurologist, or dermatologist Prevalent: 0 days minimum observation period before index	All databases	0		

* CCAE, Commercial Claims and Encounters; CI, confidence interval; DM, dermatomyositis; JMDC, Japan Medical Data Center; MDCD, Merative Medicaid; MDCR, Merative Medicare; NR, not reported due to the extra specific cohort being of inadequate size; Optum DOD, Optum's Clinformatics Data Mart; Optum EHR, Optum's longitudinal electronic health record repository; PPV, positive predictive value; STARR, Stanford University Research Repository.

Table 2. Comparison between manual chart review and probabilistic gold standard (PheValuator) using data at Johns Hopkins University*

	Sensitivity			PPV		
Case identification algorithm	Reference: manual chart review for ACR/EULAR 2017 classification criteria	Reference: probabilistic gold standard (PheValuator)	Delta	Reference: manual chart review for ACR/EULAR 2017 classification criteria	Reference: probabilistic gold standard (PheValuator)	Delta
1	0.55	0.53	0.02	0.93	0.86	0.07
2	0.35	0.39	−0.04	0.91	0.91	0.01
3	0.41	0.64	−0.22	0.88	0.74	0.14
4	0.40	0.44	−0.04	0.90	0.95	−0.06
5	0.68	0.71	−0.03	0.92	0.93	−0.004
6	0.24	0.19	0.04	0.83	0.91	−0.08

* ACR, American College of Rheumatology; PPV, positive predictive value.

Evaluation of case identification algorithms using gold standard manual chart review. A subgroup of patients (n = 2,590) within the Johns Hopkins Hospital OMOP CDM underwent manual chart review to determine whether the patient met ACR/EULAR 2017 classification criteria for DM.[10] Table 2 shows the sensitivity and PPV for each algorithm using manual chart review (where the reference was ACR/EULAR Classification Criteria for DM[10]) compared with using PheValuator (where the reference was a probabilistic gold standard for DM). Cohen's kappa statistic was calculated to evaluate agreement between the cases identified by manual chart review and PheValuator, yielding $\kappa = 0.70$, which signifies substantial agreement between the two reference standards applied to the Johns Hopkins University patient sample.

Evaluation of case identification algorithms using CohortDiagnostics. Using algorithm 2, which was created to identify incident DM and be of higher specificity, incidence rates of adult DM in claims databases ranged from 6 to 30 patients/million person-years, consistent with prior epidemiologic reports ranging from 2 to 20 patients/million person-years (Table 3).[12] Expectedly, the incidence rates across academic centers (4–67/million patient-years) were on average higher, given these are tertiary care centers enriched for rare diseases such as DM. Consistent with prior epidemiologic literature, the incidence rate was 2- to 3-fold higher for women compared with men across all data sources and algorithms. In addition, the highest incidence rates were in the age deciles 50 to 59 years and 60 to 69 years, consistent with prior literature (Supplemental Table 3). For the more specific algorithm 2 and more sensitive algorithm 1, the frequency of DM-relevant measurements and procedures are summarized in Table 4. Creatinine kinase (CK), aldolase, and skin and muscle biopsies were commonly observed for the DM cohorts across all databases, except for two databases in which no skin and muscle biopsies were recorded (STARR and JMDC). CK measurements were taken for 86% and 92%, and aldolase measurements were recorded for 61% and 71% of patients with DM identified by algorithm 1 and algorithm 2, respectively. The variability in frequency of these DM-relevant measurements was relatively small across databases and consistently less among the more specific algorithm 2 compared to algorithm 1, with an interquartile range (IQR) of 6.5% and 3.0% for CK and IQR of 13.5% and 7.0% for aldolase, respectively.

Evaluation of case identification algorithms using probabilistic gold standard (PheValuator). We analyzed the performance of each algorithm across 9 of the 11 databases using PheValuator (Table 1). The model output for identifying DM

Table 3. Cohort counts, person-years, and incidence rates across all data sources*

Database source	Cohort count, n	Person-years, n	Incidence rate (per 1,000,000 PY)
Columbia University	248	57,068,630	4.35
MDCD	975	144,532,193	6.75
Optum EHR	4,999	570,549,689	8.76
CCAE	7,098	492,792,920	14.4
Pharmetrics	5,853	401,283,864	14.59
HealthVerity	13,567	881,978,718	15.38
Optum DOD	5,729	323,501,218	17.71
Stanford University	434	24,398,358	17.79
JMDC	2,070	85,029,987	24.34
MDCR	1,398	45,377,736	30.81
Johns Hopkins University	485	7,231,328	67.07

* Data shown for case identification algorithm 2.
CCAE, Commercial Claims and Encounters; JMDC, Japan Medical Data Center; MDCD, Merative Medicaid; MDCR, Merative Medicare; Optum DOD, Optum's Clinformatics Data Mart; Optum EHR, Optum's longitudinal electronic health record repository; PY, person-years.

Table 4. Measurements and procedures on patient cohorts defined by case identification algorithms 1 and 2*

Database	CK measurement (ever)	CK measurement 31–365 days before and 31–365 days after index date	Aldolase measurement (ever)	Aldolase measurement 31–365 days before and 31–365 days after index date	Muscle biopsy, deep and superficial (ever)	Skin biopsy (ever)
OMOP concept(s)	4,265,595 or 3,007,220	4,265,595 or 3,007,220	4,196,218 or 3,027,953	4,196,218 or 3,027,953	4,208,408 + 4,010,111	2,101,931
Phenotype algorithm 1, %						
Merative MarketScan CCAE	90	56/69	66	30/37	17	35
HealthVerity	91	49/66	68	28/37	13	24
JDMC	97	71/85	44	18/25	NR	NR
Merative MarketScan Multi-State Medicaid Database	83	48/57	55	24/30	14	17
Merative MarketScan Medicare Supplemental	62	31/36	29	11/14	16	40
Optum DOD	90	56/67	64	35/39	15	29
Optum EHR	80	31/52	52	15/25	8	13
PharMetrics Plus	87	57/66	63	32/36	10	18
Johns Hopkins OMOP CDM	91	50/49	84	36/35	6	4
Stanford University (STARR)	85	49/55	85	30/60	NR	16
Columbia University (CUIMC)	85	31/56	57	19/31	9	13
Average overall	86	44/60	61	25/34	12	21
Phenotype algorithm 2, %						
Merative MarketScan CCAE	93	60/75	70	33/41	19	36
HealthVerity	95	55/73	73	32/42	15	25
JDMC	98	75/93	55	23/31	NR	NR
Merative MarketScan Multi-State Medicaid Database	91	60/58	67	33/39	16	18
Merative MarketScan Medicare Supplemental	91	60/68	67	33/39	16	18
Optum DOD	94	61/74	69	31/39	13	30
Optum EHR	86	36/61	60	18/31	11	15
PharMetrics Plus	92	63/72	69	37/40	13	19
Johns Hopkins OMOP CDM	94	55/55	88	41/39	6	5
Stanford University (STARR)	90	32/68	91	32/68	NR	20
Columbia University (CUIMC)	91	36/67	75	25/44	11	17
Average overall	92	54/69	71	31/41	13	20

* Values are the number. OMOP concept identifiers used for each column are included.
CCAE, Commercial Claims and Encounters; CDM, Common Data Model; CUIMC, Columbia University Irving Medical Center; CK, creatinine kinase; JMDC, Japan Medical Data Center; NR, not reported; OMOP, Observational Medical Outcomes Partnership; Optum DOD, Optum's Clinformatics Data Mart; Optum EHR, Optum's longitudinal electronic health record repository; STARR, Stanford University Research Repository.

cases through PheValuator (ie, probabilistic gold standard) can be found in Supplemental Table 4, reporting the top 10 weighted covariates used by the model. Columbia University and Optum DOD did not have sufficient patient counts for the xSpec cohort to be used for training a reliable predictive model and were thus excluded from the PheValuator analysis.

We found that the more criteria added to an algorithm, the lower the sensitivity and higher the PPV for identifying a patient with DM. For example, adding the criterion for prescription of an immunosuppressive medication reduced the sensitivity by approximately 15% but increased the PPV by 10% on average over all nine databases. The inclusion of some criteria–such as having a myositis autoantibody panel ordered or receiving a DM diagnosis from a specialty provider–yielded low or even zero results across several databases. Using the same algorithm, the sensitivity and PPV could change as much as >30%, depending on which database was used. For example, for algorithm 1, the sensitivity ranged from 38% for MDCR to 66% for JMDC, with PPV ranging from 57% for HealthVerity to 89% for Stanford University.

Examining across the databases used, some databases in general performed better for all algorithms. The academic centers (Johns Hopkins University, Stanford University, and Columbia University) in general have higher PPV (mean 85%–96%) across all algorithms compared with claims databases (70%–80%). Sensitivity was worst for Medicare and Medicaid databases (MDCD and MDCR, average sensitivities across all algorithms were 47% and 44%, respectively).

DISCUSSION

In this study, we have demonstrated the performance of several DM case identification algorithms across a range of different databases, including academic medical centers and US and Japanese insurance claims data. We found a spectrum of

performance for these algorithms depending on the database used and overall identified an incident algorithm (2) and prevalence algorithm (4) with good performance, as evidenced by their sensitivity (42% and 49%, respectively) and PPV (83% and 84%, respectively). A major advantage of this study is understanding how a given algorithm performs on several different data sources, which can not only help researchers select the appropriate algorithm and database when conducting a study but also can enable better clinical trial recruitment by understanding the incidence and prevalence of DM in different countries and academic medical centers.

The observed heterogeneity of algorithm performance across data sources is likely a combination of factors, including variance in underlying patient population characteristics (older individuals versus general population), differences in site-specific mapping and enrichment of OMOP data, variety in clinical documentation and/or data extraction practices, or incongruent data quality. For example, the algorithms requiring immunosuppressant medication orders had similar sensitivity on the Medicare claims database (MDCR) compared with the Johns Hopkins University EHR database, but performance was not as comparable between these two databases for algorithms requiring a second DM diagnosis code. One possible explanation is that MDCR, as a Medicare claims dataset, contains an older population in which a second DM diagnostic code is less likely to appear in the defined follow-up window (for example, if the patient died from cancer, interstitial lung disease, infection, or a non-DM–related cause), whereas prescription-based criteria may still be reliably captured.

A strength of this study is the multiple methods used to validate each case identification algorithm. We reviewed a large number of manual records in one data source (Johns Hopkins University) to support the use of PheValuator ($\kappa = 0.70$) across other data sources. PheValuator enabled the quantification of sensitivity and PPV, despite not having access to patient-level data. To evaluate the clinical plausibility of the cohorts generated by each algorithm, we used CohortDiagnostics to summarize demographic and clinical characteristics. Although the tool itself does not evaluate algorithm validity, its output allowed us to assess alignment with known epidemiologic features of adult DM, such as female predominance and peak onset in the fifth and sixth decades of life. These multiple, complementary approaches increase the likelihood the patients identified were truly those with adult DM.

We chose to focus our DM case identification algorithms leveraging *International Classification of Diseases* (*ICD*) coding, immune-suppressing and/or immune-modulating medications, laboratory testing, and specialty provider interaction, as has been done previously. The requirement for two diagnoses of DM increases the PPV of the algorithm as others have documented; however, this comes at the cost of excluding patients with one-time consultations or those that died after a single visit. The decision to include medications was based on improvements in PPV (patients with "rule out" DM are unlikely to be prescribed disease-modifying antirheumatic drugs and biologic agents), as well as the fact that many key research questions in the field require knowledge of medications, such as comparative effectiveness studies and prediction of medication adverse events.

Our study was greatly enhanced by leveraging the OHDSI community and OMOP CDM. Given that all OHDSI data partners use OMOP, no transfer of patient-level data was necessary. Instead, an R package was created and distributed to all participating data partners, executed in their local environment, and aggregated data shared back to our central site. Such an approach is conducive to generalizable and reproducible research because all code and software is public and open source. Further, given over 10% of the world's population has at least some of its data in OMOP format, the study of rare diseases becomes more achievable. In this study, depending on the algorithm selected, 30,000 to 50,000 adult patients with DM are available to study. With the inclusion of additional data sources, more patients could easily be studied with minimal effort and cost. In addition, this approach is agnostic to disease; other stakeholders in the rheumatologic community can engage the OHDSI network to perform similar observational studies. Finally, many more sophisticated resources within the OHDSI community exist, including open-source software for running comparative effectiveness studies and patient-level prediction, which are of increasing interest as the amount of real-world data continues to grow substantially. With the data generated in this study, analytical techniques such as quantitative bias analysis can be employed to correct for misclassification and provide more accurate estimates.

There are also several limitations to this work. Given many of our case identification algorithms used *ICD*, *Tenth Revision* (*ICD-10*) coding, we could not focus on other IIM subgroups such as antisynthetase syndrome and immune-mediated necrotizing myopathy. Although the Systematized Nomenclature of Medicine - Clinical Terms/OHDSI Standardized Vocabularies include standard concepts for many conditions that extend beyond the clinical granularity covered by *ICD-10, Clinical Modification* (*ICD-10-CM*) codes (eg, antisynthetase syndrome), the extraction of conditions from EHR data is prioritized toward billing codes that constitute the majority of condition codes available in patient records, at least for US health systems mandated to use *ICD-10-CM* codes for reimbursement. This is a particularly relevant limitation when using secondary data to study rare diseases like DM, with very specific subgroups among already small patient populations. It is clear from our study that not all desirable concepts are mapped appropriately from source data to OMOP, such as specialty provider and myositis-specific autoantibodies. As the mapping process from source data to OMOP becomes more refined over time, we expect these data sources to become available, enabling identification of autoantibody-driven subtypes.

Although we used manual chart review at Johns Hopkins University to validate our case identification algorithms and assess the output of PheValuator, we did not perform chart review in the other databases. The Johns Hopkins University review cohort was enriched for DM cases, which may have inflated the PPV of the algorithm in this subset. However, this enrichment does not affect sensitivity, which is independent of disease prevalence. Because our algorithms were applied to large populations and DM is a rare condition, the number of true negatives was very high. Consequently, performance metrics that include true negatives in the denominator, such as specificity and negative predictive value (NPV), had limited interpretability due to the extremely low signal. For this reason, we focused on reporting sensitivity and PPV to assess diagnostic performance. It is also important to note that both PPV and NPV are affected by disease prevalence. Therefore, differences in adult DM prevalence across databases likely contributed to some of the observed variation in PPV estimates. Lastly, due to the deidentified nature of the data, we could not determine whether patients were included in multiple databases. As such, there is a possibility of duplicate individuals (eg, a patient seen at Johns Hopkins University also appearing in a claims dataset), which could introduce additional bias.

In conclusion, we have developed and validated several case identification algorithms for adult DM across multiple data sources, enabling the use of real-world data to study this rare disease. This approach can be used for other rheumatic diseases to improve the generalizability and reproducibility of research findings.

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 Mecoli 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. Lundberg IE, Fujimoto M, Vencovsky J, et al. Idiopathic inflammatory myopathies. Nat Rev Dis Primers 2021;7(1):86.
2. Khoo T, Lilleker JB, Thong BYH, et al H. Epidemiology of the idiopathic inflammatory myopathies. Nat Rev Rheumatol 2023;19(11): 695–712.
3. Kronzer VL, Kimbrough BA, Crowson CS, et al. Incidence, prevalence, and mortality of dermatomyositis: a population-based cohort study. Arthritis Care Res (Hoboken) 2023;75(2):348–355.
4. Bernatsky S, Joseph L, Pineau CA, et al. Estimating the prevalence of polymyositis and dermatomyositis from administrative data: age, sex and regional differences. Ann Rheum Dis 2009;68(7):1192–1196.
5. Bernatsky S, Linehan T, Hanly JG. The accuracy of administrative data diagnoses of systemic autoimmune rheumatic diseases. J Rheumatol 2011;38(8):1612–1616.
6. Dobloug C, Garen T, Bitter H, et al. Prevalence and clinical characteristics of adult polymyositis and dermatomyositis; data from a large and unselected Norwegian cohort. Ann Rheum Dis 2015;74(8):1551–1556.
7. Kwa MC, Ardalan K, Laumann AE, et al. Validation of International Classification of Diseases Codes for the epidemiologic study of dermatomyositis. Arthritis Care Res (Hoboken) 2017;69(5):753–757.
8. Svensson J, Arkema EV, Lundberg IE, et al. Incidence and prevalence of idiopathic inflammatory myopathies in Sweden: a nationwide population-based study. Rheumatology (Oxford) 2017;56(5): 802–810.
9. DeFalco F, Sena AG, Knoll C, et al. ATLAS. OHDSI; 2023. Accessed April 8, 2025. https://github.com/OHDSI/Atlas/wiki
10. Lundberg IE, Tjärnlund A, Bottai M, et al; International Myositis Classification Criteria Project consortium, The Euromyositis register and The Juvenile Dermatomyositis Cohort Biomarker Study and Repository (UK and Ireland). 2017 European League Against Rheumatism/American College of Rheumatology classification criteria for adult and juvenile idiopathic inflammatory myopathies and their major subgroups. Arthritis Rheumatol 2017;69(12):2271–2282.
11. Rao GA, Shoaibi A, Makadia R, et al. CohortDiagnostics: phenotype evaluation across a network of observational data sources using population-level characterization. PLoS One 2025;20(1):e0310634.
12. Swerdel JN, Schuemie M, Murray G, et al. PheValuator 2.0: methodological improvements for the PheValuator approach to semi-automated phenotype algorithm evaluation. J Biomed Inform 2022; 135:104177.
13. Swerdel JN, Hripcsak G, Ryan PB. PheValuator: development and evaluation of a phenotype algorithm evaluator. J Biomed Inform 2019;97:103258.

Arthritis Care & Research
Vol. 78, No. 3, March 2026, pp 371–377
DOI 10.1002/acr.25635

Identifying High-Impact Solutions to Address Racial and Ethnic Health Disparities in Lupus: A Consensus-Based Approach

Joy Buie,[1] Michael P. Fisher,[1] Kristen Backor,[2] Hannah Tyldsley,[2] Ashira Blazer,[3] Candace Feldman,[4] Andrea Knight,[5] S. Sam Lim,[6] Barbara Ann Polk,[7] Ed Yelin,[8] Edith Williams,[9] and Karen H. Costenbader[4]

Objective. We conducted formative research aimed at identifying solutions that address inequitable health outcomes in lupus due to adverse social determinants of health (SDoH).

Methods. We conducted a search for keywords, which provided insights into potential solutions and initiatives underway. An advisory panel of lupus experts iteratively reviewed the list of literature-scoped solutions in working sessions, filling knowledge gaps, which allowed for further defining and classifying solutions based on area of focus, feasibility, and impact. Seven-in-depth semistructured discussions and a modified Delphi survey approach were leveraged to align the advisory panel based on feasibility, impact, and costs of the proposed solutions.

Results. Thirty-three solutions were identified and classified into four key categories: financial safety net, patient education and shared decision-making, physician education, and other solutions. High-impact solutions that were prioritized included the following: "collecting granular information like patient-reported outcomes to provide personalized care and accelerate development of new products," "expanding Medicaid coverage via infrastructure," and "supporting people living with lupus in applying and getting approval for disability."

Conclusion. Addressing health and health care disparities linked to negative SDoH is a key goal in the management of lupus, as disparities in outcomes can be stark. Increasing the visibility of potential solutions and aligning the community on top priorities can enable more efficient and effective contributions to health care equity and ultimately better health outcomes for people living with lupus.

INTRODUCTION

Systemic lupus erythematosus (SLE) is a chronic, heterogeneous autoimmune disease characterized by systemic inflammation and potential for irreversible multiorgan damage. Although advancements in care since the 1940s have markedly improved 5- and 10-year survival rates to over 90%,[1] substantial disparities persist in health outcomes and death. Lupus mortality rates are disproportionately higher in historically marginalized racial and ethnic groups within the United States.[2] Specifically, Black/African American, Hispanic/Latino, and Indigenous populations face remarkably higher rates of poor health outcomes and mortality compared to White patients.[2]

Despite the recent approval of new therapies, such as anifrolumab, belimumab, and voclosporin, which has expanded treatment options for SLE, cutaneous lupus erythematosus, and lupus nephritis, access to these innovations remains inequitable due to lack of economic resources among minority groups. Structural barriers, such as socioeconomic instability and broader negative social determinants of health (SDoH), hinder equitable

Supported by Genentech, Inc, a member of the Roche Group; AstraZeneca PLC; Mallinckrodt Pharmaceuticals; Novartis, Inc; Biogen; EMD Serono; GlaxoSmithKline; and Bristol Meyers Squibb.

[1]Joy Buie, PhD, MSCR, RN, Michael P. Fisher, PhD: Lupus Foundation of America, Washington, DC; [2]Kristen Backor, PhD, Hannah Tyldsley, MS: Charles River Associates, San Francisco, California; [3]Ashira Blazer, MD: University of Maryland School of Medicine, Baltimore; [4]Candace Feldman, MD, ScD, PhD, Karen H. Costenbader, MD, MPH: Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts; [5]Andrea Knight, MD: The Hospital for Sick Children and University of Toronto, Toronto, Ontario, Canada; [6]S. Sam Lim, MD: Emory University School of Medicine, Atlanta, Georgia; [7]Barbara Ann Polk, BS: Jobs for the Future, Boston, Massachusetts; [8]Ed Yelin, PhD: University of California, San Francisco; [9]Edith Williams, PhD: University of Rochester Medical Center, Rochester, New York.

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.25635).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25635.

Address correspondence via email to Joy Buie, PhD, MSCR, RN, at buie@lupus.org.

Submitted for publication April 16, 2025; accepted in revised form August 4, 2025.

SIGNIFICANCE & INNOVATIONS

- This work demonstrates a systematic, comprehensive approach to identifying priority interventions that have the potential to help mitigate inequitable health outcomes among individuals with lupus, particularly those most impacted by adverse social determinants of health.
- Our findings underscore that facilitating access to disability benefits, bolstering access to treatment, ensuring early detection of lupus risk factors, and implementing nurse or social worker–led strategies to enhance follow-up care are pivotal for addressing the multilayered impact of adverse social determinants of health on lupus disease management and outcomes, especially among historically marginalized communities in the United States.

distribution of these advancements, leaving vulnerable populations unable to benefit equally.[3]

Socioeconomic status (SES) plays a pivotal role in shaping health disparities in SLE. Low SES restricts housing choices,[4] often confining individuals to neighborhoods marked by higher crime rates, environmental toxins, and food insecurity.[5] Exposure to low SES exacerbates disease progression and results in greater cumulative damage, contributing to higher mortality rates in low SES populations.[6] Additionally, reliance on public health insurance in the United States often leads to fragmented and inconsistent care, limiting access to specialists and safe, comprehensive therapies.[7] Such discontinuity is especially detrimental in conditions such as lupus nephritis, in which unmanaged disease significantly increases the risk of irreversible organ damage, kidney failure, cardiovascular complications, and other adverse outcomes.

Neighborhood-level disadvantages, as reflected by higher Area Deprivation Index (ADI) scores, further compound disparities. High ADI scores are associated with increased disease severity and damage in conditions such as discoid lupus erythematosus and lupus nephritis.[8,9] Without targeted interventions addressing these structural inequities, patients from historically marginalized communities remain at heightened risk for morbidity, mortality, and reduced quality of life.

Although the link between negative SDoH, low SES, and poor lupus outcomes is well recognized,[10–12] systematic efforts to identify and prioritize actionable solutions remain lacking. Currently, there is no comprehensive road map for strategies to address these challenges. Instead, public health and medical professionals often work in fragmented ways, tackling health disparities without coordination. Given the complexity and scale of SDoH-related issues in lupus, engaging experts with diverse knowledge of the disease is essential for creating impactful solutions.

The Lupus Addressing Health Inequities in Minorities (Lupus AIM) Advisory Panel was established to address this gap. Through a structured Delphi process, the panel identified key causes of lupus-related health disparities and proposed targeted solutions to reduce them. This article presents a prioritized, expert-informed road map for addressing SDoH in lupus, which were presented in a previous article.[10] By aligning these efforts across the lupus community, we aim to close existing health equity gaps and improve the quality of life for individuals living with lupus.

METHODS

Overview of the Lupus AIM program. The Lupus AIM program is composed of multiple phases to systematically identify health disparities and prioritize solutions addressing health disparities in SLE. Phase 1 focused on identifying challenges, whereas phases 2 and 3 involved solution generation and prioritization, which are the focus of this consensus statement. In phase 1, the Lupus AIM Advisory Panel conducted a systematic literature review to identify major epidemiologic and social determinants that negatively impact lupus outcomes. This review was initially conducted in January 2021 and updated in 2023. Relevant research was identified through PubMed using the following search terms: [“lupus” or “lupus nephritis” or “systemic lupus erythematosus”] AND [“disparities” or “social determinants of health” or “inequity”] OR [[“race” or “ethnicity”] AND [“prevalence” or “outcome” or “treatment” or “therapy”]]. The review was limited to English-language human studies published between 2011 and 2023. Findings from the literature review informed a scientific statement published[10] in 2023.

Building on these findings, a modified Delphi approach was employed in phases 2 and 3 to achieve expert consensus on priority solutions for addressing SDoH and reducing racial and ethnic health disparities in lupus outcomes. Delphi methods are widely used in health sciences to obtain structured input from experts.[13] The approach used here included four key steps[14,15]:

1. Panel identification: convene a multidisciplinary panel of experts in lupus and health disparities.
2. Problem identification: define key factors driving health disparities in lupus, focusing on social inequities and barriers to access.
3. Iterative refinement: conduct iterative rounds of solution development and refinement based on panel input.
4. Consensus determination: prioritize solutions based on expert consensus.

This approach differed from the traditional Delphi method because anonymity of participants was not a goal. Participant anonymity offers both benefits and drawbacks[16] and was not deemed necessary in this study given the productive dialogue expected from advisory panel meetings and workshops.

Composition of the Lupus AIM Advisory Panel. The Lupus AIM Advisory Panel was convened in October 2020 to address health disparities in SLE among racial and ethnic minority communities in the United States. The panel comprises 13 members, including 2 patients, 6 clinician-scientists, 3 research professors, and 2 representatives from nonprofit and voluntary health organizations. Members were selected for their expertise in lupus, health disparities, and public health, with recruitment led by Dr Karen H. Costenbader, a researcher in the field. Panel diversity ensured that multiple perspectives were represented, enabling a comprehensive approach to identifying solutions. This study was reviewed by the Advarra, Inc Institutional Review Board and deemed exempt.

Delphi methodology for solution identification and prioritization. A four-round iterative Delphi process was implemented to develop and refine solutions for reducing health disparities in lupus outcomes as outlined in Table 1. Each round involved structured discussions, workshops, and surveys to iteratively identify, refine, and prioritize solutions.

The solution identification and generation phase involved a series of advisory panel meetings from May 2021 to August 2022. During these sessions, the panel identified structural factors contributing to racial and ethnic inequities in lupus outcomes and access barriers exacerbating health disparities. These discussions resulted in a preliminary list of 42 solutions. The panel conducted a prioritization exercise, ranking each solution based on its potential impact on addressing disparities and its feasibility for implementation by the Lupus Foundation of America (LFA) using a 1 (low) to 3 (high) scale. Combined scores for feasibility and impact were used to rank solutions, with the top ~50% forming an initial prioritized list of 22 solutions.

The solution refinement phase, conducted between September 2023 and January 2024, updated the list to reflect recent advancements in industry and academic thought. This phase incorporated input from structured 30-minute interviews with seven industry experts and members of the Lupus AIM Advisory Panel as well as insights from a panel workshop. Feedback from these activities resulted in an updated list of 34 solutions, integrating previously prioritized solutions from 2022, findings from 2023 research, and new ideas generated during discussions. These solutions were categorized into four themes: financial safety net, patient education and shared decision-making, physician education, and other solutions.

The Solution Prioritization Survey 1 was conducted during a final workshop in January 2024. Panel members re-evaluated the updated solutions based on their impact, feasibility, and financial cost using a 1 (low) to 3 (high) scale. Qualitative feedback from the workshop was also incorporated, leading to a refined list of 33 solutions (4 solutions consolidated into 2 solutions, 2 solutions updated, and 1 new solution added). Quantitative data from the survey were analyzed in Microsoft Excel to calculate mean scores for each metric.

In Solution Prioritization Survey 2, conducted online in April 2024, panel members validated the updated solutions by reassessing their impact on addressing disparities in lupus outcomes. Using the final list of 33 solutions, members rated each solution's impact on a scale of 1 (low) to 3 (significant/major impact). Survey 2 measured consistency in response, as all respondents had also participated in survey 1. Consensus was determined using a two-factor process assessing both the percentage of panel members who rated a solution as having significant/major impact (score of 3) and the mean impact rating.

Both surveys were conducted anonymously; although panel members knew the identity of other panelists, ratings for specific solutions were collected independently using an online survey. Panelists did not know others' individual responses. Data from Solution Prioritization Survey 2 was analyzed in Microsoft Excel to determine the final priority solutions.

Consensus determination. Consensus was defined as follows:

Table 1. Lupus AIM solution identification and prioritization research and consensus process*

Round	Date(s)	Activity/activities	Output(s)
Round 1a: Solution identification and generation	May 2021 to August 2022	Advisory panel meetings Solution item generation	42 solutions
Round 1b: Solution identification and generation		Preliminary prioritization exercise	Prioritized list of 22 solutions
Round 2: Solution refinement	September 2023 to January 2024	Literature search 1:1 discussion with industry experts	34 solutions, categorized by topic
Round 3: Solution Prioritization Survey 1	January 2024	Advisory panel workshop Solution prioritization survey (feasibility, impact, cost)	Solution Prioritization Survey 1 rating for each of the 34 solutions Refined list of 33 solutions
Round 4: Solution Prioritization Survey 2	April 2024	Solution prioritization survey (impact)	Solution Prioritization Survey 2 results for each of the 33 solutions Final prioritized list of top 12 solutions based on impact

* Lupus AIM, Lupus Addressing Health Inequities in Minorities.

1. Consensus criterion: >50% of the panel rated a solution as having significant/major impact (score of 3).
2. High-impact criterion: the mean impact rating for a solution exceeded the median rating across all solutions.

Solutions meeting both criteria were identified as priority interventions for addressing disparities in lupus outcomes. This structured and proven process ensured a robust and transparent methodology for prioritization.[17]

The second survey identified 12 priority solutions based on both consensus (>50% of panel members rating the solution as having a significant/major impact) and a mean impact rating exceeding the median across all solutions (2.36). The results are summarized in Table 2 (sorted by level of consensus). Large language models were used minimally for checking spelling and grammar errors within the article.

RESULTS

Rank 1 priority solution. The most impactful solution was identified as follows:

1. Supporting individuals with lupus in navigating and securing disability benefits

This solution received a mean rating of 2.73, with 82% of the panel agreeing it would have a significant impact on reducing disparities in lupus outcomes. The complex process of applying for disability benefits can hinder access for many, and resources such as LFA's "Applying for Social Security Disability" webpage (https://www.lupus.org/resources/what-to-know-about-social-security-disability-insurance) provide critical guidance to improve approval rates.

Rank 2 priority solutions. Three solutions were ranked equally as the second-highest priority, each receiving a mean impact rating of 2.64, with 64% panel consensus on their potential to address disparities:

1. Expanding third-party grants and funding for lupus treatments, exemplified by initiatives such as the Patient Navigator Grant Program by Aurinia, to enhance treatment affordability
2. Increasing screening for lupus risk factors in high-risk communities to facilitate earlier diagnosis and health care access
3. Leveraging nurse and social worker–guided resource identification during follow-up visits to provide additional support, including community-focused and health-related resources

Rank 3 priority solutions. Four solutions ranked as the third priority, with a mean impact rating of 2.45 and 64% panel consensus. However, the panel expressed greater variability in their assessments, with some members considering these solutions to have limited impact:

Table 2. Results of Lupus AIM Advisory Panel solution prioritization exercise (in order of consensus)*

Ranking	Category	Solution	Consensus, %	Mean rating
1	Other	Support people living with lupus in applying and getting approval for disability	82	2.73
2	Financial safety net	Third-party grants and funding for lupus initiatives	64	2.64
2	Education and shared decision-making	Aim to increase screening for lupus risk factors in high-risk communities and increase early access to health care	64	2.64
2	Physician education	Employ nurse and social worker–guided resource identification and interview to enhance patient follow-up visits	64	2.64
3	Financial safety net	Create resources that will help patients and providers locate patient-level financial assistance and other programs supporting care-related costs for commercially insured patients	64	2.45
3	Education and shared decision-making	Provide education regarding pregnancies and fetal death in diverse communities with lupus	64	2.45
3	Education and shared decision-making	Provide nonmedical support for un(der)insured patients navigating the health care system through creating resources to complement clinical education	64	2.45
3	Physician education	Strengthen the social care aspect of lupus in addition to the health care aspect by leveraging social workers and CHWs and integrating social workers with practices	64	2.45
4	Physician education	Provide comprehensive lupus education to frontline providers practicing in underresourced communities	55	2.55
5	Financial safety net	Expand Medicaid coverage via an infrastructure bill by lobbying	55	2.45
5	Other	Leverage government resources and coalitions for funding opportunities and collaborations	55	2.45
5	Education and shared decision-making	Multilingual education resources created by native speakers of other languages (ie, not just Spanish)	55	2.45

* CHW, Community Health Worker; Lupus AIM, Lupus Addressing Health Inequities in Minorities.

1. Developing patient-level financial assistance resources by building on initiatives such as LFA's "Financial Assistance Resources" webpage, the Lupus Research Alliance's Financial Resource Center, and the Patient Access Network Foundation's SLE Medicine Assistance Program
2. Enhancing education on pregnancy and fetal death risks in diverse lupus populations to address maternal health disparities
3. Providing nonmedical support for underinsured patients, such as through LFA's "Ask a Health Educator" program, to complement clinical care and help patients navigate the health care system
4. Strengthening the social care aspect of lupus management by incorporating social workers and leveraging programs such as the National Institutes of Health Integrated Care Management Program.

Rank 4 priority solution. The fourth priority, with 55% panel consensus and a mean impact rating of 2.55, was the following:

1. Providing comprehensive lupus education to frontline providers in underresourced communities

Rank 5 priority solutions. Three solutions were ranked as the fifth priority, each with 55% panel consensus but a lower mean impact rating of 2.45, reflecting mixed opinions on their potential impact:

1. Expanding Medicaid coverage through advocacy for legislative initiatives such as infrastructure bills to improve treatment access for low-income populations
2. Using government resources and coalitions to secure additional funding opportunities (no specific examples identified)
3. Developing multilingual educational resources, such as GSK's Patient Learning and Community Education Hub, to address racial and ethnic health disparities in lupus outcomes.

DISCUSSION

Using a modified Delphi process, a multidisciplinary panel of health equity experts from the Lupus AIM Advisory Panel prioritized 12 solutions to address SDoH contributing to racial and ethnic disparities in lupus outcomes. Through iterative rounds of solution nominations, structured feedback, surveys, and a comprehensive literature review, the panel reached consensus on solutions spanning four key areas: improving financial safety nets, enhancing physician education, advancing patient education and shared decision-making, and addressing miscellaneous challenges categorized as "other." The diversity of these solutions underscores the broad scope of SDoH factors that, if left unaddressed, negatively impact health outcomes in SLE.

The top 12 priority solutions (Table 1) span all four thematic categories: financial safety net, patient education and shared decision-making, physician education, and other solutions. Patient education and shared decision-making emerged as the most represented category, accounting for 33% (4 of 12) of the prioritized solutions despite representing only 24% (7 of 33) of the comprehensive solutions. This suggests that the panel viewed these solutions as particularly impactful in addressing health disparities. Financial safety net solutions accounted for 25% of the prioritized solutions, consistent with their representation in the comprehensive list. Physician education and other solutions were slightly underrepresented in the priority list, making up 25% and 17%, respectively, compared to 30% and 21% in the comprehensive list. These patterns highlight the panel's focus on patient-centric interventions.

The panel identified supporting individuals with lupus in applying for and securing disability benefits as the most impactful and highly feasible solution, with 82% consensus and a mean impact rating of 2.73. This solution addresses a critical need for the most vulnerable populations, particularly those with severe symptoms that limit employability. Resources such as LFA's "Applying for Social Security Disability" webpage were identified as playing a key role in facilitating this process. Financial solutions also ranked highly, including third-party grants and funding programs such as the Patient Navigator Grant Program and Lupus Copay Assistance Program. These initiatives aim to reduce the financial burden of lupus care, particularly for underserved populations, which represents a major need in the lupus community and, when addressed, has implications for improving health outcomes.[18] Panelists emphasized the importance of extending financial support to cover care-related costs, even for patients with commercial insurance. Timely diagnosis and early intervention were identified as pivotal for reducing disparities. Solutions such as increased screening for lupus risk factors in high-risk communities and improving early access to care were prioritized. The prioritization of this solution aligns with the current observation that it takes, on average, six years and four physicians to receive a definite lupus diagnosis.[19] This problem is the most severe in underresourced communities. African Americans with lupus are 10 times more likely than White individuals to live in highly disadvantaged neighborhoods,[20] where limited access to primary care and preventive services exacerbates health inequities. This lack of early intervention contributes to the more severe disease manifestations observed shortly after diagnosis in these populations.[21] Interventions such as Medicaid expansion, leveraging government resources for funding opportunities, and creating multilingual resources beyond Spanish received lower consensus. Panelists, however, expressed concerns regarding the feasibility of these recommendations, thereby casting doubt on their potential for immediate impact. Although these solutions address critical systemic

issues, their complexity, resource demands, and the current political climate may have influenced their lower prioritization.

This study represents the first systematic effort to prioritize solutions targeting SDoH-related disparities in lupus outcomes using a rigorous modified Delphi process. The inclusion of a multidisciplinary panel with expertise in lupus and health equity from across the United States ensured diverse perspectives. The iterative methodology facilitated inclusive exploration of potential interventions, allowing the panel to refine and prioritize actionable, feasible recommendations. Additionally, the integration of a systematic literature review added robustness to the findings, and the panelists' extensive experience in domestic and international research on lupus disparities bolstered the study's credibility.

The relatively small size of the panel may limit the generalizability of the findings. Although the panel included representation from the West Coast, most members were based in the Southeast and Northeast, potentially biasing the results due to underrepresentation from other US regions, such as the Midwest or rural areas. Solutions for lupus management in these underrepresented regions may require different emphases; for instance, rural communities often face significant scarcities of rheumatology specialists and greater travel burdens, potentially necessitating a greater reliance on telehealth initiatives, mobile health units, or enhanced roles for primary care physicians in comanaging SLE; rural residence has also been independently associated with adverse outcomes, such as an increased risk of myocardial infarction hospitalization in patients with SLE,[22] and may involve delays in diagnosis.[23] It is important to note, however, that the panel's geographic makeup aligns with the current distribution of lupus specialist concentrations and academic rheumatology centers in the United States. Furthermore, these represented regions correspond with areas of high lupus prevalence and where significant health disparities are extensively documented. In addition, several panel members possess dedicated expertise and focus their research on traditionally underrepresented and underserved populations, including Indigenous communities, who experience a high burden of lupus and significant barriers to care, as well as individuals residing in rural settings. Their contributions were invaluable in helping to extrapolate findings and consider the applicability of proposed solutions to diverse contexts beyond their immediate geographic locations. Although the regional emphasis is acknowledged as a limitation, these factors related to regional differences in patient distribution, as well as specific panelist expertise, provide important context. Furthermore, one panelist was unable to participate in both prioritization surveys, and their insights are not reflected in the final consensus. The use of a modified Delphi process, including a three-point rating scale rather than a traditional binary scale, posed methodologic challenges. Although this approach captured nuanced opinions, determining consensus was more complex and required alternative analytical methods (as detailed in the Methods section). Additionally, the study included limited input from patients, whose perspectives are essential for understanding the real-world impact of the proposed solutions (see Supplemental Document 1 for additional details on this perspective). Future work should incorporate more robust patient voices to validate and refine these recommendations.

This rigorous data and expert consensus–driven study identified 12 priority solutions to address SDoH-related disparities in lupus outcomes, emphasizing the importance of patient empowerment, financial support, and early intervention. These findings provide a road map for addressing health inequities in racially and ethnically diverse populations with lupus that can guide subsequent work to assess optimal next steps, responsible parties, and metrics for success. Future efforts should integrate patient perspectives to ensure that these solutions align with their needs.

ACKNOWLEDGMENT

We would like to thank the Lupus AIM Advisory Panel for participating in this work and making valuable contributions to this effort.

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. Buie 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 SPONSORS

Genentech, Inc; AstraZeneca PLC; Mallinckrodt Pharmaceuticals; Novartis, Inc; Biogen; EMD Serono; GlaxoSmithKline; and Bristol Meyers Squibb had no role in the study design or in the collection, analysis, or interpretation of the data, the writing of the manuscript, or the decision to submit the manuscript for publication. Publication of this article was not contingent upon approval by Genentech, Inc; AstraZeneca PLC; Mallinckrodt Pharmaceuticals; Novartis, Inc; Biogen; EMD Serono; GlaxoSmithKline; and Bristol Meyers Squibb.

REFERENCES

1. Mak A, Cheung MW, Chiew HJ, et al. Global trend of survival and damage of systemic lupus erythematosus: meta-analysis and meta-regression of observational studies from the 1950s to 2000s. Semin Arthritis Rheum 2012;41(6):830–839.
2. Yen EY, Shaheen M, Woo JMP, et al. 46-Year trends in systemic lupus erythematosus mortality in the United States, 1968 to 2013: a nationwide population-based study. Ann Intern Med 2017;167(11):777–785.
3. Liang MH, Lew ER, Fraser PA, et al. Choosing to end African American health disparities in patients with systemic lupus erythematosus. Arthritis Rheumatol 2024;76(6):823–835.

4. Yelin E, Yazdany J, Trupin L. Relationship between poverty and mortality in systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2018;70(7):1101–1106.
5. Yelin E, Trupin L, Bunde J, et al. Poverty, neighborhoods, persistent stress, and systemic lupus erythematosus outcomes: a qualitative study of the patients' perspective. Arthritis Care Res (Hoboken) 2019;71(3):398–405.
6. Garg S, Sweet N, Boderman B, et al. Multiplicative impact of adverse social determinants of health on outcomes in lupus nephritis: a meta-analysis and systematic review. Arthritis Care Res (Hoboken) 2024; 76(9):1232–1245.
7. Walunas TL, Jackson KL, Chung AH, et al. Disease outcomes and care fragmentation among patients with systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2017;69(9):1369–1376.
8. Faden DF, Xie L, Stone C, et al. Area deprivation and disease severity in adult patients with discoid lupus erythematosus. JAMA Dermatol 2024;160(9):984–988.
9. Soulsby WD, Olveda R, He J, et al; CARRA Registry Investigators. Racial disparities and achievement of the low lupus disease activity state: a CARRA Registry study. Arthritis Care Res (Hoboken) 2025; 77(1):38–49.
10. Buie J, McMillan E, Kirby J, et al. Disparities in lupus and the role of social determinants of health: current state of knowledge and directions for future research. ACR Open Rheumatol 2023;5(9):454–464.
11. González LA, Ugarte-Gil MF, Pons-Estel GJ, et al. Addressing health disparities as a function of ethnicity in systemic lupus erythematosus patients. Lupus 2022;31(14):1691–1705.
12. Gilcrease GW, Sciascia S, Padovan D, et al. Health inequalities and social determinants of health: the role of syndemics in rheumatic diseases. Autoimmun Rev 2023;22(7):103351.
13. Shang Z. Use of Delphi in health sciences research: a narrative review. Medicine (Baltimore) 2023;102(7):e32829.
14. Scott TE, Costich M, Fiorino EK, et al. Using a modified Delphi methodology to identify essential telemedicine skills for pediatric residents. Acad Pediatr 2023;23(3):511–517.
15. Nasa P, Jain R, Juneja D. Delphi methodology in healthcare research: how to decide its appropriateness. World J Methodol 2021;11(4): 116–129.
16. Khodyakov D. Generating evidence using the Delphi method. October 17, 2023. Accessed January 9, 2025. https://www.rand.org/pubs/commentary/2023/10/generating-evidence-using-the-delphi-method.html
17. Schmajuk G, Hoyer BF, Aringer M, et al; SLE Classification Criteria Steering Committee and the International SLE Expert Panel of the Initiative. Multicenter Delphi exercise to identify important key items for classifying systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2018;70(10):1488–1494.
18. Yelin E, Trupin L, Yazdany J. A prospective study of the impact of current poverty, history of poverty, and exiting poverty on accumulation of disease damage in systemic lupus erythematosus. Arthritis Rheumatol 2017;69(8):1612–1622.
19. Bruce IN, Buie J, Bloch L, et al. Lupus spectrum ambiguity has long-term negative implications for patients. Lupus Sci Med 2023;10(1): e000856.
20. Parodis I, Lanata C, Nikolopoulos D, et al. Reframing health disparities in SLE: a critical reassessment of racial and ethnic differences in lupus disease outcomes. Best Pract Res Clin Rheumatol 2023;37(4):101894.
21. Lim SS, Helmick CG, Bao G, et al. Racial disparities in mortality associated with systemic lupus erythematosus - Fulton and DeKalb Counties, Georgia, 2002-2016. MMWR Morb Mortal Wkly Rep 2019;68(18):419–422.
22. Singh JA, Chandrupatla S. Rural-urban disparities in hospitalisation for myocardial infarction in systemic lupus erythematosus in the USA. Lupus Sci Med 2025;12(1):e001516.
23. Lennep DS, Crout T, Majithia V. Rural health issues in rheumatology: a review. Curr Opin Rheumatol 2020;32(2):119–125.

Arthritis Care & Research
Vol. 78, No. 3, March 2026, pp 378–387
DOI 10.1002/acr.25640

Prevalence, Determinants, and Outcomes of Low Disease Activity and Remission Attainment in Patients With Systemic Lupus Erythematosus That Is Clinically Active

Yanjie Hao,[1] Dylan Hansen,[2] Rangi Kandane-Rathnayake,[3] Worawit Louthrenoo,[4] Yi-Hsing Chen,[5] Jiacai Cho,[6] Aisha Lateef,[7] Laniyati Hamijoyo,[8] Shue Fen Luo,[9] Yeong-Jian Jan Wu,[9] Sandra Navarra,[10] Leonid Zamora,[10] Zhanguo Li,[11] Sargunan Sockalingam,[12] Yasuhiro Katsumata,[13] Masayoshi Harigai,[13] Zhuoli Zhang,[14] Madelynn Chan,[15] Jun Kikuchi,[16] Tsutomu Takeuchi,[16] Sang-Cheol Bae,[17] Fiona Goldblatt,[18] Sean O'Neill,[19] Kristine (Pek Ling) Ng,[20] Annie Law,[21] B. M. D. B. Basnayake,[22] Nicola Tugnet,[23] Sunil Kumar,[24] Cherica Tee,[25] Michael Tee,[25] Naoaki Ohkubo,[26] Yoshiya Tanaka,[26] Shirley Chan,[27] C. S. Lau,[27] Vera Golder,[3] Alberta Hoi,[3] Shereen Oon,[28] Eric Morand,[3] and Mandana Nikpour,[29] for the Asia-Pacific Lupus Collaboration

Objective. This study aimed to identify in patients with systemic lupus erythematosus (SLE) with clinically active disease the attainment of frequency and determinants of Lupus Low Disease Activity State (LLDAS) and Definition of Remission in SLE (DORIS) and the frequency and determinants of flare and damage accrual after target attainment.

Methods. Patients in a multinational cohort with SLE who had clinical disease activity but were not in LLDAS or DORIS were observed prospectively.

Results. A total of 1,991 patients (93.2% female) were observed for a median (interquartile range) of 2.5 (0.7–4.5) years, with 70.9% and 55.6% achieving LLDAS and DORIS, respectively. Nephritis and low complements were associated with a longer time, and antimalarial and immunosuppressant use were associated with a shorter time to LLDAS attainment. After the first LLDAS and DORIS attainment, 47.0% and 47.5% of the patients experienced flare(s), respectively, and 9.5% and 7.9 % of patients accrued organ damage within 24 months, respectively. Longer cumulative time at target and antimalarial use was associated with a longer time to flare and damage accrual, whereas dose reduction in glucocorticoids and immunosuppressants was associated with a shorter time to flare. Reduction in immunosuppressants also correlated with a shorter time to damage accrual.

Conclusion. In patients with SLE with clinical disease activity, the proportion attaining LLDAS and DORIS under usual care conditions is suboptimal. Longer maintenance of these states is significantly associated with reduced risk of flare. Because flares and damage accrual still occur frequently following initial target attainment, further research is needed to inform strategies for maintaining these targets.

INTRODUCTION

There is increasing interest in adopting the principle of treat-to-target in systemic lupus erythematosus (SLE). The recommendations for the management of SLE have emphasized the importance of treat-to-target in SLE and proposed "remission, as defined by the Definition of Remission in SLE (DORIS) criteria, or alternately, low disease activity states, such as the Lupus Low Disease Activity State (LLDAS), as treatment targets."[1] Multiple large prospective and retrospective studies have demonstrated

This work was presented as an oral presentation at the American College of Rheumatology 2023 Convergence in San Diego, California, on November 10–15, 2023. Hao Y, Hansen D, Kandane-Rathnayake R, et al. Prevalence, determinants and outcomes of target attainment in SLE patients with clinically active disease in a large multinational prospective lupus cohort. *Arthritis Rheumatol* 2023; 75 (suppl 9). https://acrabstracts.org/abstract/prevalence-determinants-and-outcomes-of-target-attainment-in-sle-patients-with-clinically-active-disease-in-a-large-multinational-prospective-lupus-cohort/.

The funders had no role in study design, data collection, data analysis, or writing of the report.

Supported by AstraZeneca. Dr Hao's work was supported by a Melbourne research scholarship from The University of Melbourne. Dr Bae's work was supported in part by a Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Education (NRF-2021R1A6A1A03038899). Dr Morand's work was supported by a National Health and Medical Research Council (NHMRC) Investigator Grant. Dr Nikpour's work was supported by an NHMRC Investigator Grant (GNT1176538). The Asia Pacific Lupus Collaboration was supported by unrestricted project grants from AstraZeneca, Bristol-Myers Squibb, Eli Lilly, GlaxoSmithKline, Janssen Research and Development, EMD Serono,

SIGNIFICANCE & INNOVATIONS

- The present study focused on patients with systemic lupus erythematosus (SLE) with clinically active disease, defined according to the clinical domains of the Systemic Lupus Erythematosus Disease Activity Index-2000.
- Target attainment, including low disease activity and remission, is not optimal in this group of patients under usual care conditions, primarily managed with conventional therapies, with 30% not achieving either target over a median of 2.5 years.
- Flare and damage accrual are still substantial after reaching low disease activity or remission. Maintaining these targets for longer may reduce the risk of flares and organ damage, and tapering immunosuppressive treatment after target attainment is a risk factor for these adverse outcomes.
- This is the first study to comprehensively evaluate target attainment and subsequent clinical outcomes in patients with clinically active disease in a large prospective cohort with SLE in the Asia–Pacific region. Prospective clinical trials of a treating-to-target strategy in SLE are needed, and data from the present study can inform the design of such clinical trials.

that in patients with SLE, attainment of these targets is associated with improved outcomes, including reduced damage accrual, better health-related quality of life (HRQoL), and improved survival.[2–6] Although studies have indicated that attainment of low disease activity or remission is protectively associated with reduced damage accrual in patients with high disease activity (Systemic Lupus Erythematosus Disease Activity Index-2000 [SLEDAI-2K] ≥ 6) at baseline,[3] significant knowledge gaps remain concerning the timing and impacts of remission or low disease activity attainment in the larger subset of patients who have any clinically active disease, who are likely to be the patients recruited into future treating-to-target intervention studies.[7]

We undertook a retrospective analysis of patients in the large multinational prospective Asia-Pacific Lupus Collaboration (APLC) cohort, under usual care conditions. We observed patients from their first visit with any clinical disease activity to identify (i) frequency and time to attainment of low disease activity and remission, defined by the LLDAS and DORIS criteria, respectively,[2,8] (ii) determinants of attaining these targets, (iii) frequency and time to flare and damage accrual subsequent to LLDAS or DORIS attainment, and (iv) determinants of developing flare and damage accrual following target attainment.

PATIENTS AND METHODS

Patients. This study included all patients enrolled in the APLC cohort between March 2013 and December 2020 who newly developed any clinical disease activity following at least one visit with clinical inactivity. Clinical disease activity was defined as the clinical domains of the SLEDAI-2K (cSLEDAI-2K) >0 but were not in DORIS or LLDAS. Baseline was defined as the first observed clinical activity following an observed period of inactivity to ensure time from the onset of activity to LLDAS or DORIS could be correctly calculated. Patients were observed from this baseline. A subgroup of patients who newly developed high disease activity, defined as SLEDAI-2K ≥10, was also observed from their first fulfilling high disease activity criteria.

The APLC cohort is a prospective multinational observational cohort of patients with SLE, including 25 sites across

and Union Chimique Belge to support data collection and data and project management.

[1]Yanjie Hao, MD: The University of Melbourne and St. Vincent's Hospital Melbourne, Melbourne, Victoria, Australia, and Peking University First Hospital, Beijing, China; [2]Dylan Hansen, MBioStat: St. Vincent's Hospital Melbourne, Melbourne, Victoria, Australia; [3]Rangi Kandane-Rathnayake, PhD, Vera Golder, PhD, Alberta Hoi, PhD, Eric Morand, PhD: Monash University, Melbourne, Victoria, Australia; [4]Worawit Louthrenoo, MD: Chiang Mai University Hospital, Chiang Mai, Thailand; [5]Yi-Hsing Chen, MD: Taichung Veterans General Hospital, Taichung, Taiwan; [6]Jiacai Cho, MBBS: National University Hospital, University Medicine Cluster, Singapore; [7]Aisha Lateef, MBBS: Woodlands Health, Singapore; [8]Laniyati Hamijoyo, MD: Padjadjaran University, Hasan Sadikin General Hospital, Bandung, Indonesia; [9]Shue Fen Luo, MD, Yeong-Jian Jan Wu, MD: Chang Gung Memorial Hospital, Taoyuan City, Taiwan; [10]Sandra Navarra, MD, Leonid Zamora, MD: University of Santo Tomas Hospital, Manila, Philippines; [11]Zhanguo Li, PhD: People's Hospital Peking University Health Science Center, Beijing, China; [12]Sargunan Sockalingam, MBBS: University of Malaya, Kuala Lumpur, Malaysia; [13]Yasuhiro Katsumata, PhD, Masayoshi Harigai, PhD: Tokyo Women's Medical University, Tokyo, Japan; [14]Zhuoli Zhang, PhD: Peking University First Hospital, Beijing, China; [15]Madelynn Chan, MD: Tan Tock Seng Hospital, Singapore; [16]Jun Kikuchi, MD, Tsutomu Takeuchi, PhD: Keio University, Tokyo, Japan; [17]Sang-Cheol Bae, PhD: Hanyang University Hospital for Rheumatic Diseases, Hanyang University Institute for Rheumatology Research, and Hanyang Institute of Bioscience and Biotechnology, Seoul, South Korea; [18]Fiona Goldblatt, PhD: Flinders Medical Centre, Adelaide, South Australia, Australia; [19]Sean O'Neill, PhD: Liverpool Hospital, Liverpool, New South Wales, Australia; [20]Kristine (Pek Ling) Ng, BScMed, MBBS: Health New Zealand Waitemata, Auckland, New Zealand; [21]Annie Law, MBBS: Singapore General Hospital, Singapore; [22]B. M. D. B. Basnayake, MD: Teaching Hospital Kandy, Kandy, Sri Lanka; [23]Nicola Tugnet, MBChB: Greenlane Clinical Centre, Auckland, New Zealand; [24]Sunil Kumar, MBBS: Middlemore Hospital, Auckland, New Zealand; [25]Cherica Tee, MD, Michael Tee, MD: University of the Philippines, Manila, Philippines; [26]Naoaki Ohkubo, MD, Yoshiya Tanaka, PhD: University of Occupational and Environmental Health, Kitakyushu, Japan; [27]Shirley Chan, MBBS, C. S. Lau, MD: The University of Hong Kong, Hong Kong, China; [28]Shereen Oon, PhD: The University of Melbourne and St. Vincent's Hospital Melbourne, Melbourne, Victoria, Australia; [29]Mandana Nikpour, PhD: The University of Melbourne and St. Vincent's Hospital, Melbourne, Victoria and The University of Sydney and Royal Prince Alfred Hospital, Sydney, New South Wales, Australia.

Additional supplementary information cited in this article can be found online in the Supporting Information section (http://onlinelibrary.wiley.com/doi/10.1002/acr.25640).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25640.

Address correspondence via email to Mandana Nikpour, PhD, at mandana.nikpour@sydney.edu.au.

Submitted for publication February 5, 2025; accepted in revised form August 14, 2025.

12 countries in the Asia–Pacific region. Patients eligible for enrollment into the APLC cohort are those with newly diagnosed or pre-existing adult patients with SLE (≥18 years old), fulfilling either the 1997 American College of Rheumatology (ACR) Modified Classification Criteria or the 2012 Systemic Lupus International Collaborating Clinics (SLICC) Classification Criteria for SLE.[9,10] Patients are treated under standard care conditions and are seen every three to six months. No specific treatment algorithm was predefined.

Ethics. Human research ethical approval for data collection, storage of the central dataset, analysis, and publication of data collected by the APLC were obtained from the Monash University Human Research Ethics Committee (project 18778). Local ethics approvals have also been obtained at each participating site. Informed consent was obtained from each participant before enrollment.

Data collection. In the APLC cohort, patient demographics and SLE-related history were collected at enrollment. SLE manifestations; SLEDAI-2K score[11]; Physician Global Assessment (PGA; scale 0–3) score[12]; medication use including antimalarials, prednisolone (or equivalent), methotrexate (MTX), leflunomide (LEF), azathioprine (AZA), mycophenolate/mycophenolic acid (MMF/MPA), cyclosporin (CsA), tacrolimus (Tac), cyclophosphamide (CYC), rituximab, and belimumab; and disease flares were collected at each visit. The SLICC/ACR Damage Index (SDI) and the HRQoL 36-Item Short-Form Health Survey (SF36) were completed annually.[13,14] All data were recorded in a standardized electronic case report form.[15]

Data availability. Access to APLC pooled data is subject to the specific guidelines outlined in the APLC Data Access Policy (available on request). The APLC welcomes requests for aggregate (summary) data or to perform analyses of new research questions, and such requests can be submitted to the APLC Steering Committee via the APLC Project Manager.

Definitions. LLDAS was defined as SLEDAI-2K ≤4, with no activity in any major organ, no new disease activity since the previous assessment, PGA ≤1, and prednisolone (or equivalent) dosage ≤7.5 mg/day, with standard maintenance doses of immunosuppressant (IS) and approved biologic agents allowed.[3] The DORIS criteria require a cSLEDAI-2K of 0 (disregarding serology including anti–double-stranded DNA and complements) and PGA <0.5. It allows a maximum prednisolone dosage of 5 mg/day and IS treatment.[8] Disease flare was assessed using the flare definition used in the Safety of Estrogens in Lupus Erythematosus National Assessment (SELENA) trial—the SELENA-SLEDAI Flare Index (SFI).[16] Cumulative organ damage was defined according to SDI. An SDI score >0 was an indication of the presence of organ damage overall, and an increase in SDI of 1 unit or greater was defined as damage accrual.[13] Minimum clinically important difference (MCID) in the mental component summary (MCS) and physical component summary (PCS) scores of the SF36 were defined as a +2.5 increase in MCS and PCS scores.[17]

Statistical analysis. Continuous variables are presented as medians with interquartile ranges (IQRs), and categorical variables as numbers (percentages). Time to first LLDAS or DORIS attainment was measured from the first visit that met the inclusion criteria (study baseline) to the first subsequent visit that met the LLDAS or DORIS criteria. After first reaching LLDAS or DORIS, patients were further followed up, and time to first flare or damage accrual was defined from the first LLDAS or remission point to the first subsequent visit that met the definition for SFI or damage accrual. The duration of the first episode of LLDAS or DORIS was measured from the first visit that met the LLDAS or DORIS criteria to the first subsequent visit that did not. Univariable and multivariable time-varying covariate Cox regression analyses across consecutive visits were used to account for time-varying clinical features to determine factors associated with time to first LLDAS or DORIS attainment and factors associated with time to first flare and damage accrual following first LLDAS or DORIS attainment. Age, gender, disease duration, ethnicity, current smoking status, education level, clinical and serologic features according to SLEDAI-2K, and treatment-related variables, including prednisolone equivalent glucocorticoids dose, antimalarial use, and IS use (categorized as MMF/MPA use and other IS use including MTX, LEF, AZA, CsA, Tac, and CYC), were included in the univariable analysis for determinants of target attainment. Age, gender, disease duration, ethnicity, current smoking status, education level, serologic features, cumulative time in LLDAS or DORIS post-first target attainment until the current visit, time-adjusted mean (TAM) daily prednisolone dose from cohort enrollment up to the current visit, antimalarial use, reduction of prednisolone dose, and reduction of IS dose were included in the univariate analyses for flares and damage accrual after first LLDAS or DORIS attainment. Dose reduction was defined as any decrease in dose from the previous visit in prednisolone equivalent glucocorticoids or oral IS therapy, including MTX, LEF, AZA, MMF/MPA, CsA, and Tac, excluding antimalarials. All the dose reductions were at the physician's discretion. We selected variables for multivariable Cox regression analysis based on the results of univariate analyses and clinical relevance. All analyses were performed with STATA version 17.0 (StataCorp), and a P value of ≤0.05 was considered statistically significant.

RESULTS

Demographic and clinical characteristics at baseline. Among 4,106 patients enrolled in the APLC cohort from March 2013 to December 2020, 1,991 patients newly developed clinical disease activity during follow-up, including

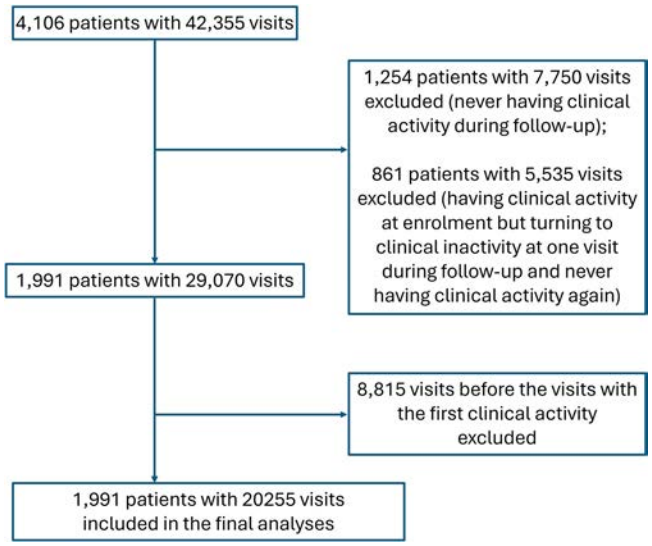

Figure 1. Flow diagram depicting selection of patients and visits for inclusion in analyses.

756 patients who newly developed high disease activity (SLEDAI-2K ≥ 10) on at least one occasion. For the included 1,991 patients, 8,815 visits that occurred before their first clinical activity visit were excluded, leaving 20,255 visits in the analyses (Figure 1). The 1,991 patients had a median (IQR) age at baseline of 41 (32–51) years; 93.2% were female, and 87.7% had Asian ethnicity. The median (IQR) disease duration at baseline was 10 (5–17) years, and the median (IQR) follow-up duration from baseline was 2.5 (0.7–4.5) years. The median (IQR) SLEDAI-2K score at baseline was 6 (4–8), with nephritis (41.2%), mucocutaneous involvement (32.5%), and arthritis (16.0%) as the most common clinical features (Table 1). During follow-up, antimalarial, prednisolone, IS, and biologics were used in 76.1%, 85.8%, 73.5%, and 3.1% of patients, respectively, with a median (IQR) TAM prednisolone dosage of 5 (2.5–9.2) mg/day. Compared to the analyzed cohort overall, the high disease activity subgroup was younger (36 [28–47] years old), had a shorter disease duration at baseline (9.0 [4.0–15.0] years), were more likely to be of Asian ethnicity (88.7%), and had a significantly higher frequency of nephritis (78.7%) and a higher median SLEDAI-2K score (12 [10–14]) at baseline. During follow-up, more patients with high disease activity were taking prednisolone (95.5%), IS (85.2%), and biologics (6.9%), with a higher TAM prednisolone dosage (7.9 [5–12.9] mg/day; Supplementary Table 1).

Target attainment from baseline. During follow-up, 1,412 (70.9%) patients attained LLDAS at least once, with a median (IQR) time to first LLDAS attainment of 0.5 (0.3–1.0) years. 1,107 (55.6%) patients attained DORIS, with a median (IQR) of 0.6 (0.3–1.4) years to first attainment. The duration of the first episodes of LLDAS and DORIS were 0.5 (0.3–1.0) and 0.6 (0.3–1.2) years, respectively (Table 2). Compared to the analyzed cohort overall, the high disease activity subgroup showed a markedly lower proportion of LLDAS (56.1%) and DORIS (39.7%) attainment, with a longer time to attain the targets and a shorter duration in the targets (Supplementary Table 2).

Table 1. Patient characteristics*

Baseline	Cohort (n = 1,991)
Age at baseline visit,[a] median (IQR), y	41 (32–51)
Age at diagnosis, median (IQR), y	29 (21–39)
Disease duration at baseline visit,[a] median (IQR), y	10.0 (5.0–17.0)
Female, n (%)	1,856 (93.2)
Family history of SLE, n (%)	153 of 1,963 (7.8)
Asian ethnicity, n (%)	1,738 of 1,982 (87.7)
Smoker at baseline visit,[a] n (%)	110 of 1,972 (5.6)
Tertiary education, n (%)	908 of 1,819 (49.2)
Clinical features at baseline visit according to SLEDAI-2K,[a] n (%)	
Nephritis	820 (41.2)
Mucocutaneous	647 (32.5)
Arthritis	319 (16.0)
Leukopenia	241 (12.1)
Thrombocytopenia	101 (5.1)
NPSLE	34 (1.7)
Serositis	29 (1.5)
Fever	23 (1.2)
Vasculitis	24 (1.2)
Myositis	8 (0.4)
SLEDAI-2K at baseline visit,[a] median (IQR)	6 (4–8)
PGA at baseline visit,[a] median (IQR)	0.7 (0.3–1.0)
Follow-up	
Follow-up duration from baseline visit,[a] median (IQR), y	2.5 (0.7–4.5)
Total visits per patient including baseline visit, median (IQR)	8 (3–14)
Medication ever used during follow-up	
Antimalarial use,[b] n (%)	1,515 (76.1)
Prednisolone use, n (%)	1,709 (85.8)
TAM daily prednisolone dosage, median (IQR), mg/day	5 (2.5–9.2)
Immunosuppressant use,[c] n (%)	1,464 (73.5)
Biologic use,[d] n (%)	61 (3.1)
Deaths during follow-up, n (%)	45 (2.3)

* IQR, interquartile range; NPSLE, neuropsychiatric systemic lupus erythematosus; PGA, Physician Global Assessment; SLE, systemic lupus erythematosus; SLEDAI-2K, Systemic Lupus Erythematosus Disease Activity Index-2000; TAM, time-adjusted mean.
[a] Baseline visit was defined as the first visit with clinically active disease following a visit without clinical activity.
[b] Antimalarials include hydroxychloroquine and chloroquine.
[c] Immunosuppressants include methotrexate, leflunomide, azathioprine, mycophenolate/mycophenolic acid, cyclosporin, tacrolimus, and cyclophosphamide.
[d] Biologics include rituximab and belimumab.

Main outcomes following first target attainment. After the first attainment of LLDAS, 663 of 1,412 (47.0%) of the patients experienced flare(s) within a median (IQR) follow-up of 2.4 (0.8–4.0) years. The median (IQR) time to first flare was 0.6 (0.3–1.4) years. After the first attainment of DORIS, 526 of 1,107 (47.5%) patients had flare(s) within a median (IQR) of 2.7 (1.0–4.2)

Table 2. Target attainment from baseline and associated outcomes*

LLDAS- and DORIS-associated outcomes	Cohort (n = 1,991)
LLDAS attainment from baseline,[a] n (%)	1,412 of 1,991 (70.9)
Time to first LLDAS attainment, median (IQR), y	0.5 (0.3–1.0)
Duration of first LLDAS, median (IQR), y	0.5 (0.3–1.0)
Flare following first attainment of LLDAS, n (%)	663 of 1,412 (47.0)
Time to first flare, median (IQR), y	0.6 (0.3–1.4)
Damage accrual over 24 months following first attainment of LLDAS, n (%)	68 of 717 (9.5)
Time to first damage accrual, median (IQR), y	1.9 (1.0–3.1)
MCS MCID[b] attainment over 24 months following first attainment of LLDAS, n (%)	207 of 642 (32.2)
PCS MCID[b] attainment over 24 months following first attainment of LLDAS, n (%)	155 of 642 (24.1)
TAM prednisolone dosage over 24 months following first attainment of LLDAS, median (IQR), mg/day	4.5 (1.5–6.0)
Death following first attainment of LLDAS, n (%)	22 of 1,412 (1.6)
DORIS attainment from baseline,[a] n (%)	1,107 of 1,991 (55.6)
Time to first DORIS attainment, median (IQR), y	0.6 (0.3–1.4)
Duration of first DORIS, median (IQR), y	0.6 (0.3–1.2)
Flare following first attainment of DORIS, n (%)	526 of 1,107 (47.5)
Time to first flare, median (IQR), y	0.7 (0.3–1.5)
Damage accrual over 24 months following first attainment of DORIS, n (%)	47 of 592 (7.9)
Time to first damage accrual, median (IQR), y	2.1 (1.2–3.4)
MCS MCID[b] attainment over 24 months following first attainment of DORIS, n (%)	177 of 543 (32.6)
PCS MCID[b] attainment over 24 months following first attainment of DORIS, n (%)	144 of 543 (26.5)
TAM prednisolone dosage over 24 months following first attainment of DORIS, median (IQR), mg/day	3.7 (0.9–5.1)
Death following first attainment of DORIS, n (%)	16 of 1,107 (1.4)

* DORIS, Definitions of Remission in SLE; IQR, interquartile range; LLDAS, Lupus Low Disease Activity State; MCID, minimum clinically important difference; MCS, mental component summary; PCS, physical component summary; SLE, systemic lupus erythematosus; TAM, time-adjusted mean.
[a] Baseline visit was defined as the first visit with clinically active disease following a visit without clinical activity.
[b] MCID in the MCS or PCS scores of the 36-Item Short-Form Health Survey was defined as a +2.5 increase in MCS or PCS score.

years, and the time to first flare was 0.7 (0.3–1.5) years. The incidences of flare during the follow-up period from first reaching LLDAS and DORIS to the end of follow-up (*after* first LLDAS and DORIS attainment) were 45 and 44 per 100 person-years, respectively. However, for the same group of patients, the incidences of flare from baseline to first reaching LLDAS and DORIS (*before* first LLDAS and DORIS attainment) were 200 and 116 per 100 person-years, respectively. After the first LLDAS and remission attainment, 68 of 717 (9.5%) and 47 of 592 (7.9%) patients had damage accrual over two years, respectively, with a median (IQR) of 1.9 (1.0–3.1) and 2.1 (1.2–3.4) years to first damage accrual, respectively (Table 2). With HRQoL, the proportions of patients attaining SF36 PCS and MCS MCID over two years after the first LLDAS or DORIS attainment were similar (24.1% vs 26.5% for PCS and 32.2% vs 32.6% for MCS, respectively). The median (IQR) TAM prednisolone dosages over two years after the first LLDAS and DORIS attainment were 4.5 (1.5–6.0) and 3.7 (0.9–5.1) mg/day, respectively. Concerning mortality rates, 22 of 1,412 (1.6%) patients died following the first LLDAS attainment over a median (IQR) follow-up of 2.4 (1.1–3.8) years; similarly, 16 of 1,107 (1.4%) of patients died over a median (IQR) of 3.1 (1.4–4.2) years following the first DORIS attainment (Table 2). In the high disease activity subgroup, more patients flared (58.0% and 59.3% after the first LLDAS and DORIS attainment, respectively) and accrued damage (11.2% and 9.6% after the first LLDAS and DORIS attainment over 24 months, respectively), compared to the analyzed cohort overall, with a relatively shorter time to first flare and damage accrual. However, the frequency of PCS and MCS MCID attainment was higher than the overall cohort, and the mortality rate was similar to the overall cohort (1.4% and 1.7% after the first LLDAS and DORIS attainment, respectively; Supplementary Table 2).

Determinants of target attainment from baseline. Multivariable time-varying covariate Cox regression analysis in the analyzed overall cohort identified nephritis, low complements, and higher prednisolone dose were independently associated with a longer time to attainment of LLDAS or DORIS. In addition, thrombocytopenia showed a significant association with a longer time to DORIS attainment. In contrast, older age, arthritis, antimalarial use, and MMF/MPA or other IS use were associated with a shorter time to attainment of LLDAS (Table 3).

Determinants of flare from first target attainment. In the analyzed cohort overall, after first achieving LLDAS or DORIS, longer cumulative time in LLDAS or DORIS (hazard ratio [HR] 0.91, 95% confidence interval [CI] 0.89–0.94, $P < 0.0001$ for LLDAS, and HR 0.92, 95% CI 0.89–0.94, $P < 0.0001$ for DORIS) and antimalarial use were each found to be protectively associated with subsequent flare, but prednisolone dose reduction increased the risk of subsequent flare (HR 1.21, 95% CI 1.02–1.45, $P = 0.033$ in the LLDAS model, and HR 1.23, 95% CI 1.00–1.51, $P = 0.047$ in the DORIS model). In addition, IS dose reduction was also associated with a shorter time to flare after LLDAS attainment (HR 1.27, 95% CI 1.02–1.60, $P = 0.036$) but did not reach statistical significance

Table 3. Multivariable models for time to first LLDAS or DORIS attainment from baseline*

Variables[a]	LLDAS		DORIS	
	HR (95% CI)	*P* value	HR (95% CI)	*P* value
Age, per 10 years	1.05 (1.00–1.10)	0.036	1.02 (0.97–1.08)	0.47
Clinical features according to SLEDAI-2K				
Mucocutaneous	0.98 (0.85–1.13)	0.78	0.88 (0.75–1.03)	0.11
Arthritis	1.19 (1.02–1.40)	0.031	1.09 (0.98–1.32)	0.42
Leukopenia	1.03 (0.87–1.23)	0.73	0.83 (0.67–1.02)	0.072
Thrombocytopenia	0.80 (0.58–1.11)	0.178	0.36 (0.21–0.61)	0.0002
Vasculitis	0.48 (0.21–1.12)	0.091	0.21 (0.03–1.50)	0.12
Nephritis	0.58 (0.51–0.67)	<0.0001	0.69 (0.59–0.80)	<0.0001
Serological features according to SLEDAI-2K				
Anti-dsDNA positive	0.94 (0.84–1.05)	0.30	0.90 (0.79–1.03)	0.13
Low complements	0.85 (0.84–1.05)	0.0024	0.78 (0.68–0.89)	0.0003
Treatments				
Prednisolone dosage, mg/day	0.93 (0.91–0.94)	<0.0001	0.89 (0.87–0.90)	<0.0001
Antimalarial use	1.16 (1.02–1.31)	0.024	0.98 (0.86–1.13)	0.81
Immunosuppressant use				
No immunosuppressant	Reference	Reference	Reference	Reference
MMF/MPA	1.16 (1.00–1.34)	0.044	1.16 (0.97–1.37)	0.095
Other immunosuppressants[b]	1.15 (1.00–1.31)	0.046	0.98 (0.84–1.14)	0.75

* Multivariable time-varying covariate Cox regression analysis was performed. CI, confidence interval; DORIS, Definitions of Remission in SLE; dsDNA, double-stranded DNA; HR, hazard ratio; LLDAS, Lupus Low Disease Activity State; MMF/MPA, mycophenolate/mycophenolic acid; SLE, systemic lupus erythematosus; SLEDAI-2K, Systemic Lupus Erythematosus Disease Activity Index-2000.
[a] Variables were at any visit from baseline to the first LLDAS or DORIS attainment. The baseline was defined as the first visit with clinically active disease following a visit without clinical activity.
[b] Other immunosuppressants include methotrexate, leflunomide, azathioprine, cyclosporin, tacrolimus, and cyclophosphamide.

after DORIS attainment (HR 1.23, 95% CI 0.96–1.60, $P = 0.10$; Table 4).

Determinants of damage accrual from first target attainment. In terms of determinants of damage accrual after first LLDAS or DORIS attainment, older age and higher TAM daily prednisolone dose from cohort enrollment were associated with a shorter time to damage accrual, whereas antimalarial use was associated with a longer time to damage accrual in both LLDAS and DORIS models. Longer cumulative time in target since first attainment positively related to a longer time to damage accrual, especially the cumulative time in DORIS, which showed a statistically significant protective effect for subsequent damage accrual (HR 0.97, 95% CI 0.94–1.01, $P = 0.109$ for LLDAS, and HR 0.96, 95% CI 0.93–0.99, $P = 0.036$ for DORIS). In addition, IS dose reduction was associated with an increase in risk of damage accrual, and this association reached statistical significance in the LLDAS model (HR 1.87, 95% CI 1.29–2.73, $P = 0.001$ in the

Table 4. Multivariable models for time to first flare from first LLDAS or DORIS attainment*

Variables[a]	LLDAS		DORIS	
	HR (95% CI)	*P* value	HR (95% CI)	*P* value
Age, per 10 years	0.94 (0.88–1.01)	0.085	0.91 (0.85–0.99)	0.020
Disease duration, per 5 years	1.01 (0.95–1.06)	0.84	1.02 (0.96–1.08)	0.46
Cumulative time in LLDAS/DORIS, per 3 months[b]	0.91 (0.89–0.94)	<0.0001	0.92 (0.89–0.94)	<0.0001
Treatments				
Antimalarial use	0.66 (0.56–0.78)	<0.0001	0.70 (0.58–0.83)	<0.0001
Prednisolone reduction[c]	1.21 (1.02–1.45)	0.033	1.23 (1.00–1.51)	0.047
Immunosuppressant reduction[d]	1.27 (1.02–1.60)	0.036	1.23 (0.96–1.60)	0.10

* Multivariable time-varying covariate Cox regression analysis was performed. CI, confidence interval; DORIS, Definitions of Remission in SLE; HR, hazard ratio; LLDAS, Lupus Low Disease Activity State; SLE, systemic lupus erythematosus.
[a] Variables were at any visit between the first LLDAS or DORIS attainment and the first flare.
[b] Cumulative time in LLDAS or DORIS was defined as being since the first LLDAS or DORIS attainment up to this visit. It was cumulative time in LLDAS for the LLDAS model and cumulative time in DORIS for the DORIS model.
[c] Prednisolone reduction was defined as any decrease in dose from the previous visit in prednisolone equivalent glucocorticoids.
[d] Immunosuppressant reduction was defined as any decrease in dose from the previous visit in any of the oral immunosuppressants, including methotrexate, leflunomide, azathioprine, mycophenolate/mycophenolic acid, cyclosporin, and tacrolimus, excluding antimalarials.

Table 5. Multivariable models for time to first damage accrual from first LLDAS or DORIS attainment*

Variables[a]	LLDAS		DORIS	
	HR (95% CI)	*P* value	HR (95% CI)	*P* value
Age, per 10 years	1.26 (1.11–1.42)	<0.0001	1.38 (1.20–1.57)	<0.0001
Disease duration, per 5 years	1.07 (0.97–1.18)	0.155	1.11 (1.01–1.23)	0.035
Cumulative time in LLDAS/DORIS, per 3 months[b]	0.97 (0.94–1.01)	0.109	0.96 (0.93–0.99)	0.036
Treatments				
Antimalarial use	0.66 (0.49–0.90)	0.008	0.66 (0.47–0.92)	0.016
TAM prednisolone dosage, per 1 mg/day[c]	1.03 (1.00–1.06)	0.029	1.04 (1.01–1.08)	0.022
Immunosuppressants reduction[d]	1.87 (1.29–2.73)	0.001	1.47 (0.92–2.35)	0.108

* Multivariable time-varying covariate Cox regression analysis was performed. CI, confidence interval; DORIS, Definitions of Remission in SLE; HR, hazard ratio; LLDAS, Lupus Low Disease Activity State; SLE, systemic lupus erythematosus; TAM, time-adjusted mean.
[a] Variables were at any visit between the first LLDAS or DORIS attainment and the first damage accrual.
[b] Cumulative time in LLDAS or DORIS was defined as being since the first LLDAS or DORIS attainment up to this visit. It was cumulative time in LLDAS for the LLDAS model and cumulative time in DORIS for the DORIS model.
[c] TAM daily dose of prednisolone equivalent glucocorticoids was from cohort enrollment up to this visit.
[d] Immunosuppressant reduction was defined as any decrease in dose from the previous visit in any of the oral immunosuppressants, including methotrexate, leflunomide, azathioprine, mycophenolate/mycophenolic acid, cyclosporin, and tacrolimus, excluding antimalarials.

LLDAS model, and HR 1.47, 95% CI 0.92–2.35, $P = 0.108$ in the DORIS model; Table 5).

DISCUSSION

Patients enrolled in typical clinical trials of novel therapies in SLE are recruited based on having moderate or severe disease activity.[18,19] In contrast, future studies of treating-to-target approaches in SLE are likely to include a wider subset of patients, namely all those who are not in remission or LLDAS and have any clinical activity. Less is known about the rates, determinants, and consequences of treating-to-target goal attainment in these patients, and filling this knowledge gap may improve the design of such trials. In this study, we have reported the prevalence of attainment of LLDAS and DORIS "targets" in patients with any clinically active disease in a large multinational, multiethnic cohort in the Asia–Pacific region receiving standard-of-care therapy and some clinically important outcomes such as flare, damage accrual, and HRQoL following target attainment. In addition to the wider subset of patients with any clinical activity, we also reported these outcomes in a subgroup with high disease activity.

Cohort studies reported variable rates of attainment of LLDAS and remission. For example, Golder et al previously reported that 78.0% of the patients in the APLC cohort attained LLDAS during a mean follow-up of 2.2 years,[3] whereas 60.9% of the patients attained DORIS.[6] However, the reports were based on the entire APLC population, including a substantial proportion of patients who were in LLDAS or remission throughout the observation period; such patients would not be enrolled in a treating-to-target intervention study. The present study analyzed all patients with any clinically active disease who were not in LLDAS or remission and found relatively lower frequencies of target attainment. This means that almost one-third of patients with active disease were unable to attain a low activity disease state after a median of 2.6 years of usual care, and this number even increased to 44% in patients with high disease activity, underlining the need for a target-driven approach to SLE management. In the present study, the median time to achieve LLDAS and DORIS was 0.5 and 0.7 years, respectively; this is shorter than those reported by Babaoğlu et al and Wilhelm et al from the Hopkins Lupus cohort (1.1 years for LLDAS and 1.8 years for DORIS) despite a lower mean disease activity at baseline in the Hopkins patients.[20,21] However, the duration of the DORIS in the present report was longer than the Hopkins Lupus cohort (7.2 vs 3 months).[21] Different cohort compositions may explain these differences. In the Hopkins cohort, 61.8% of the patients were Caucasian, 30.1% were African American, and less than 10% were of other ethnicities, and patients with African American ethnicity had a lower frequency of LLDAS attainment than the other ethnicities.[20] There was another cohort study from Europe by Pitsigavdaki et al evaluating LLDAS and DORIS attainment rates in a population with moderate to high disease activity, defined as PGA ≥1.5 and/or SLEDAI-2K ≥6. Similar to the Hopkins cohort, they also reported a relatively longer time to target attainment (9 months for LLDAS and 15 months for DORIS) compared to our cohort. Different patient populations with varying ethnicities and disease activities may influence this result. In this European cohort, 96.3% of the patients were of White ethnicity, and the median disease activity was higher than ours.[22]

The subgroup analyses showed that patients with high disease activity were less likely to attain treatment targets than the overall cohort. Moreover, they were less likely to maintain low disease activity or remission status and had a higher chance of flaring, indicating that current management strategies still do not meet the needs of this group of patients. HRQoL is a key outcome, and improvement in HRQoL is one of the primary treatment objectives for patients with SLE. In the current study, over

30% and approximately 25% of the patients achieved an MCID in the mental and physical components of SF36 after reaching the targets, respectively. The high-activity subgroup achieved MCID at a greater frequency than the group overall. These findings suggest that target attainment can benefit HRQoL, especially in patients with high disease activity. Additionally, mental health improved more with target attainment than physical health, and DORIS attainment contributed more to improvement than LLDAS attainment in physical components.

Our results showed that lupus nephritis prolonged the time to LLDAS and remission attainment, whereas thrombocytopenia may impede remission attainment. Several previous studies from large cohorts have also shown that lupus nephritis was negatively associated with low disease activity or remission attainment,[20,23,24] but only one study by Zen et al reported that hematologic manifestations were negatively associated with remission attainment.[24] We also found that low complement was associated with a longer time to both LLDAS and DORIS attainment, consistent with some previous studies.[20,21,23] Regarding treatment effects, our analyses showed a positive association of antimalarial use with LLDAS attainment. In addition, IS use also showed a positive association with LLDAS attainment.

Although a series of studies have reported flare frequencies in patients with SLE,[25–28] we investigated flare frequency *following* LLDAS or remission attainment; this has rarely been reported in large multinational cohorts. We found that the incidence of flare after the target attainment was significantly lower compared to the period before target attainment in the same group of patients (45 vs 200 per 100 person-years after and before LLDAS attainment; 44 vs 116 per 100 person-years after and before DORIS attainment), confirming the protective association of target achievement with flare. However, almost half of the patients still experienced flare(s) following LLDAS and remission attainment within less than one year. In addition, organ damage accrued in almost 10% of the patients after target attainment. Longer cumulative time in LLDAS or DORIS states following the first target attainment was significantly associated with reduced subsequent flare, and longer cumulative time in remission was significantly associated with less subsequent damage accrual. Therefore, management strategies that aim to maintain low disease activity or remission states after attaining these targets could help delay or prevent flares and damage accrual, and longer duration in remission compared to LLDAS is more desirable given its stronger protective association with reduced damage accrual.

Because long-term adverse effects of glucocorticoids are widely recognized, tapering and maintaining the lowest possible dose are recommended in treatment guidelines.[1] Our damage accrual models demonstrated that higher glucocorticoid exposure increased the risk of damage accrual, which further emphasizes the importance of maintaining the lowest possible glucocorticoid dose. However, we found that reducing the dose of glucocorticoids below the 7.5 mg/day prednisolone threshold after LLDAS attainment or the 5 mg/day prednisolone threshold after DORIS attainment increased the risk of flare. The CORTICOLUP randomized trial by Mathian et al also demonstrated that withdrawal of low-dose prednisolone in clinically inactive disease resulted in excess flares compared to the maintenance of prednisolone.[29] These results indicate that more effective and safer glucocorticoid tapering strategies need to be defined and that close monitoring is necessary during the tapering process. Our models showed that IS tapering was associated with a higher risk of flare after LLDAS attainment, suggesting that maintaining IS treatment could potentially prevent flares because glucocorticoids are tapered. A meta-analysis has also reported glucocorticoid-sparing and flare-prevention effects of IS medications.[30] In addition to preventing flare, our models also suggested the effects of IS in preventing damage accrual because IS dose reduction was associated with an increased risk of subsequent damage accrual after LLDAS attainment. However, IS reduction did not significantly increase the risk of flare or damage accrual in the DORIS models, indicating that tapering IS during DORIS may be safer than tapering during LLDAS in terms of risk of flare and damage accrual. Overall, our findings suggest that current treatments are inadequate for flare and damage accrual prevention and highlight the need for interventions and strategies that maintain LLDAS or remission while allowing for the tapering of glucocorticoids.

A key strength of this study is the inclusion of a large number of patients with clinically active disease at baseline. This is important in treat-to-target clinical trial design because these are the patients most suitable to enroll while differing from those enrolled in classic phase 2 or 3 trials. Subgroup analyses for patients with high disease activity provide important information for clinical practice and inform the design of studies and clinical trials that target those with high disease activity.

There are, however, some limitations to this study. First, clinically active disease in the present study was defined according to the clinical domains of SLEDAI-2K. However, SLEDAI-2K omits several clinical manifestations, including hemolytic anemia and cardiac, pulmonary, and gastrointestinal involvement. Secondly, there is increasing use of biologics such as rituximab, belimumab, and anifrolumab in SLE and voclosporin in lupus nephritis. However, we were unable to evaluate the impact of this on target attainment and subsequent outcomes due to their very limited use in this cohort at the time the data were collected. Therefore, the current analysis offers an overview of how targets are achieved and maintained in patients with clinically active disease mainly under the current standard of care, primarily with conventional therapies. Varying factors may contribute to the low usage of biologics, including rituximab, belimumab, and anifrolumab, in this cohort, such as socioeconomic conditions, the availability of these biologics by the time of data collection, and regulatory limitations for access to biologics imposed by local authorities. In addition to limited biologic use, the proportion of conventional IS

use was also relatively low in this cohort (73.5% in the overall cohort and 85.2% in the high disease activity subgroup), which may also have some impact on glucocorticoid use, target attainment, and subsequent outcomes because our models showed that IS use was associated with increased target attainment and IS reduction was associated with increased flare and damage accrual. Thirdly, although the APLC cohort represents a multinational and multiethnic population, non-Asian patients make up less than 20% of the total group. As a result, the generalizability of the findings on a global scale is limited. However, it is important to note that the Asian ethnicity encompasses various races in the APLC cohort, including North-East Asian, South-East Asian, and Southern and Central Asian. The non-Asian ethnicity in the APLC cohort includes Caucasian, Hispanic, Aboriginal and/or Torres Strait Islander, Pacific Islander, Sub-Saharan African, African Caribbean, Egyptian, Latin American, Samoan, and Māori.

In conclusion, LLDAS and DORIS are both achievable goals in patients with clinically active disease. However, a significant proportion of patients receiving usual care, primarily with conventional therapies, do not attain these targets. This could potentially be improved with more target-driven approaches and more effective treatments. Flares and damage accrual occurred frequently following target attainment, highlighting the remaining unmet need in this population. Longer durations of target maintenance and antimalarial use protected patients against flare and damage accrual. Conversely, glucocorticoid dose reduction after target attainment increased the risk of flare, and IS dose reduction increased the risk of both flare and damage. Prospective treat-to-target strategy clinical trials are needed to inform monitoring and treatment adjustment in SLE to maintain states associated with optimal outcomes.

ACKNOWLEDGMENT

We thank all the participants in the APLC for their participation. Open access publishing facilitated by The University of Sydney, as part of the Wiley - The University of Sydney agreement via the Council of Australian University Librarians.

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

ROLE OF THE STUDY SPONSOR

AstraZeneca had no role in the study design or in the collection, analysis, or interpretation of the data or the decision to submit the manuscript for publication. Publication of this article was not contingent upon approval by AstraZeneca.

REFERENCES

1. Fanouriakis A, Kostopoulou M, Andersen J, et al. EULAR recommendations for the management of systemic lupus erythematosus: 2023 update. Ann Rheum Dis 2024;83:15–29.
2. Franklyn K, Lau CS, Navarra SV, et al; Asia-Pacific Lupus Collaboration. Definition and initial validation of a Lupus Low Disease Activity State (LLDAS). Ann Rheum Dis 2016;75:1615–1621.
3. Golder V, Kandane-Rathnayake R, Huq M, et al; Asia-Pacific Lupus Collaboration. Lupus Low Disease Activity State as a treatment endpoint for systemic lupus erythematosus: a prospective validation study. Lancet Rheumatol 2019;1:e95–e102.
4. Hao YJ, Ji LL, Gao D, et al. The flare rates and factors determining flare occurrence in systemic lupus erythematosus patients who achieved low disease activity or remission: results from a prospective cohort study. Lupus Sci Med 2022;9:e000553.
5. Sharma C, Raymond W, Eilertsen G, et al. Association of achieving Lupus Low Disease Activity State fifty percent of the time with both reduced damage accrual and mortality in patients with systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2020;72:447–451.
6. Golder V, Kandane-Rathnayake R, Huq M, et al; Asia-Pacific Lupus Collaboration. Evaluation of remission definitions for systemic lupus erythematosus: a prospective cohort study. Lancet Rheumatol 2019;1:e103–e110.
7. Mucke J, Kuss O, Brinks R, et al. LUPUS-BEST—treat-to-target in systemic lupus erythematosus: study protocol for a three-armed cluster-randomised trial. Lupus Sci Med 2021;8:e000516.
8. van Vollenhoven R, Bertsias G, Doria A, et al. 2021 DORIS Definition of Remission in SLE: final recommendations from an international task force. Lupus Sci Med 2021;8:e000538. Erratum in: *Lupus Sci Med* 2022;9:e000538corr1.
9. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1997;40:1725.
10. Petri M, Orbai AM, Alarcón GS, et al. Derivation and validation of the Systemic Lupus International Collaborating Clinics classification criteria for systemic lupus erythematosus. Arthritis Rheum 2012;64:2677–2686.
11. Ibañez D, Gladman DD, Urowitz MB. Adjusted mean Systemic Lupus Erythematosus Disease Activity Index-2K is a predictor of outcome in SLE. J Rheumatol 2005;32:824–827.
12. Petri M, Hellmann D, Hochberg M. Validity and reliability of lupus activity measures in the routine clinic setting. J Rheumatol 1992;19:53–59.
13. 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;27:373–376.
14. Ware JE Jr, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;30:473–483.
15. Kandane-Rathnayake R, Golder V, Louthrenoo W, et al. Development of the Asia Pacific Lupus Collaboration cohort. Int J Rheum Dis 2019;22:425–433.
16. Petri M, Buyon J, Kim M. Classification and definition of major flares in SLE clinical trials. Lupus 1999;8:685–691.
17. Strand V, Crawford B. Improvement in health-related quality of life in patients with SLE following sustained reductions in anti-dsDNA antibodies. Expert Rev Pharmacoecon Outcomes Res 2005;5:317–326.

18. Morand EF, Furie R, Tanaka Y, et al; TULIP-2 Trial Investigators. Trial of anifrolumab in active systemic lupus erythematosus. N Engl J Med 2020;382:211–221.

19. Petri M, Bruce IN, Dörner T, et al. Baricitinib for systemic lupus erythematosus: a double-blind, randomised, placebo-controlled, phase 3 trial (SLE-BRAVE-II). Lancet 2023;401:1011–1019.

20. Babaoğlu H, Li J, Goldman D, et al. Time to Lupus Low Disease Activity State in the Hopkins Lupus cohort: role of African American ethnicity. Arthritis Care Res (Hoboken) 2020;72:225–232.

21. Wilhelm TR, Magder LS, Petri M. Remission in systemic lupus erythematosus: durable remission is rare. Ann Rheum Dis 2017;76: 547–553.

22. Pitsigavdaki S, Nikoloudaki M, Garantziotis P, et al. Pragmatic targets for moderate/severe SLE and their implications for clinical care and trial design: sustained DORIS or LLDAS for at least 6 months is sufficient while their attainment for at least 24 months ensures high specificity for damage-free progression. Ann Rheum Dis 2024;83: 464–474.

23. Golder V, Kandane-Rathnayake R, Hoi A, et al; Asia-Pacific Lupus Collaboration. Frequency and predictors of the Lupus Low Disease Activity State in a multi-national and multi-ethnic cohort. Arthritis Res Ther 2016;18:260.

24. Zen M, Iaccarino L, Gatto M, et al. Prolonged remission in Caucasian patients with SLE: prevalence and outcomes. Ann Rheum Dis 2015; 74:2117–2122.

25. Petri M, Singh S, Tesfasyone H, et al. Prevalence of flare and influence of demographic and serologic factors on flare risk in systemic lupus erythematosus: a prospective study. J Rheumatol 2009;36:2476–2480.

26. Mirzayan MJ, Schmidt RE, Witte T. Prognostic parameters for flare in systemic lupus erythematosus. Rheumatology (Oxford) 2000;39: 1316–1319.

27. Inês L, Duarte C, Silva RS, et al. Identification of clinical predictors of flare in systemic lupus erythematosus patients: a 24-month prospective cohort study. Rheumatology (Oxford) 2014;53:85–89.

28. Nikpour M, Urowitz MB, Ibañez D, et al. Frequency and determinants of flare and persistently active disease in systemic lupus erythematosus. Arthritis Rheum 2009;61:1152–1158.

29. Mathian A, Pha M, Haroche J, et al. Withdrawal of low-dose prednisone in SLE patients with a clinically quiescent disease for more than 1 year: a randomised clinical trial. Ann Rheum Dis 2020;79:339–346.

30. Pego-Reigosa JM, Cobo-Ibáñez T, Calvo-Alén J, et al. Efficacy and safety of nonbiologic immunosuppressants in the treatment of nonrenal systemic lupus erythematosus: a systematic review. Arthritis Care Res (Hoboken) 2013;65:1775–1785.

Arthritis Care & Research
Vol. 78, No. 3, March 2026, pp 388–397
DOI 10.1002/acr.25628

American College of Rheumatology
Empowering Rheumatology Professionals

Virtual or In-Person: Does It Matter? Comparing Pain, Function, Quality of Life, Self-Efficacy, and Physical Function Outcomes of Virtual, Hybrid, and In-Person Education and Exercise Program Participants

Jill Van Damme,[1] Vanina Dal Bello-Haas,[1] Ayse Kuspinar,[1] Patricia Strachan,[1] Michael Zywiel,[2] James Young,[3] Rhona McGlasson,[4] and Lisa C. Carlesso[1]

Objective. This study aimed to determine if program format (in-person, virtual, or hybrid) results in differences in 3-month outcomes of pain, function, quality of life, self-efficacy, and chair stands in a hip/knee osteoarthritis-management program.

Methods. A secondary analysis of the Good Life with osteoArthritis in Denmark (GLA:D) Canada database was completed. Multiple linear regression was completed for pain and function, analysis of covariance for quality of life and self-efficacy, and negative binomial regression to analyze chair stands. Outcome measures included the 12-item Knee/Hip Injury and Osteoarthritis Outcome Score (pain, quality of life, and physical function subscales), Arthritis Self-Efficacy Scale (self-efficacy), and 30-second chair stand test. Models were adjusted for different covariates.

Results. The analyses included 5,062 individuals with knee and/or hip osteoarthritis who completed the program between January 2019 and March 2024 (76.7% female sex, mean age 67.27 years, mean body mass index 29.51 kg/m^2). When compared with in-person formats, there was no difference in virtual or hybrid formats at 3 months for pain, quality of life, or self-efficacy. When compared with in-person formats, the virtual format resulted in lower function scores (B = –1.71; 95% confidence interval [CI] –2.78 to –0.63) and the hybrid format performed 3% fewer chair stands at 3 months (incidence rate ratio 0.97; 95% CI 0.93–0.99), which is not a clinically important change.

Conclusion. The GLA:D Canada program appears effective in virtual, hybrid, and in-person formats. With the known barriers of strictly in-person formats, these results provide further support for the research and implementation of virtual and hybrid approaches for helping individuals manage osteoarthritis.

INTRODUCTION

An estimated 595 million individuals around the globe are living with symptomatic osteoarthritis (OA),[1] resulting in chronic pain, limited physical function, mental health concerns, reduced quality of life (QoL) and/or increased risk for numerous chronic conditions (hypertension, dyslipidemia, or back pain).[2,3] The prevalence of OA increases with age[1] and commonly affects the hip and knee joints.[1] Current American College of Rheumatology (ACR) guidelines for managing OA and its symptoms strongly recommend first-line treatment including education, exercise, weight management, and pharmaceuticals.[4]

Many programs designed to help individuals self-manage their OA are in-person, introducing a substantial barrier for some populations.[5] Alternatively, there are some virtual OA programs that have demonstrated feasibility and acceptability and improved outcomes (eg, decreased pain and disability, improved function, or higher QoL) for individuals with OA.[6,7] However, many programs focus on education,[7] focus on self-guided exercise modules without feedback on technique or safety,[7,8] or only target individuals before or after total joint replacement (TJR).[6] Additionally, several of the identified virtual programs describe an in-person component,[7] better reflecting a hybrid approach.

This study was completed in partial fulfillment of Jill Van Damme's PhD program, dual-option MSc Physiotherapy and PhD Rehabilitation Science, School of Rehabilitation Science, McMaster University.

[1]Jill Van Damme, PT, PhD, Vanina Dal Bello-Haas, PT, PhD, Ayse Kuspinar, PT, PhD, Patricia Strachan, RN, PhD, Lisa C. Carlesso, PT, PhD: McMaster University, Hamilton, Ontario, Canada; [2]Michael Zywiel, MD: University Health Network, Toronto, Ontario, Canada; [3]James Young, DC, PhD: University Health Network, Toronto, Ontario, Canada, and University of Southern Denmark, Odense, Denmark; [4]Rhona McGlasson, PT: Bone and Joint Canada, Toronto, Ontario, Canada.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25628.

Address correspondence via email to Lisa C. Carlesso, PT, PhD, at carlesl@mcmaster.ca.

Submitted for publication October 28, 2024; accepted in revised form July 31, 2025.

SIGNIFICANCE & INNOVATIONS

- Osteoarthritis-management programs can be effective at improving pain, but individuals cannot always access in-person programs, warranting investigation into virtual and hybrid approaches.
- This study found that virtual and hybrid formats demonstrated equivalency to in-person at 3 months for pain, function, quality of life, self-efficacy, and chair stand scores. Although significant differences were observed in function and chair stand scores for virtual and hybrid, respectively, when compared with in-person, they were not clinically meaningful.
- Given the supporting evidence on virtual assessments for rehabilitation, this study provides additional support for the implementation and continuation of the Good Life with osteoArthritis in Denmark program via virtual and hybrid formats.

OA-management programs can reduce the direct (eg, medications, assistive devices, or rehabilitation) and indirect (eg, transportation costs, time off work, or childcare) health care costs associated with OA.[9] Virtual OA-management programs may be more appealing than in-person formats because they save participants both time and money.[10] However, health care providers have raised concerns over the reliability of virtual assessments and the quality of virtual care in relation to patient outcomes.[11] There has been a significant focus since the onset of the COVID-19 pandemic on validating virtual patient-reported measures[12] and performance-based measures,[13] which have found virtual assessments to be feasible, reliable, and valid.[12,13] Preliminary evidence supports that virtual program outcomes may be equivalent to in-person programs for participant pain and function.[14]

The Good Life with OA in Denmark (GLA:D) Canada[15] program provides a unique opportunity as it offers the same program for hip and knee OA with in-person, virtual, and hybrid formats. The GLA:D program originated in Denmark and was implemented in Canada (GLA:D Canada) in 2016.[15] The GLA:D Canada program consists of two group education and 12 group exercise sessions over 6 weeks (described elsewhere[16,17]) and collects patient-reported and performance-based outcomes.[17] Unlike many virtual and hybrid programs, each education and exercise session of the GLA:D program is therapist-led. The objective of this study was to compare program outcomes across in-person, virtual, and hybrid formats. Specifically, we aimed to assess differences in 3-month patient-reported outcomes of pain, function, QoL, self-efficacy, and chair stands between participants who complete an in-person, virtual, or hybrid GLA:D Canada program. We hypothesized that participants who completed the GLA:D Canada program virtually or hybrid would report outcomes (pain, function, QoL, self-efficacy, and chair stands) equivalent to those in the GLA:D Canada in-person formats.

MATERIALS AND METHODS

This study is a secondary analysis of the GLA:D Canada Database. Ethics approval was obtained from the Hamilton integrated Research Ethics Board (project #13700).

Participants and sample. The GLA:D Canada registry is an ongoing national database of individuals with hip and/or knee OA who have participated in the GLA:D Canada program. Individuals with symptomatic hip or knee OA are eligible to participate in a GLA:D program. The GLA:D program does not require radiographic confirmation of OA, and this information is not currently available within the database. Clinicians determine if an individual's symptoms would be responsive to the GLA:D program.[17] The GLA:D Canada program is offered through a variety of care pathways (private or publicly funded clinics and hospital settings), and there may be a cost associated with participation that is set by individual clinics.[17] Participants choose to contribute their data into the database through informed consent at the clinic level and are responsible for the input of all data via a secure log-in. We requested data for individuals who participated in a GLA:D program between January 2019 and March 2024 as this was inclusive of the time at which the virtual and hybrid formats were introduced. At the time of data transfer there were 20,272 individuals in the database.

Only individuals who had data on the type of GLA:D format completed (in-person, virtual, or hybrid), had a formal diagnosis of OA, had not received an injection (cortisone or hyaluronic acid) 2 weeks before starting the program, and had not had surgical intervention since starting the program were included in the analysis. A flowchart depicting participant selection is provided in Figure 1.

Program format. Participants chose the program format based on what was offered in their geographic regions. In-person format was operationalized as having attended the assessment, education, and exercise sessions within an outpatient clinic. Virtual format was operationalized as having completed the entirety of the GLA:D Canada program (assessment, education, and exercise) virtually from home or another nonclinical setting using audiovisual technology to participate. The hybrid format was operationalized as having participated in a combination of in-person and virtual format for one or more sessions (assessment, education session, or one or more exercise sessions with a clinician).

Primary outcome. The primary outcome was pain at 3 months after the program as measured by the 12-item Knee Injury OA Outcome Score (KOOS-12)/12-item Hip Injury OA Outcome Score (HOOS-12)[18,19] pain subscale. The KOOS-12/HOOS-12 subscales have demonstrated strong internal consistency ($\alpha > 0.70$ for all subscales), responsiveness to change and

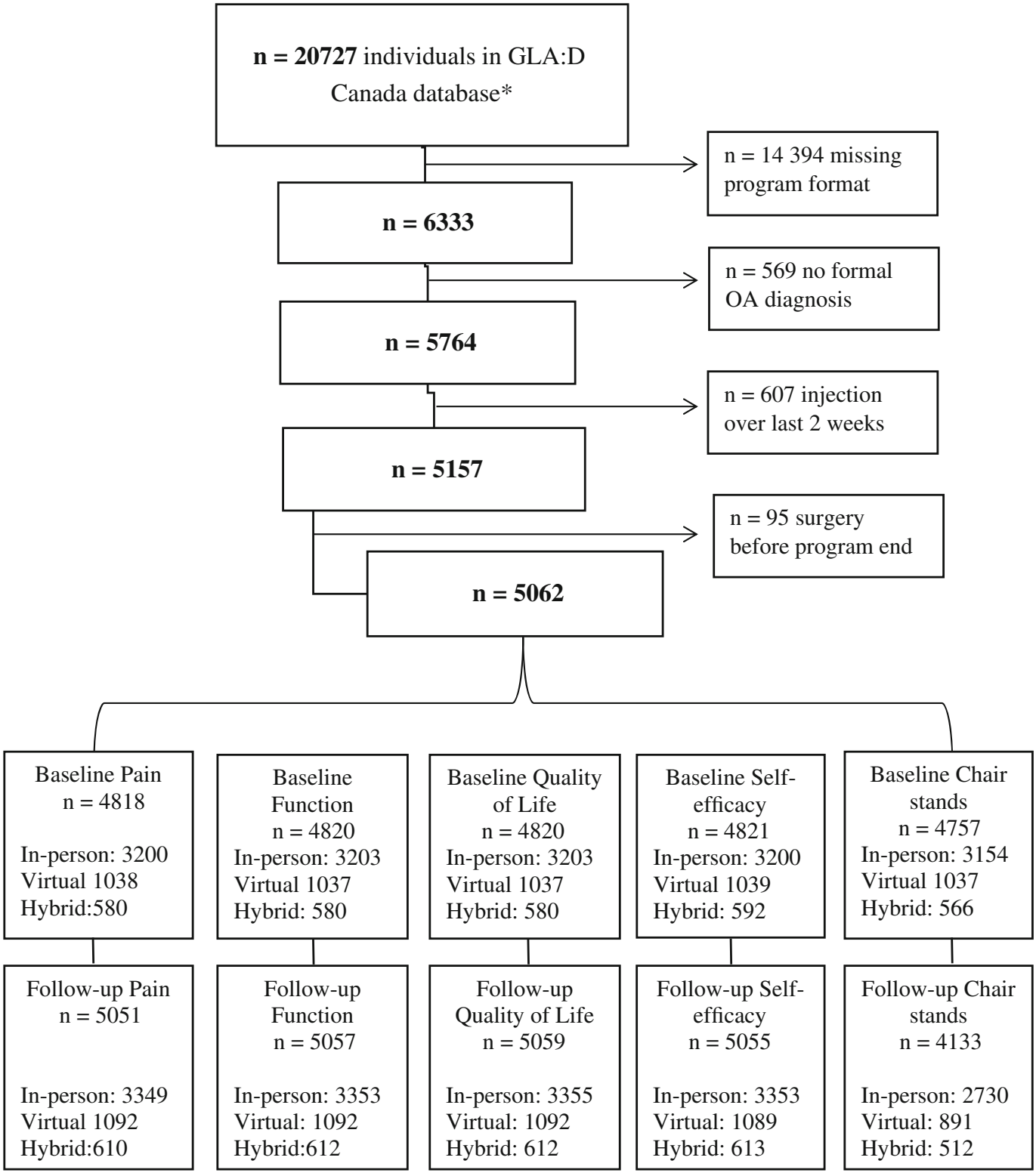

Figure 1. Flow chart of participant selection from the GLA:D Canada database. Complete case analysis was completed for each model with participant numbers as indicated above. Baseline and follow-up data are provided as total and by format (in-person, virtual, or hybrid). Follow-up data are from 3 months after the program. *Between January 2019 and March 2024. GLA:D, Good Life with osteoArthritis in Denmark; OA, osteoarthritis.

validity (known groups, construct),[18,19] with a minimal clinically important difference (MCID) of 14.2 to 16.5 points or 15.6 to 19.2 points for KOOS-12 and HOOS-12, respectively, for individuals before and after a TJR.[20]

Secondary outcomes. Secondary outcomes included QoL and function measured by the KOOS-12/HOOS-12[18,19] respective subscales, self-efficacy measured by the eight-item Arthritis Self-Efficacy Scale (ASES-8),[21] and the 30-s chair stand test (30-s CST),[22] which is a physical performance measure of lower extremity strength.[22] The ASES-8 has also demonstrated strong internal consistency (α = 0.89) and concurrent validity in diverse populations,[21] but the MCID has not yet been reported in the literature. The 30-s CST had demonstrated test-retest reliability and criterion validity and construct validity for community-dwelling older adults.[22] Additionally, the 30-s CST has demonstrated interrater reliability, intrarater reliability for in-person assessments of individuals awaiting a TJR (interclass correlation coefficient [ICC] = 0.97),[23] and test-retest reliability for virtual assessments in community-dwelling veterans, most of whom reported an arthritis diagnosis (ICC = 0.99),[13] with an MCID of two repetitions for individuals with OA.[24]

Covariates. Covariates and confounders were identified by reviewing relevant OA literature, the outcomes of interest, and available database variables. Self-reported demographic data in the GLA:D database include age (years), sex (male, female, or prefer not to answer), body mass index (BMI; calculated based on height in meters and weight in kilograms), comorbidities (cardiovascular disease, hypertension, cholesterol, pulmonary disease, kidney disease, liver disease, hematologic disease, gastrointestinal disease, depression, cancer, rheumatologic disease, back pain, and neurologic disease), highest obtained education (elementary school, high school, trade or college, university, or other), current employment status (full time, part time, not working and on benefits, not working and seeking work, retired, homemaker, or other), marital status (single, married, common law, living with partner, separated, divorced, or widowed), and physical activity levels (number of days physically active for 30 min over the last 7 days). For participants with multiple osteoarthritic joints, the outcome measure was completed based on the most affected joint. The following nominal covariates were dichotomized for the analysis: marital status (partnered vs nonpartnered), employment status (working vs not working) and education level (less than or equal to high school vs postsecondary). All models included age, sex, BMI (kg/m^2), number of comorbidities, and the baseline outcome measure as covariates.

Additional covariates included baseline self-efficacy, education, employment, and marital status for the pain model; baseline pain and baseline self-efficacy for the function and QoL models; baseline pain, physical activity levels, and marital status for self-efficacy; and baseline pain, baseline self-efficacy, and baseline function for chair stands. All outcome measures were completed as per their respective standardized instructions. Chair stands were completed with the supervision of the clinician as per the protocol described by Jones et al.[22]

Sample size. A priori sample size estimates were completed for 3-month pain using G*Power 3.1.9.4[25] with moderate effect size, 0.05 alpha error probability, power of 0.95, and inclusion of nine covariates (age, sex, BMI, number of comorbidities, education, employment status, marital status, self-efficacy, and baseline pain). For three groups (in-person, virtual, or hybrid) a minimum sample size of 172 for each strata for a total of 516 participants was required for one-way analysis of covariance (ANCOVA).

Statistical analysis. Demographic information was analyzed using descriptive statistics for continuous variables (mean and SD) and categorical variables (counts and percentages). A one-way ANCOVA was planned to compare continuous dependent variables (pain, function, QoL, and self-efficacy), and Poisson regression was planned to compare count variables (chair stands).[22] Assumptions were checked for ANCOVA and Poisson regression, including the interaction between format and baseline values, and if violated ($P < 0.05$), a multiple linear regression (MLR) and negative binomial regression were undertaken, respectively.

The amount of missing data was examined for each outcome. For outcomes with <5% missing data at the 3-month follow-up, a complete case analysis was planned. For outcomes with >5% missing data, a comparison between individuals with and without missing data was completed. Multiple imputation for outcome and covariates within the model was completed using logistic regression for categorical variables and linear regression for scale variables using the fully conditional specification (FCS) approach.[26] FCS is appropriate for data that are either monotone or arbitrary missing patterns[26] and preferable when the outcome variable and baseline value of that outcome variable have missing values.[27] Five imputed datasets were generated based on theoretical grounds that five imputations obtain accurate outputs,[28] and the results of the pooled estimates were reported. Analysis was completed using SPSS 29.2 (IBM SPSS Statistics for Windows, IBM Corp).

RESULTS

A total of 5,062 individuals with hip or knee OA were included in analysis, of whom 3,357 completed the program in-person, 1,092 completed the program virtually, and 613 completed a hybrid program. The mean ± SD age for participants was 67 ± 8.2 years, mean ± SD BMI was 29.5 ± 6.0 kg/m^2, and 76.7% were female (n = 3,880); see Table 1 for sociodemographic information. Missing data for pain, function, QoL, and self-efficacy was ≤5%; therefore, a complete case analysis was performed for

Table 1. Sociodemographic information of included participants by format of program*

	Format			
	In-person	Virtual	Hybrid	Total
Totals, n (%)	3,357 (66.3)	1,092 (21.6)	613 (12.1)	5,062 (100)
Age, mean (SD), y	67.7 (8.3)	66.3 (8.2)	66.7 (7.7)	67.3 (8.2)
Sex, n (%)				
Female	2,577 (76.8)	848 (77.7)	455 (74.2)	3,880 (76.7)
Male	776 (23.1)	241 (22.1)	156 (25.4)	1,173 (23.2)
Prefer not to answer	4 (0.1)	2 (0.2)	2 (0.3)	8 (0.2)
BMI, mean (SD), kg/m^2	29.6 (5.9)	29.6 (6.3)	29.1 (6.1)	29.5 (6.0)
Education, n (%)				
High school or less	571 (17)	143 (13.2)	69 (11.3)	783 (15.5)
Postsecondary	2,628 (78.3)	892 (81.7)	513 (83.7)	4,033 (79.7)
Other	6 (0.2)	2 (0.2)	0 (0.0)	8 (0.2)
Marital status, n (%)				
Single	900 (26.8)	279 (25.5)	165 (26.9)	1,344 (26.6)
Partnered	2,297 (68.4)	759 (69.5)	417 (68.0)	3,473 (68.6)
Employment status, n (%)				
Working	731 (21.8)	277 (25.4)	158 (25.8)	1,166 (23.0)
Not working	100 (3.0)	54 (4.9)	23 (3.8)	177 (3.5)
Retired	2,250 (71.5)	669 (64.9)	383 (67.1)	3,302 (69.5)
Homemaker	61 (1.9)	28 (2.7)	6 (1.1)	95 (2.0)
Other	6 (0.2)	3 (0.3)	1 (0.2)	10 (0.2)
Smoking (yes), n (%)	88 (2.7)	41 (3.9)	10 (1.7)	139 (2.9)
Comorbidities, mean (SD)	2,130 (1.678)	1,767 (1.624)	1,865 (1.581)	2,019 (1.662)
Back pain, n (%)	1,394 (43.6)	320 (30.8)	240 (41.5)	1,954 (40.6)
Cancer, n (%)	204 (6.4)	67 (6.5)	39 (6.7)	310 (6.5)
Cardiovascular, n (%)	440 (18.2)	123 (27.9)	48 (13.6)	611 (19.0)
Cholesterol, n (%)	1,116 (35.2)	317 (30.6)	194 (33.6)	1,627 (34.0)
Depression, n (%)	494 (15.5%)	173 (16.7)	91 (15.7)	758 (15.8)
Diabetes, n (%)	345 (10.8)	116 (11.2)	53 (9.2)	514 (10.7)
Gastrointestinal, n (%)	783 (24.6)	155 (14.9)	102 (17.7)	1,040 (21.7)
Hematologic, n (%)	108 (3.4)	35 (3.4)	20 (3.5)	163 (3.4)
Hypertension, n (%)	1,351 (42.4)	378 (36.5)	213 (36.9)	1,942 (40.5)
Kidney, n (%)	102 (3.2)	18 (1.7)	8 (1.4)	128 (2.7)
Liver, n (%)	53 (1.7)	13 (1.3)	8 (1.4)	74 (1.5)
Neurologic, n (%)	120 (5.2)	24 (6.5)	19 (5.6)	163 (5.4)
Pulmonary, n (%)	374 (11.7)	119 (11.5)	67 (11.6)	560 (11.7)
Rheumatologic, n (%)	265 (8.3)	72 (7.0)	41 (7.1)	378 (7.9)
Pain frequency, n (%)				
Never	45 (1.4)	12 (1.2)	13 (2.2)	70 (1.5)
Monthly	83 (2.6)	26 (2.5)	14 (2.4)	123 (2.6)
Weekly	410 (12.8)	123 (11.8)	61 (10.6)	594 (12.4)
Daily	2,045 (64.0)	663 (63.9)	390 (67.5)	3,098 (64.4)
Always	610 (19.1)	214 (20.6)	100 (17.3)	924 (19.2)
Pain in, n (%)				
Right hip	1,162 (70.1)	373 (72.1)	215 (71.0)	1,750 (70.6)
Left hip	1,022 (66.3)	345 (69.0)	184 (64.6)	1,551 (66.7)
Right knee	2,233 (89.1)	753 (91.1)	393 (87.1)	3,379 (89.3)
Left knee	2,106 (88.5)	699 (90.3)	375 (86.6)	3,180 (88.6)
Number of days physically active, mean (SD)[a]	4.4 (2.2)	4.3 (2.2)	4.5 (2.1)	4.4 (2.2)

* BMI, body mass index.
[a] Number of days physically active ≥30 minutes per day.

these models. Table 2 provides baseline and follow-up data, and Table 3 provides a summary of results.

Pain. Because of violations of ANCOVA assumptions (interaction between baseline pain and format, P = 0.02), an MLR model with the interaction term compared 3-month postprogram pain across formats while adjusting for age, sex, BMI, comorbidities, marital status, education level, employment status, baseline self-efficacy, and baseline pain. The interaction term was not statistically significant in the MLR model (P = 0.29, confirmed with visual inspection of graphs) and was not included in the final model. All assumptions for MLR were met. When compared with in-person, there was no statistically significant difference in 3-month pain in the virtual format (B = –0.79, 95% confidence interval [CI] –1.73 to 0.16; P = 0.10) or the hybrid format (B = 0.27, 95% CI –0.93 to 1.46; P = –0.66).

Table 2. Baseline and follow-up data of outcome variables by setting*

Variables	Baseline, mean (SD)	Follow-up, mean (SD)
Pain		
In-person	51.9 (15.7)	58.2 (17.5)
Virtual	51.0 (16.1)	57.2 (17.8)
Hybrid	53.1 (15.1)	59.4 (16.0)
Function		
In-person	58.5 (19.0)	64.7 (20.0)
Virtual	58.9 (19.6)	63.6 (21.0)
Hybrid	60.8 (18.7)	66.5 (18.7)
Quality of life		
In-person	42.2 (17.6)	49.5 (19.1)
Virtual	41.5 (17.0)	48.2 (19.2)
Hybrid	43.2 (16.9)	49.9 (17.7)
Self-efficacy		
In-person	6.2 (1.8)	6.5 (1.9)
Virtual	6.2 (1.8)	6.5 (1.9)
Hybrid	6.3 (1.8)	6.6 (1.9)
30-s chair stand test		
In-person	12.2 (5.3)	16.2 (6.6)
Virtual	11.8 (4.9)	16.4 (6.6)
Hybrid	12.1 (4.9)	15.7 (6.2)

* Pain was measured by the 12-item Hip Injury Osteoarthritis Outcome Score (HOOS-12)/12-item Knee Injury Osteoarthritis Outcome Score (KOOS-12) pain subscale, scored 0 to 100 with higher scores indicating less pain. Function was measured by the HOOS-12/KOOS-12 function subscale, scored 0 to 100 with higher scores indicating better function. Quality of life was measured by the HOOS-12/KOOS-12 quality of life subscale scored 0 to 100 with higher scores indicating better quality of life. Self-efficacy was measured by the eight-item Arthritis Self-Efficacy Scale, scored 1 to 10 with higher scores indicating more self-efficacy. The 30-s chair stand test measured the number of sit-to-stands completed in 30 seconds.

Function. Because of violations of ANCOVA assumptions (interaction between baseline function and format, $P = 0.01$), an MLR model with the interaction term compared 3-month post-program function scores after the program across formats while controlling for age, sex, BMI, comorbidities, baseline self-efficacy, and baseline function. The interaction term was not statistically significant in the MLR model ($P = 0.30$, confirmed with visual inspection of graphs) and was not included in the final model. All assumptions for MLR were met. When compared with in-person, the virtual group reported statistically significant lower function scores compared with in-person (B = –1.71, 95% CI –2.78 to –0.63). There was no statistical difference when comparing the hybrid format with in-person (B = 0.36, 95% CI –0.99 to 1.72).

QoL. An ANCOVA was completed to compare 3-month QoL scores among the in-person, virtual, and hybrid formats. All assumptions were met. After adjustment for age, sex, BMI, comorbidities, baseline pain, baseline self-efficacy, and baseline QoL there was no statistically significant difference in QoL at 3 months when comparing virtual (B = –0.98, 95% CI –1.97 to 0.01; $P = 0.051$) or hybrid (B = –0.31, 95% CI –1.56 to 0.94; $P = 0.63$) with in-person formats (F[2, 4,522] = 1.91, $P = 0.15$, partial $n^2 < 0.001$).

Self-efficacy. An ANCOVA was completed to compare 3-month self-efficacy ratings after the program among the in-person, virtual, and hybrid formats. After adjustment for age, sex, BMI, comorbidities, marital status, physical activity levels, baseline pain, and baseline self-efficacy there was no statistically significant difference in self-efficacy at 3 months when comparing virtual (B = –0.08, 95% CI –0.19 to 0.04; $P = 0.20$) or hybrid (B = 0.03, 95% CI –0.12 to 0.17; $P = 0.72$) with in-person formats (F[2, 4,508] = 1.03, $P = 0.36$, partial $n^2 < 0.001$).

Chair stands. There were 18.4% missing data for 30-s CST at 3-month follow-up, and as such we compared individuals with and without missing data using independent samples *t*-test. We found statistically significantly differences between individuals with and without 30-s CST data at 3 months for age (mean difference = 1.56, 95% CI 0.95–2.16; $P < 0.001$), self-efficacy (mean difference = –0.23, 95% CI –0.36 to –0.09; $P < 0.001$), baseline 30-s CST (mean difference = 0.47, 95% CI 0.05–0.90; $P = 0.03$), and marital status (mean difference = –0.05, 95% CI –0.08 to –0.02; $P = 0.003$). Given the large sample size, we decided to do

Table 3. Summary of results*

Setting	Pain, B (95% CI)[a]	Function, B (95% CI)[a]	Quality of life, B (95% CI)[b]	Self-efficacy, B (95% CI)[b]	Chair stands, IRR (95% CI)[c]
Virtual	−0.79 (−1.73 to 0.16)	−1.71 (−2.78 to −0.63)	−0.98 (−1.97 to 0.01)	−0.08 (−0.19 to 0.04)	1.02 (0.99 to 1.05)
Hybrid	0.27 (−0.93 to 1.46),	0.36 (−0.99 to 1.72)	−0.31 (−1.56 to 0.94)	0.03 (−0.12 to 0.17)	0.97 (0.93 to 0.99)

* Results are in comparison with in-person data. All models were adjusted for age, sex, body mass index, and comorbidities. Additionally, pain was adjusted for marital status, education, employment, baseline self-efficacy, and baseline pain. Function was adjusted for baseline self-efficacy and baseline function. Quality of life was adjusted for baseline pain, baseline self-efficacy, and baseline quality of life. Self-efficacy was adjusted for marital status, physical activity levels, baseline pain, and baseline self-efficacy. Chair stands was adjusted for baseline pain, baseline function, baseline self-efficacy, and baseline chair stands. Function was significant for virtual format, and chair stands was significant for hybrid format. B, regression coefficient; CI, confidence interval; IRR, incidence rate ratio.
[a] Multiple linear regression using complete case analysis.
[b] Analysis of covariance using complete case analysis.
[c] Negative binomial regression using complete case analysis.

both a complete case analysis and multiple imputation and compare the results. Individuals with a value of zero chair stands (ie, those who were not able to stand up from the chair within 30 s) were not considered missing data.

Overdispersion was present in the initial Poisson regression model (Pearson χ^2/df = 1.91) comparing 3-month 30-s CST across in-person, virtual, and hybrid formats while controlling for age, sex, BMI, comorbidities, baseline pain, baseline function, baseline self-efficacy, and baseline chair stands. A negative binomial model was undertaken and demonstrated better fit (Pearson χ^2/df = 1.05). The omnibus test indicated that the overall model was significant ($P < 0.001$). The number of chair stands performed in the hybrid group was 3% lower than the in-person group (incidence rate ratio [IRR] = 0.97, 95% CI 0.93–0.99; $P = 0.03$). No difference was observed between virtual and in-person formats (IRR = 1.02, 95% CI 0.99–1.05; $P = 0.18$). Multiple imputation was completed, and the updated model remained significant ($P < 0.001$). Pooled estimates demonstrated that the number of chair stands performed in the hybrid group was 5% lower than the in-person (IRR = 0.95, 95% CI 0.92–0.97; $P = 0.005$); no difference was observed between virtual and in-person (IRR = 1.02, 95% CI 0.99–1.05; $P = 0.18$).

Post hoc analysis. Given the statistically significant result between hybrid and in-person, we completed a post hoc sensitivity analysis to compare the original data from before and during the COVID-19 restrictions (from 2019 to October 1, 2022, the date restrictions were lifted across Canada) with after the pandemic.[29] We found that the hybrid format remained statistically significant before and during the pandemic, with the hybrid group performing 7% fewer chair stands when compared with in-person (IRR = 0.93, 95% CI 0.88–0.97; $P = 0.002$) but was not statistically significant after the pandemic (IRR = 1.0, 95% CI 0.95–1.05; $P = 0.99$).

DISCUSSION

We used a national database of people with hip and knee OA participating in the GLA:D Canada program to assess whether there are differences in 3-month outcomes based on the format of delivery. In line with our initial hypotheses, we found no differences when comparing in-person with virtual formats at 3 months for pain, QoL, self-efficacy, and 30-s CST; however, function was found to be lower in those attending virtually. Similarly, there were no differences when comparing in-person with hybrid at 3 months for pain, QoL, self-efficacy, and self-reported function, but those in the hybrid group performed slightly fewer chair stands when compared with in-person.[16]

Although this study identified that there were statistical differences between in-person and virtual for function and in-person and hybrid for chair stands, these differences were small and likely do not represent clinically meaningful difference.[30] The virtual group function scores were only 1.71 points less than in-person, well below the MCID of 14.[18,19] Similarly, the hybrid group performed 3% to 5% fewer stands compared with the in-person group, which translates to only roughly half a chair stand and does not surpass the MCID of 2.[13] For a 3% to 5% difference to be clinically meaningful, a participant would need to have completed 40 chair stands. However, this remains an interesting finding and, given how hybrid data are defined in the GLA:D dataset, it is plausible that this finding may be because of a lack of consistency regarding the conducting of baseline and 3-month assessments in the same format (ie, virtual or in-person) and with the same equipment (ie, standardization of chair).

As the data span the COVID-19 pandemic, we hypothesized that the significant finding may be explained by incongruent assessment formats for those attending in the hybrid group. Individuals may have completed baseline and follow-up assessments in two different formats (eg, baseline in-person and follow-up virtually or vice versa), introducing measurement error. Post hoc analysis found that the hybrid format remained statistically significant before and during the pandemic but was not statistically significant after the pandemic. As a result, we suggest interpretation of our result with caution given that the program, including assessments, had to be adapted in response to pandemic restrictions to accommodate virtual participation.[31] Therefore, it is likely that with consistent assessment formats, this difference may not exist. As we are unable to confirm our hypothesis because of limitations in the data, additional research is warranted comparing chair stands among formats moving forward.

Some argue that participants and clinicians should determine what is a meaningful change for each patient circumstance as it relates to their goals, values, and beliefs.[30] As previous research has demonstrated, participants' previous experience with virtual programs influenced their belief and acceptance in these approaches.[32] Offering patients an opportunity to select a program format that best suits them may help enable clinically meaningful outcomes.

This study provides further support for virtual programs for those with hip and knee OA and contributes additional support to a recent meta-analysis that found that remote OA interventions are effective at improving pain, QoL, and physical function for individuals with knee OA.[33] A cohort study also found that both pain and function were equivalent between in-person and virtual programs[14] (the authors describe an in-person introductory session for the virtual program[14] better reflecting a hybrid format), which aligns with our findings.

Alternative approaches to in-person formats such as virtual or hybrid formats can be beneficial to participants and clinicians. Reported benefits for participants may include improved access to specialized programs and clinicians, convenience, scheduling, costs savings, enhanced communication with clinicians, improved program adherence, and increased confidence in managing their OA.[11,34,35] Virtual and hybrid formats may

appeal to clinicians because of the benefits associated with more efficient caseload management, communication with participants, and the flexibility to continue care in different geographic locations.[11,34]

Despite the potential benefits of virtual programs, their acceptance remains somewhat contentious after COVID-19, with some individuals supporting this new direction for rehabilitation and other citing concerns over the process and outcomes.[11,34] Qualitative findings from GLA:D programs in Australia and Canada indicated that although individuals with no previous exposure to virtual rehabilitation were initially skeptical, this resolved with experience of the virtual format.[11,32] With the results of our study suggesting equivalency of virtual or hybrid OA-management program for pain, self-efficacy, and QoL, there is a need to advocate for the inclusion of virtual therapies during clinical training to increase exposure and start familiarizing the next generation of clinicians with virtual care approaches. Integrating virtual formats into clinical training provides the opportunity to develop expertise in these approaches. With exposure to either virtual or hybrid formats, it appears that the acceptance and confidence of clinicians and participants increases.[32,34] However, there are few resources available to understand how to make virtual and hybrid approaches most effective.[31] Many resources to support virtual or hybrid formats focus on *what* is important for maintaining respective college standards (privacy and confidentiality, consent, and patient safety), clearly important aspects of clinical care, but these documents do not delve into the *how* clinicians can be most effective virtually.[31]

Additionally, it is important to acknowledge identified barriers to implementing and accessing virtual care. Virtual programs require access to internet services and telecommunication devices, which may be cost prohibitive,[36] inaccessible in certain geographic regions,[36,37] or simply uncomfortable for individuals to use because of privacy and security concerns.[36] Clinicians have also described concerns over their ability to appropriately monitor patients' form.[37] Qualitative work has identified that clinicians must adjust their approach to virtual program delivery, including the monitoring of form and how feedback is provided.[11]

There are limitations with secondary data analyses, as the available data were not collected for specific research questions.[38] As the measures and data collected were not determined by the research team, there may be other variables that would help provide valuable information related to the outcomes of in-person, virtual, and hybrid formats. There was no information on race or ethnicity, pain catastrophizing, or kinesiophobia, which are linked to OA outcomes.[39,40] Similarly, there is no confirmation of radiographic evidence of OA for participants, including the Kellgren and Lawrence (KL) classification.[41] Although the current standard for diagnosing hip and knee OA is clinical signs and symptoms,[1] evidence supports a strong correlation between radiographic OA and pain when controlling for individual factors.[1] Future work would benefit from the inclusion of KL classification in statistical models. There were also no measures to assess the acceptance of virtual or hybrid formats (eg, Telehealth Acceptance Questionnaire[42]) or adherence to the program, which are valuable for program evaluation. The GLA:D 2021 to 2022 annual report describes >75% attendance,[43] but the program format is not delineated.

This dataset is specific to the Canadian context and results may not be generalizable to other countries implementing the GLA:D program. Participants were also primarily female (76.7%) and well educated (79.7% with postsecondary education), which influences health care use[44] and can impact adherence and engagement with rehabilitative programs. Because of significant missing data at the 12-month follow-up, these data were not included in the analysis and our results are limited to the 3-month follow-up.

This is an observational study completed in the context of the COVID-19 pandemic during which virtual programs were some of the only available options for clinicians and participants to engage in rehabilitation services. As such the value and engagement in virtual and hybrid approaches during this time may be different after COVID-19. Similarly, the conceptualization of hybrid is very broad, requiring only one session each to be completed virtually and in-person. As such the numbers of sessions participated in virtually or in-person may vary from participant to participant, which may influence the results. Future work may benefit from a stricter definition of hybrid formats.

Additionally, there may be differences in participant outcomes related to the specific clinic or clinician offering the program. As previous qualitative work has identified, there are clinician-specific characteristics that participants feel improve virtual or hybrid experiences as well as clinician beliefs and acceptance of virtual or hybrid formats.[11] However, we did not complete the analysis to determine if the clinic site and/or clinician would contribute to the outcomes of the virtual or hybrid formats, as this was considered out of scope for this project. Future work may benefit from examining more specific clinical characteristics that lend themselves to success in virtual and hybrid formats.

There are few other studies that compare the same program delivered in different formats, often comparing "usual care" with a virtual or hybrid program.[7] Additionally, the terms "virtual" or "digital" may be misleading in other studies as many include or require an in-person component[7] that better reflects hybrid approaches to care and makes an evaluation of a truly virtual format compared with in-person more difficult to discern.

Multiple outcomes measures were evaluated in this study including self-reported pain, function, QoL, and self-efficacy in addition to a physical performance measure, all of which are recommended in the evaluation of OA-management

programs.[45] Our study also reflects current recommendations on the evaluation of function by including both self-report and physical performance measures to provide a more complete picture of function.[46]

Lastly, our study compared a clinician-led, synchronous, group-based, OA-management program that included guideline recommended components of education and exercise.[47] All three formats also allow for real-time feedback from clinicians to participants, opportunities to ask questions, and peer support; all components that have been highlighted within chronic disease management theory[48,49] and OA-management literature.[47,50]

Virtual and hybrid formats continue to demonstrate favorable outcomes for OA populations that appear to be equivalent to in-person approaches. Virtual and hybrid approaches may improve access to OA rehabilitation for individuals and improve outcomes while also reducing program costs and mitigating scheduling issues. The results of this study suggest that ongoing support for the offering of virtual and hybrid approaches is warranted while also acknowledging the need for additional research. Based on the results of this study some outcome measures, like chair stands, merit additional research to identify if true differences exist between groups. Future work on virtual and hybrid approaches may benefit from a focus on how to best implement these formats to optimize clinician and participant experience and outcomes, which may help advocate for investment into these alternative approaches to care. For now, clinicians and participants can remain confident in the virtual and hybrid formats of the GLA:D Canada program.

ACKNOWLEDGMENT

Access to the GLA:D Canada database was given in kind by GLA:D Canada.

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 Carlesso 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.

Data availability

The data are owned by the GLA:D Canada team and housed securely at the University Health Network in Toronto, Canada. A data-sharing agreement was required to obtain the data as the dataset is not publicly available.

REFERENCES

1. Tang S, Zhang C, Oo WM, et al. Osteoarthritis. Nat Rev Dis Primers 2025;11(1):10.
2. Swain S, Sarmanova A, Coupland C, et al. Comorbidities in osteoarthritis: a systematic review and meta-analysis of observational studies. Arthritis Care Res (Hoboken) 2020;72(7):991–1000.
3. Aw NMY, Yeo SJ, Wylde V, et al. Impact of pain sensitisation on the quality of life of patients with knee osteoarthritis. RMD Open 2022; 8(1):e001938.
4. Kolasinski SL, Neogi T, Hochberg MC, et al. 2019 American College of Rheumatology/Arthritis Foundation guideline for the management of osteoarthritis of the hand, hip, and knee. Arthritis Rheumatol 2020;72(2):220–233.
5. Wallis JA, Ackerman IN, Brusco NK, et al. Barriers and enablers to uptake of a contemporary guideline-based management program for hip and knee osteoarthritis: a qualitative study. Osteoarthr Cartil Open 2020;2(4):100095.
6. Doiron-Cadrin P, Kairy D, Vendittoli PA, et al. Feasibility and preliminary effects of a tele-prehabilitation program and an in-person prehablitation program compared to usual care for total hip or knee arthroplasty candidates: a pilot randomized controlled trial. Disabil Rehabil 2020;42(7):989–998.
7. Safari R, Jackson J, Sheffield D. Digital self-management interventions for people with osteoarthritis: systematic review with meta-analysis. J Med Internet Res 2020;22(7):e15365.
8. Gohir SA, Eek F, Kelly A, et al. Effectiveness of internet-based exercises aimed at treating knee osteoarthritis: the iBEAT-OA randomized clinical trial. JAMA Netw Open 2021;4(2):e210012.
9. Leifer VP, Katz JN, Losina E. The burden of OA-health services and economics. Osteoarthritis Cartilage 2022;30(1):10–16.
10. Molina-Garcia P, Mora-Traverso M, Prieto-Moreno R, et al. Effectiveness and cost-effectiveness of telerehabilitation for musculoskeletal disorders: a systematic review and meta-analysis. Ann Phys Rehabil Med 2024;67(1):101791.
11. Van Damme J, Dal Bello-Haas V, Strachan P, et al. Client and clinician perspectives about a virtual education and exercise chronic disease management programme for people with hip and knee osteoarthritis. Musculoskeletal Care 2024;22(2):e1881.
12. Bernhardsson S, Larsson A, Bergenheim A, et al. Digital physiotherapy assessment vs conventional face-to-face physiotherapy assessment of patients with musculoskeletal disorders: a systematic review. PLoS One 2023;18(3):e0283013.
13. Ogawa EF, Harris R, Dufour AB, et al. Reliability of virtual physical performance assessments in veterans during the COVID-19 pandemic. Arch Rehabil Res Clin Transl 2021;3(3):100146.
14. Jönsson T, Dell'Isola A, Lohmander LS, et al. Comparison of face-to-face vs digital delivery of an osteoarthritis treatment program for hip or knee osteoarthritis. JAMA Netw Open 2022;5(11):e2240126.
15. Davis AM, Kennedy D, Wong R, et al. Cross-cultural adaptation and implementation of Good Life with osteoArthritis in Denmark (GLA:D™): group education and exercise for hip and knee osteoarthritis is feasible in Canada. Osteoarthritis Cartilage 2018;26(2):211–219.
16. Ageberg E, Nilsdotter A, Kosek E, et al. Effects of neuromuscular training (NEMEX-TJR) on patient-reported outcomes and physical function in severe primary hip or knee osteoarthritis: a controlled before-and-after study. BMC Musculoskelet Disord 2013;14(1):232.
17. Young JJ, Perruccio AV, Veillette CJH, et al. The GLA:D® Canada program for knee and hip osteoarthritis: a comprehensive profile of program participants from 2017 to 2022. PLoS One 2023;18(8): e0289645.
18. Gandek B, Roos EM, Franklin PD, et al. A 12-item short form of the Hip disability and Osteoarthritis Outcome Score (HOOS-12): tests of

reliability, validity and responsiveness. Osteoarthritis Cartilage 2019; 27(5):754–761.

19. Gandek B, Roos EM, Franklin PD, et al. A 12-item short form of the Knee injury and Osteoarthritis Outcome Score (KOOS-12): tests of reliability, validity and responsiveness. Osteoarthritis Cartilage 2019; 27(5):762–770.
20. Soh SE, Harris IA, Cashman K, et al. Minimal clinically important changes in HOOS-12 and KOOS-12 scores following joint replacement. J Bone Joint Surg Am 2022;104(11):980–987.
21. Wilcox S, Schoffman DE, Dowda M, et al. Psychometric properties of the 8-item English Arthritis Self-Efficacy Scale in a diverse sample. Arthritis (Egypt) 2014;2014:385256.
22. Jones CJ, Rikli RE, Beam WCA. A 30-s chair-stand test as a measure of lower body strength in community-residing older adults. Res Q Exerc Sport 1999;70(2):113–119.
23. Gill S, McBurney H. Reliability of performance-based measures in people awaiting joint replacement surgery of the hip or knee. Physiother Res Int 2008;13(3):141–152.
24. Wright AA, Cook CE, Baxter GD, et al. A comparison of 3 methodological approaches to defining major clinically important improvement of 4 performance measures in patients with hip osteoarthritis. J Orthop Sports Phys Ther 2011;41(5):319–327.
25. Faul F, Erdfelder E, Lang AG, et al. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007;39(2):175–191.
26. IBM SPSS Statistics. Overview (MULTIPLE IMPUTATION command). Accessed September 18, 2024. https://www.ibm.com/docs/en/spss-statistics/30.0.0?topic=imputation-analyze-patterns-multiple
27. Jakobsen JC, Gluud C, Wetterslev J, et al. When and how should multiple imputation be used for handling missing data in randomised clinical trials - a practical guide with flowcharts. BMC Med Res Methodol 2017;17(1):162.
28. Carpenter JR, Kenward MG. Chapter 4: Multiple imputation for quantitative data. In: Missing Data in Randomised Controlled Trials - A Practical Guide. Health Technology Assessment Methodology Programme; 2007:75–91.
29. Government of Canada. COVID-19: Canada's response. Accessed September 4, 2024. https://www.canada.ca/en/public-health/services/diseases/2019-novel-coronavirus-infection/canadas-reponse.html
30. Ranganathan P, Pramesh CS, Buyse M. Common pitfalls in statistical analysis: clinical versus statistical significance. Perspect Clin Res 2015;6(3):169–170.
31. Van Damme J, Dal Bello-Haas V, Kuspinar A, et al. Guiding documents for engaging with remote chronic disease management programs as a healthcare provider: a scoping review. Int J Telerehabil 2023;15(2):e6583.
32. Ezzat AM, Bell E, Kemp JL, et al. "Much better than I thought it was going to be": telehealth delivered group-based education and exercise was perceived as acceptable among people with knee osteoarthritis. Osteoarthr Cartil Open 2022;4(3):100271.
33. Yang Y, Li S, Cai Y, et al. Effectiveness of telehealth-based exercise interventions on pain, physical function and quality of life in patients with knee osteoarthritis: a meta-analysis. J Clin Nurs 2023; 32(11-12):2505–2520.
34. Sia LL, Sharma S, Ing JBM, et al. Physiotherapists' perceptions, readiness, enablers, and barriers to use telerehabilitation: a scoping review. J Back Musculoskelet Rehabil 2024;37(6):1441–1454.
35. Molina-Garcia P, Mora-Traverso M, Prieto-Moreno R, et al. Effectiveness and cost-effectiveness of telerehabilitation for musculoskeletal disorders: a systematic review and meta-analysis. Ann Phys Rehabil Med 2024;67(1):101791.
36. Hirko KA, Kerver JM, Ford S, et al. Telehealth in response to the COVID-19 pandemic: implications for rural health disparities. J Am Med Inform Assoc 2020;27(11):1816–1818.
37. Franco JB, Maximino LP, Barretti Secchi LL, et al. What are the barriers to telerehabilitation in the treatment of musculoskeletal diseases? Port J Public Health 2023;42(1):33–42.
38. Cheng HG, Phillips MR. Secondary analysis of existing data: opportunities and implementation. Shanghai Jingshen Yixue 2014;26(6): 371–375.
39. Cijs B, Stekelenburg R, Veenhof C, et al. Prognostic factors on changes in pain, physical functioning and participation in patients with hip- and/or knee OA: a systematic review. Arthritis Care Res (Hoboken) 2025;77(2):228–239.
40. Helminen EE, Arokoski JPA, Selander TA, et al. Multiple psychological factors predict pain and disability among community-dwelling knee osteoarthritis patients: a five-year prospective study. Clin Rehabil 2020;34(3):404–415.
41. Coppola C, Greco M, Munir A, et al. Osteoarthritis: insights into diagnosis, pathophysiology, therapeutic avenues, and the potential of natural extracts. Curr Issues Mol Biol 2024;46(5):4063–4105.
42. Gagnon MP, Orruño E, Asua J, et al. Using a modified technology acceptance model to evaluate healthcare professionals' adoption of a new telemonitoring system. Telemed J e-Health 2012;18(1):54–59.
43. Zywiel M, McGlasson R. GLA:DTM Canada annual report 2021/2022. Bone and Joint Canada. Published June 1, 2023. Accessed March 14, 2024. https://gladcanada.ca/
44. Thompson AE, Anisimowicz Y, Miedema B, et al. The influence of gender and other patient characteristics on health care-seeking behaviour: a QUALICOPC study. BMC Fam Pract 2016;17(1):38.
45. Allen KD, Huffman K, Cleveland RJ, et al. Evaluating osteoarthritis management programs: outcome domain recommendations from the OARSI Joint Effort Initiative. Osteoarthritis Cartilage 2023;31(7): 954–965.
46. Selzer F, Zarra MB, MacFarlane LA, et al. Objective performance tests assess aspects of function not captured by self-report in knee osteoarthritis. Osteoarthr Cartil Open 2022;4(4):100311.
47. Arden NK, Perry TA, Bannuru RR, et al. Non-surgical management of knee osteoarthritis: comparison of ESCEO and OARSI 2019 guidelines. Nat Rev Rheumatol 2021;17(1):59–66.
48. Gee PM, Greenwood DA, Paterniti DA, et al. The eHealth enhanced chronic care model: a theory derivation approach. J Med Internet Res 2015;17(4):e86.
49. Wagner EH, Austin BT, Davis C, et al. Improving chronic illness care: translating evidence into action. Health Aff (Millwood) 2001;20(6): 64–78.
50. Ethgen O, Vanparijs P, Delhalle S, et al. Social support and health-related quality of life in hip and knee osteoarthritis. Qual Life Res 2004;13(2):321–330.

Arthritis Care & Research
Vol. 78, No. 3, March 2026, pp 398–406
DOI 10.1002/acr.25637

American College of Rheumatology
Empowering Rheumatology Professionals

Knee Crepitus and Osteoarthritis Features in Young Adults Following Traumatic Knee Injury

Jamon L. Couch,[1] Brooke E. Patterson,[2] Kay M. Crossley,[2] Ali Guermazi,[3] Matthew G. King,[4] Danilo De Oliveira Silva,[2] Jackie L. Whittaker,[5] Michael A. Girdwood,[2] and Adam G. Culvenor[2]

Objective. This study explored the association between knee crepitus and the presence, and worsening, of structural osteoarthritis features and self-reported outcomes in young adults following traumatic knee injury.

Methods. One year following anterior cruciate ligament reconstruction (ACLR), 112 participants (41 female participants; median age 28 years old) self-reported the presence and/or absence of knee crepitus using an item from the Knee Injury and Osteoarthritis Outcome Score (KOOS). Patellofemoral and tibiofemoral osteoarthritis features (ie, cartilage lesions, osteophytes, and bone marrow lesions) were assessed from magnetic resonance imaging scans at one and five years after ACLR. Self-reported outcomes were assessed with two KOOS subscales (pain and quality of life [QoL]) and the International Knee Documentation Committee subjective evaluation form (ie, self-reported function). Poisson regression evaluated the relationship between self-reported crepitus and the presence/worsening of structural osteoarthritis features. General linear models explored the relationship between crepitus and self-reported outcomes.

Results. Self-reported crepitus was associated with full-thickness patellofemoral cartilage lesions one year after ACLR (prevalence ratio 2.70, 95% confidence interval [CI] 1.41–6.39) but not the risk of worsening structural osteoarthritis features between one and five years after ACLR. Those with crepitus reported worse pain (β = –6.42, 95% CI –10.47 to –2.36), QoL (β = –10.39, 95% CI –18.58 to –2.20), and function (β = –5.49, 95% CI –10.92 to –0.06) one year after ACLR but greater improvement in pain and function between one and five years.

Conclusion. Self-reported knee crepitus was associated with the presence of full-thickness patellofemoral cartilage defects one year after ACLR but was not associated with a greater risk of worsening structural osteoarthritis features up to five years after ACLR. One year after ACLR, those with crepitus reported worse pain, knee-related QoL, and function.

The funders had no role in any part of the study or in any decision about publication.

Supported by Arthritis Australia (Grant in Aid), La Trobe University's Sport, Exercise and Rehabilitation Research Focus Area (Project Grant), the Queensland Orthopaedic Physiotherapy Network (Project Grant), the University of Melbourne (Research Collaboration Grant), and the University of British Columbia's Centre for Hip Health and Mobility (Society for Mobility and Health). Mr Couch's work was supported by an Australian Government Research Training Program scholarship. Dr Patterson's work was supported by an Australian Research Council Early Career Industry Fellowship (grant IE230100135) and was supported by a National Health and Medical Research Council (NHMRC) of Australia postgraduate scholarship at the time of the study (grant GNT1114296). Dr De Oliveira Silva's work was supported by an NHMRC of Australia Investigator Grant (grant GNT2033417). Dr Whittaker's work was supported by a Michael Smith Health Research British Columbia Scholar Award (grant SCH-2020-0403). Dr Girdwood's work was supported by an NHMRC of Australia postgraduate scholarship at the time of the study (grant GNT1190882). Dr Culvenor's work was supported by an NHMRC of Australia Investigator Grant (grant GNT2008523).

[1]Jamon L. Couch, BHlthSc, MPhysioPrac, MExSci(S&C): La Trobe University and Australian International Olympic Committee Research Centre, Melbourne, Victoria, Australia, and Arthritis Research Canada, Vancouver, British Columbia, Canada; [2]Brooke E. Patterson, BPhysio (Hons), PhD, Kay M. Crossley, BAppSc (Physio), PhD, Danilo De Oliveira Silva, PhD, Michael A. Girdwood, BHlthSc, MPhysioPrac, PhD, Adam G. Culvenor, BPhysio (Hons), PhD: La Trobe University and Australian International Olympic Committee Research Centre, Melbourne, Victoria, Australia; [3]Ali Guermazi, MD, PhD, MSc: Boston University School of Medicine, Boston, Massachusetts; [4]Matthew G. King, BPhysio (Hons), PhD: La Trobe University, Melbourne, Victoria, Australia; [5]Jackie L. Whittaker, PhD: Arthritis Research Canada and University of British Columbia, Vancouver, British Columbia, Canada.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25637).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25637.

Address correspondence via email to Adam G. Culvenor, BPhysio (Hons), PhD, at a.culvenor@latrobe.edu.au.

Submitted for publication February 24, 2025; accepted in revised form August 12, 2025.

SIGNIFICANCE & INNOVATIONS

- Knee crepitus has been proposed as an early clinical indicator of osteoarthritis (OA). The absence of studies exploring the relationship between knee crepitus and structural and/or symptomatic OA features in young adults following traumatic knee injury makes the relevance of knee crepitus in this population difficult to ascertain.
- Our study revealed that knee crepitus was cross-sectionally associated with the presence of full-thickness cartilage defects in the patellofemoral compartment and worse pain, knee-related quality of life, and self-reported function at one year after anterior cruciate ligament reconstruction (ACLR).
- Knee crepitus was not associated with greater risk of worsening structural or symptomatic OA features up to five years following ACLR.
- Although the presence of knee crepitus may moderate initial recovery from traumatic knee injury, the absence of an association between knee crepitus and worsening structural and symptomatic OA features brings into question the appropriateness of including crepitus as part of future diagnostic or classification criteria for early-stage post-traumatic OA.

INTRODUCTION

Knee crepitus, the audible crackling or grinding noise during knee joint movement, has been proposed as an early clinical indicator of structural osteoarthritis (OA).[1] A recent meta-analysis demonstrated a cross-sectional association between knee crepitus and several OA-related structural features (eg, osteophytes, bone marrow lesions [BMLs], cartilage defects) on magnetic resonance imaging (MRI); however, a paucity of longitudinal data prevented the evaluation of the prognostic implications of this clinical sign.[2]

The relationship between crepitus and self-reported outcomes related to symptomatic OA is unclear.[3–10] Previous studies have demonstrated a cross-sectional relationship between knee crepitus and higher pain scores,[4,6,9] lower self-reported function, and lower knee-related quality of life (QoL).[9] Several studies have observed nil meaningful association cross-sectionally[5,10] and longitudinally,[7,8] further obscuring the true relationship between knee crepitus and symptomatic outcomes.

A well-established population at elevated risk of early-onset knee OA are young adults with a history of traumatic knee injury, such as anterior cruciate ligament (ACL) rupture.[11] Approximately 50% of people with an ACL rupture will develop symptomatic radiographic OA within a decade of injury,[11,12] an estimated 15 years earlier than their uninjured counterparts.[13] Importantly, the absence of studies exploring the relationship between knee crepitus and OA-related features in these individuals makes the relevance of knee crepitus in this population difficult to ascertain. The pathophysiological processes associated with knee OA development in young adults following traumatic knee injury may differ from that seen in other forms of OA,[14,15] prompting further investigation of this important (and growing) demographic.[16–18]

While established OA represents an advanced and irreversible disease state, early stages of the condition are potentially modifiable, providing a window of opportunity to slow or prevent knee OA progression and reduce disease burden.[11,19,20] Identifying clinical signs that discern individuals at greater risk of structural and symptomatic progression following traumatic knee joint injury has the potential to meaningfully alter disease trajectory through the implementation of targeted secondary prevention strategies. The primary aim of this study was to explore the association between self-reported knee crepitus and the presence, and worsening, of knee OA structural features between one and five years following ACL reconstruction (ACLR). Additionally, we explored the relationship between knee crepitus and self-reported outcomes related to pain, knee-related QoL, and function.

PATIENTS AND METHODS

Study design. This is a secondary analysis of the Knee Osteoarthritis Anterior Cruciate Ligament Longitudinal Assessment (KOALA) study, a longitudinal cohort study evaluating the trajectory of, and risk factors for, symptomatic, structural, and functional outcomes following ACLR. The KOALA study design and methods are described elsewhere.[14,21] Briefly, eligible participants were aged 18 to 50 years at the time of ACLR, consecutively recruited one year following ACLR using a single-bundle hamstring-tendon autograft by two experienced orthopedic surgeons, and re-assessed at five years after ACLR.[21] Baseline exclusion criteria included more than five years between ACL rupture and ACLR, >15 months after ACLR at recruitment, previous injury and/or surgery to the ACLR knee, and subsequent injury and/or surgery to the ACLR knee (between surgery and baseline assessment). Participants with a subsequent injury between baseline and five-year follow-up were not excluded as this reflects a typical ACLR population (up to 30% of young active patients suffer a second ACL rupture in the first few years following surgery).[22,23] At one and five years following ACLR, participants were examined using imaging and self-reported outcome measures. Ethical approval was obtained through the University of Queensland (HREC 2012000567 and 2013001448), University of Melbourne (HREC 1136167), and La Trobe University (HEC 15-100) Human Research Ethics Committees. All participants provided written informed consent.

Self-reported knee crepitus. Self-reported crepitus of the index ACLR knee was assessed using item S2 of the Knee Injury and Osteoarthritis Outcome Score (KOOS) symptoms subscale (KOOS-S2: Do you feel grinding, hear clicking or any other

type of noise when your knee moves?) at one year after ACLR. Participants responded on an ordinal scale (0 = never, 1 = rarely, 2 = sometimes, 3 = often, 4 = always) in relation to symptoms experienced in the past week. The presence of knee crepitus was defined as a response of at least 3 (ie, often or always). Test positivity cutoff was established based on a recent evaluation of the diagnostic performance for this item. The selected threshold provides the greatest diagnostic accuracy (specificity 73%, 95% confidence interval [CI] 61%–83%; positive predictive value 74%, 95% CI 64%–81%; positive likelihood ratio 1.75, 95% CI 1.14–2.70) for the identification of participants with knee crepitus on physical examination.[24]

Structural OA features. Unilateral MRI scans of the index ACLR knee were acquired at one and five years following ACLR using identical sequences and a single 3.0T imaging system (Phillips Achieva) with 16-channel knee coil, as previously described.[14] The three-dimensional proton density-weighted volumetric isotropic turbo spin echo acquisition (VISTA) was acquired at 0.35 mm (repetition time/echo time [TR/TE] 1300 ms/27 ms, field of view [FOV] 150 mm^2, and echo train length 64 ms) and reconstructed in axial and coronal planes. The sagittal short-tau inversion-recovery (STIR) sequence was at 2.5 mm thickness and 1.2 mm slice gap, and an inversion time of 180 ms was applied with TR/TE 3850 ms/30 ms, FOV 160 mm^2, and voxel size 0.45 × 0.50 × 2.5 mm. The axial proton density-weighted turbo spin echo (TSE) sequence was obtained with imaging parameters of TR/TE 3850 ms/34 ms, slice thickness 2.5 mm, slice gap 2.0 mm, corresponding voxel size 0.5 × 0.55 × 2.5 mm, and FOV 140 mm^2.

All MRI scans were evaluated by an experienced musculoskeletal radiologist (AG) with 15 years of experience in semiquantitative MRI analyses (at the time of reading) and previously established interrater and intrarater reliability using the MRI OA Knee Score (MOAKS).[25] The radiologist was blinded to clinical and surgical information but not timepoint of MRI acquisition (ie, images taken at one and five years after ACLR were evaluated together). The MOAKS is a commonly used semiquantitative evaluation method used to assess OA features on MRI with very good interrater reliability (intraclass correlation coefficient [ICC] = 0.61–0.80) and near perfect intrarater reliability (ICC = 0.81–1.0).[25] MOAKS separates the knee into 14 subregions (10 from the tibiofemoral joint and 4 from the patellofemoral joint) to score articular cartilage defects (assessed using VISTA sequence) and BMLs (assessed using STIR and fluid-sensitive TSE sequences). Osteophytes are scored in 12 subregions (6 from the tibiofemoral joint and 6 from the patellofemoral joint). In this study, subregions were grouped to form patellofemoral (medial patella, lateral patella, medial trochlea, and lateral trochlea) and tibiofemoral (medial femur [central and posterior], lateral femur [central and posterior], medial tibia [anterior, central, and posterior], and lateral tibia [anterior, central, and posterior]) compartments.[21]

Articular cartilage defects were graded from 0 to 3 based on size (percentage of surface area relative to subregion; 0 = none, 1 = <33%, 2 = 33%–66%, 3 = >66%). A score of ≥1 was used to define the presence of an articular cartilage defect. This threshold is consistent with previous investigations using MOAKS to explore the association between articular cartilage defects and clinically relevant outcomes.[14,21,25] Previous research has highlighted the need to evaluate the association between knee crepitus and different degrees of cartilage damage (ie, partial thickness vs full thickness), where possible.[26] As such, the extent of full-thickness cartilage defects were also evaluated (percentage of full-thickness loss relative to the subregion; 0 = no full-thickness loss, 1 = <10%, 2 = 10%–75%, 3 = >75%). BMLs (percentage of volume relative to subregion; 0 = none, 1 = <33%, 2 = 33%–66%, 3 = >66%) and osteophytes (how far growth extends from underlying bone; 0 = none, 1 = small, 2 = medium, 3 = large) were also graded based on size. A score of ≥1 was used to define the presence of a BML. Because the definition of a definitive osteophyte has yet to be established,[25] we considered an osteophyte to be present with a score ≥2.[21,27] We defined worsening of OA features in each compartment to include either an increase in the size and/or severity of a previously identified feature (by ≥1 point) or the development of an incident feature between one and five years following ACLR.[27,28]

Self-reported outcomes. Pain and knee-related QoL were evaluated using the KOOS pain and KOOS-QoL subscales, respectively. The International Knee Documentation Committee (IKDC) subjective evaluation form assessed self-reported knee function.[29] The KOOS is comprised of 42 items across five subscales and has demonstrated internal consistency (KOOS-pain Cronbach's alpha [*a*] = 0.84–0.91; KOOS-QoL *a* = 0.64–0.90), test-retest reliability (KOOS-pain ICC = 0.85–0.93; KOOS-QoL ICC = 0.83–0.95), responsiveness to change (KOOS-pain effect size [ES] = 0.84; KOOS-QoL ES = 1.65), and construct validity in the assessment of patient-relevant outcomes following knee injury.[30,31] Items are rated on a five-point Likert scale (0 = no problems, 4 = extreme problems), and a total score for each subscale is summated and converted to a 0 to 100 scale (where 100 = no knee problems). KOOS symptoms were not evaluated because item S2 of this subscale was used to establish the presence of knee crepitus. We excluded the KOOS activities of daily living subscale owing to the ceiling effect exhibited by young active populations following ACLR.[32] Because self-reported function (including participation in sport) was assessed using the IKDC, we did not include the KOOS sport and recreation subscale. The IKDC is comprised of 18 items, which are summated and converted to a 0 to 100 scale (where 100 = no limitations in knee function). The IKDC has demonstrated internal consistency (*a* = 0.77–0.91), test-retest reliability (ICC = 0.90–0.95), responsiveness to change (ES = 1.13), and construct validity.[30,33] The absolute change between one and five years after ACLR was

used for analysis, with a negative value indicating worsening knee problems.

Statistical analyses. Descriptive statistics summarized participant characteristics, crepitus prevalence, self-reported outcomes, and the frequency of OA-related structural feature presence and worsening. The distribution of continuous data was assessed with Shapiro-Wilk tests and visual inspection of histograms and reported accordingly. Subregion-based analyses of the tibiofemoral and patellofemoral compartments were completed using Poisson regression with generalized estimating equations (to account for correlations between subregions in the same knee compartment) to explore the association between self-reported knee crepitus and (1) the presence of articular cartilage defects (+/– full thickness), BMLs, and osteophytes cross-sectionally at one year after ACLR; and (2) risk of worsening of these features longitudinally between one and five years. Prevalence ratios (PR) and risk ratios (RR), with associated 95% CI, were calculated. A PR/RR of >1.0 represents an increased OA feature prevalence/risk of worsening in the presence of knee crepitus. PR/RR with 95% CI not crossing 1.0 were considered statistically significant. General linear models (β, 95% CI) compared pain, knee-related QoL, and self-reported function at one year after ACLR and change in self-reported outcomes between one and five years between those with or without self-reported knee crepitus. Normality of residuals from linear models was evaluated by visually inspecting histograms and Q-Q plots, alongside a formal assessment of normality with Shapiro-Wilk tests. All analyses were adjusted for age, sex, body mass index, and combined injury due to their potential influence on MRI-defined OA structural features and self-reported outcomes.[21,34] This approach aligns with the recognized risk factors for symptomatic and structural OA progression following traumatic knee injury detailed in the recent OPTIKNEE consensus recommendations.[12,20] Comparison of change in self-reported outcomes between one and five years was also adjusted for baseline scores. We defined combined injury as the presence of a cartilage defect (Outerbridge grade ≥2) or meniscal damage warranting meniscectomy at the time of index ACLR, subsequent ACL, meniscal, or collateral ligament injury, or meniscectomy to the index knee between one and five years after ACLR. This definition is consistent with previous longitudinal studies after ACLR.[35,36] Statistical analyses were performed using Stata version 18.0 (StataCorp) with a significance threshold of $P < 0.05$.

RESULTS

Participants. Of the 112 participants who completed self-reported outcomes at one year after ACLR, 111 had baseline MRI data. At five years after ACLR, 81 and 78 completed self-reported outcomes and MRI assessment, respectively (Table 1).

Previous analyses of the KOALA cohort have established that there were no differences (including surgical and baseline MRI findings) between those who did and did not return for five-year follow-up, other than a greater prevalence of medial meniscal damage at baseline in retained participants.[21] The prevalence of structural OA features (at the per-person level) at one year after ACLR and the worsening of features between one-year and five-year follow-up are provided in Supplementary Tables 1 and 2, respectively. Change in KOOS and IKDC scores between one and five years after ACLR have been reported in detail previously.[34] Briefly, significant improvements were noted in KOOS-pain (mean ± SD 2.8 ± 9), KOOS-QoL (mean ± SD 10.0 ± 18.9), and IKDC (mean ± SD 4.7 ± 10.9) between one and five years after ACLR.

Association between knee crepitus and structural OA features. At one year after ACLR, knee crepitus was associated with a higher prevalence of full-thickness cartilage defects of the patellofemoral compartment in both unadjusted (PR 2.70, 95% CI 1.41–6.39) and adjusted (PR 2.77, 95% CI 1.14–6.76) models. There was no association between knee crepitus and the prevalence of other structural OA features in either knee compartment (Table 2). Longitudinally, the presence of knee crepitus at one year after ACLR was not significantly associated with the risk of worsening structural OA features up to five years after ACLR (ie, RR not statistically significant for any OA feature) in either compartment in both unadjusted and adjusted models (Table 3).

Association between knee crepitus and self-reported outcomes. At one year after ACLR, those with knee crepitus demonstrated worse (ie, lower KOOS scores) pain (β –6.42, 95% CI –10.47 to –2.36), knee-related QoL (β –10.39, 95% CI –18.58 to –2.20), and self-reported function (β –5.49, 95% CI –10.92 to –0.06). Results were similar across unadjusted and adjusted models (Table 4). Participants with knee crepitus demonstrated greater improvement (ie, higher KOOS scores) in pain (β 9.88, 95% CI 5.07–14.69) and self-reported function (β 7.40, 95% CI 1.21–13.58) between baseline and follow-up. These relationships were attenuated in adjusted models, longitudinally (Table 4).

DISCUSSION

The current study is the first to explore the relationship between self-reported knee crepitus and OA structural features and self-reported outcomes in young adults at risk of knee OA following traumatic knee injury. Knee crepitus was cross-sectionally associated with the presence of full-thickness cartilage defects in the patellofemoral compartment at one year after ACLR, but not with greater worsening of patellofemoral or tibiofemoral OA structural features up to five years following ACLR (ie, RR not

Table 1. Participant characteristics*

Characteristics	Structural outcome analysis		Self-reported outcome analysis	
	1 year after ACLR	5 years after ACLR	1 year after ACLR	5 years after ACLR
Participants, n	111[a]	78	112	81
Age, median (IQR), years	28 (13)	32 (15)	28 (13)	32 (14)
Female, n (%)	40 (36)	30 (38)	41 (37)	31 (38)
Body mass index, median (IQR), kg/m^2	26 (5)	26 (5)[b]	26 (5)	26 (5)[c]
Combined injury[d], n (%)	51 (46)	38 (49)	52 (46)	40 (49)
Subsequent injury[e], n (%)	NA	11 (14)[f]	NA	10 (12)[g]
Knee crepitus at baseline, n (%)	23 (21)	12 (15)	23 (21)	14 (17)
KOOS subscale, median (IQR)				
Pain	94 (8)	97 (8)[h]	92 (8)	97 (8)
QoL	69 (16)	81 (15)[h]	69 (25)	81 (25)
IKDC score, median (IQR)	82 (15)	91 (15)[h]	86 (16)	91 (15)

* ACLR, anterior cruciate ligament reconstruction; IKDC, Internation Knee Documentation Committee subjective knee form; IQR, interquartile range; KOOS, Knee Injury and Osteoarthritis Outcome Score; NA, not applicable; QoL, quality of life.
[a] Sagittal short-tau inversion (STIR) sequence unavailable for one participant.
[b] n = 72 (ie, those who attended follow-up clinical testing).
[c] n = 75 (ie, those who attended follow-up clinical testing).
[d] Cartilage defect (Outerbridge grade ≥2) or meniscal injury requiring meniscectomy at time of ACLR.
[e] Subsequent ACL, meniscal, or collateral ligament injury or meniscectomy between 1 and 5 years after ACLR.
[f] Partial meniscectomy n = 6; revision ACLR n = 3; collateral ligament injury n = 2.
[g] Partial meniscectomy n = 6; revision ACLR n = 3; collateral ligament injury n = 1.
[h] n = 77.

statistically significant for any OA feature). We found that those with knee crepitus demonstrated worse pain, knee-related QoL, and self-reported function at one year after ACLR; however, they exhibited greater improvements in pain and self-reported function over the long term.

The association between knee crepitus and full-thickness cartilage defects support, in part, the findings of our recent meta-analysis demonstrating a relationship between knee crepitus and cartilage pathology in those with atraumatic knee pain conditions.[2] The relationship between knee crepitus and full-thickness cartilage defects, but absence of association with partial-thickness defects, may be explained by the increased propensity for osseous surface contact during knee joint movement in the presence of full-thickness lesions (a proposed source of knee crepitus).[26] The lack of cross-sectional association between knee crepitus and other OA structural features stand in contrast to our recent meta-analysis, in which knee crepitus was associated with increased odds of MRI-defined osteophytes and BMLs.[2] This disparity is likely owing to the differences in populations investigated, with the participants included in our meta-analysis being predominately middle-aged female patients with atraumatic knee pain conditions (such as primary OA). It has been

Table 2. Association between knee crepitus and prevalence of structural osteoarthritis features on magnetic resonance imaging at one year after anterior cruciate ligament reconstruction (cross-sectional)*

Structural feature	Patellofemoral compartment			Tibiofemoral compartment		
	Prevalence of feature, n/N	Unadjusted, PR (95% CI)	Adjusted[a], PR (95% CI)	Prevalence of feature, n/N	Unadjusted, PR (95% CI)	Adjusted[a], PR (95% CI)
Cartilage defect (any)		1.43 (0.79–2.56)	1.39 (0.84–2.30)		1.14 (0.65–2.01)	1.00 (0.59–1.70)
Crepitus	19/92	–	–	23/230	–	–
No crepitus	51/352	–	–	77/880	–	–
Cartilage defect (full thickness)		**2.70 (1.41–6.39)**	**2.77 (1.14–6.76)**		1.70 (0.73–3.94)	1.39 (0.60–3.22)
Crepitus	12/92	–	–	8/230	–	–
No crepitus	17/352	–	–	18/880	–	–
Bone marrow lesion		1.44 (0.66–3.15)	1.57 (0.72–3.41)		0.99 (0.49–2.01)	0.99 (0.49–2.02)
Crepitus	8/92	–	–	11/230	–	–
No crepitus	21/351[b]	–	–	4/879[b]	–	–
Osteophyte		NA[c]	NA[c]		0.48 (0.06–3.78)	0.57 (0.06–5.06)
Crepitus	0/138	–	–	1/138	–	–
No crepitus	4/528	–	–	8/880	–	–

* Bold values indicate a statistically significant association ($P < 0.05$). Reference category is those without crepitus. CI, confidence interval; NA, not available; PR, prevalence ratio.
[a] Adjusted for age, sex, body mass index, and combined injury.
[b] Sagittal short-tau inversion-recovery (STIR) sequence unavailable for one participant, precluding evaluation of bone marrow lesions.
[c] The association between knee crepitus and structural feature could not be established owing to nonconvergence.
[Correction added on 18 December 2025, after first online publication: The texts 'no crepitus' and 4/879b were added under 'Cartilage defect (any)' and '11/230', respectively.]

Table 3. Association between knee crepitus and worsening of structural osteoarthritis features on magnetic resonance imaging between 1 and 5 years after anterior cruciate ligament reconstruction (longitudinal)*

	Patellofemoral compartment			Tibiofemoral compartment		
Structural feature	Frequency of feature worsening, n/N	Unadjusted, RR (95% CI)	Adjusted[a], RR (95% CI)	Frequency of feature worsening, n/N	Unadjusted, RR (95% CI)	Adjusted[a], RR (95% CI)
Cartilage defect (any)		0.38 (0.14–1.08)	0.41 (0.15–1.14)		1.72 (0.47–6.24)	2.52 (0.80–7.96)
Crepitus	5/48	–	–	5/120	–	–
No crepitus	49/264	–	–	16/660	–	–
Cartilage defect (full thickness)		0.81 (0.31–2.11)	0.90 (0.36–2.22)		3.67 (0.67–20.21)	3.46 (0.70–17.16)
Crepitus	5/48	–	–	4/120	–	–
No crepitus	34/264	–	–	6/660	–	–
Bone marrow lesion		2.29 (0.70–7.47)	2.86 (0.81–10.07)		NA[b]	NA[b]
Crepitus	5/48	–	–	0/120	–	–
No crepitus	12/264	–	–	20/660	–	–
Osteophyte		NA[b]	NA[b]		0.55 (0.12–2.54)	0.58 (0.11–3.07)
Crepitus	0/72	–	–	2/72	–	–
No crepitus	9/396	–	–	20/396	–	–

* Reference category is those without crepitus. CI, confidence interval; NA, not available; RR, risk ratio.
[a] Adjusted for age, sex, body mass index, and combined injury.
[b] The association between knee crepitus and worsening of structural feature could not be established owing to non-convergence.
[Correction added on 18 December 2025, after first online publication: The texts 'Crepitus' and 'No crepitus' were removed next to 'Cartilage defect (full thickness)' and '20/660' was added under '0/120'.]

proposed that the post-traumatic knee OA phenotype is unique to that of primary (non-traumatic) OA. Post-traumatic knee OA likely exhibits distinct pathophysiological processes, possibly owing to the initial osteochondral insult associated with the mechanism of injury and/or altered biomechanics/loading patterns following injury, which may explain these divergent outcomes.[14,15]

Studies aiming to establish early-stage OA classification criteria have proposed that knee crepitus is an important clinical feature in the identification of those on an accelerated trajectory toward an established disease state.[37–39] The absence of an association between knee crepitus and worsening of OA structural features in both the patellofemoral and tibiofemoral compartment brings in to question the appropriateness of including crepitus as part of future diagnostic or classification criteria for early-stage OA following traumatic knee injury. These findings are in contrast to previous research purporting knee crepitus to be the first clinical indication of radiographic patellofemoral OA,[1] yet are consistent with evidence suggesting that it is not predictive of incident radiographic tibiofemoral OA.[1,40] Despite the positive cross-sectional association between patellofemoral joint full-thickness cartilage defects and crepitus, the data suggest that knee crepitus may be associated with a lower risk of cartilage defect worsening (ie, RR < 1.0) over time. Given the wide CIs and lack of statistical significance, this observation should be interpreted with caution. It is possible that with a longer follow-up period and/or larger sample, a clear relationship may be revealed. Long-term follow-up of this cohort (10 years after ACLR) is currently underway.

Knee crepitus prevalence was significantly lower in this cohort (21% at one year after ACLR) compared with the summary crepitus prevalence estimate following ligament injury reported in our recent meta-analysis[2] of 31 studies (35%). Our recent evaluation of the diagnostic performance of the KOOS-S2 suggests that this item may exhibit poor sensitivity and negative predictive value,[24] potentially leading to false-negative test outcomes (ie, failing to detect knee crepitus in those responding ≤2 on the KOOS-S2). Consequently, the prevalence of knee crepitus in this cohort may be, in part, a product of misclassification bias in which some participants were incorrectly classified as not having knee crepitus. It is possible that the relatively low prevalence of knee crepitus (and likelihood of misclassification) may have impacted our ability

Table 4. Association between knee crepitus and self-reported outcomes at 1 year after ACLR and worsening between 1 year and 5 years after ACLR*

	Cross-sectional (1 year after ACLR; n = 112)		Longitudinal (1 to 5 years after ACLR; n = 81)	
Outcomes	Unadjusted	Adjusted[a]	Unadjusted	Adjusted[a]
KOOS-pain, β-coefficient (95% CI)	**−6.42 (−10.47 to −2.36)**	**−6.41 (−10.59 to −2.24)**	**9.88 (5.07 to 14.69)**	**5.15 (0.85 to 9.45)**
KOOS-QoL, β-coefficient (95% CI)	**−10.39 (−18.58 to −2.20)**	**−9.15 (−17.56 to −0.73)**	10.30 (−0.77 to 21.37)	2.15 (−6.64 to 10.93)
IKDC, β-coefficient (95% CI)	**−5.49 (−10.92 to −0.06)**	**−5.19 (−10.21 to −0.17)**	**7.40 (1.21 to 13.58)**	2.93 (−2.54 to 8.40)

* Bold values indicate a statistically significant association ($P < 0.05$). ACLR, anterior cruciate ligament reconstruction; CI, confidence interval; IKDC, International Knee Documentation Committee subjective knee form; KOOS, Knee Injury and Osteoarthritis Outcome Score; QoL, quality of life.
[a] Adjusted for age, sex, body mass index, combined injury, and baseline score.

to identify meaningful associations between this clinical sign and outcomes of interest.

The prevalence of OA structural features on MRI were lower in our study (Supplementary Table 1) relative to comparable cohorts.[41,42] Indeed, the prevalence of some features (ie, osteophytes) in our cohort were similar to that seen in asymptomatic uninjured young adults.[43] This disparity in OA structural feature prevalence is likely due to variable grading methods and definitions of features across studies.[25,42] The relatively low prevalence of OA structural features in this cohort may be implicated in non-significant models, impacting our ability to identify meaningful associations. This is particularly the case for the presence of patellofemoral osteophytes cross-sectionally, and the worsening of tibiofemoral BMLs and patellofemoral osteophytes longitudinally, given we were unable to achieve convergence for these features in our analyses. Worsening of OA features were similar to that reported in previous ACL-injured cohorts evaluated with MOAKS,[44] albeit with more participants exhibiting patellofemoral cartilage worsening in our cohort (Supplementary Table 2). The high proportion of patellofemoral cartilage worsening evidenced in our study is consistent with the rapid loss of patellofemoral cartilage demonstrated in the Knee Anterior Cruciate Ligament Nonsurgical versus Surgical Treatment (KANON) trial.[45]

Our results demonstrate that those with knee crepitus describe worse pain, knee-related QoL, and self-reported function cross-sectionally at one year after ACLR. Previous research evaluating the relationship between knee crepitus and pain in atraumatic knee pain populations appears conflicting, with some studies demonstrating a relationship with higher pain scores[4,6,9] and others observing no association.[5,10] Knee crepitus has also shown to be cross-sectionally associated with lower knee-related QoL in those with knee OA, albeit with small effect sizes.[9] To our knowledge, there exists only one prior study reporting the cross-sectional association between knee crepitus and self-reported outcomes following traumatic knee injury. As part of a tertiary analysis on a subset of young adults after ACLR (mean time after ACLR = 44.4 months), a moderate positive relationship was described between knee crepitus and pain.[46] It is widely reported that patients exhibit significant negative health beliefs surrounding their knee crepitus, often considering their symptoms to be a sign of advanced degenerative changes at the knee joint. These fears lead many to reduce their engagement with physical activity, recreational interests, and even rehabilitation.[47] It is possible that the catastrophizing beliefs, hypervigilance, and fear-avoidant behaviors demonstrated by those with knee crepitus may moderate the initial recovery from traumatic knee injury, as is the case in other musculoskeletal pain conditions.[48]

Those with knee crepitus demonstrated greater improvement in pain and function between one and five years after ACLR. At a cohort level, participants demonstrated improvements in all self-reported outcomes between baseline and follow-up, similar to that seen in other large ACLR cohorts at comparable timepoints.[49] As KOOS-pain, KOOS-QoL, and IKDC scores at five years after ACLR approached normative values seen in the general population,[30,31] it is likely that the greater improvement seen in those with knee crepitus are more so reflective of lower scores at baseline and subsequent reversion to the mean, rather than a product of causation. Additionally, given that effect sizes were attenuated in adjusted models, caution is warranted in the interpretation of these findings. Should our results be confirmed in future investigations, health professionals may use these findings to reassure patients that the presence of knee crepitus following traumatic knee injury may not be a prognostic indicator for long-term knee joint health, while also encouraging engagement with appropriate physical activity to maintain (or improve) articular cartilage integrity and person-centered outcomes.[50] This may help to abate the negative experiences patients report following consultation with health professionals about their knee crepitus, often believing that their concerns had not been adequately addressed.[47]

Several limitations should be considered when interpreting the findings of our study. Firstly, this investigation was a secondary analysis of previously collected cohort data and may be insufficiently powered to detect associations between several outcomes of interest. This is especially relevant given the relatively low prevalence of knee crepitus, specific OA structural features, and the worsening of these features between one and five years after ACLR in this cohort. Despite grouping subregions within the patellofemoral and tibiofemoral compartments to increase statistical power (accounting for the correlation between subregions using generalized estimating equations), the imprecision evidenced in the wide confidence intervals for the worsening of several OA features suggests that results should be interpreted with caution. Conversely, the numerous statistical comparisons in our investigation have the potential to elicit spurious findings and may be implicated in significant models. Due to the exploratory nature of our study, we did not adjust for multiple comparisons. Although this was done to reduce the risk of a Type II error, it should be acknowledged that this may come at the expense of a Type I error. Additionally, despite residuals for our cross-sectional linear models being non-normally distributed, we maintained the use of general linear models, because given the sample size (n = 112), the sampling distribution of the coefficient approximates normality independent of the distribution of the residuals. There were no other major violations to independence, linearity, or homoscedasticity, supporting the appropriateness of linear regression. Finally, participant retention at five years after ACLR was 70% for MRI data and 72% for self-reported outcomes, which may introduce attrition bias. Despite there being no demographic, surgical, or baseline MRI differences between those who did, and did not, return for five-year follow-up (other than a greater prevalence of medial meniscal damage in retained participants), we cannot be certain that our results would not differ were these data included.

Self-reported knee crepitus was associated with the presence of full-thickness patellofemoral cartilage defects one year after ACLR but was not associated with a greater risk of worsening patellofemoral or tibiofemoral OA structural features between one and five years after ACLR. Those with knee crepitus reported worse pain, knee-related QoL, and function cross-sectionally at one year after ACLR; however demonstrated greater improvements in pain and self-reported function longitudinally.

ACKNOWLEDGMENTS

Open access publishing facilitated by La Trobe University, as part of the Wiley - La Trobe University agreement via the Council of Australian University Librarians.

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 Culvenor 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. Schiphof D, van Middelkoop M, de Klerk BM, et al. Crepitus is a first indication of patellofemoral osteoarthritis (and not of tibiofemoral osteoarthritis). Osteoarthritis Cartilage 2014;22(5):631–638.
2. Couch JL, King MG, De Oliveira Silva D, et al. Noisy knees - knee crepitus prevalence and association with structural pathology: a systematic review and meta-analysis. Br J Sports Med 2025;59(2):126–132.
3. Belo JN, Berger MY, Koes BW, et al. Prognostic factors in adults with knee pain in general practice. Arthritis Rheum 2009;61(2):143–151.
4. Clark JA, Spiro A III, Fincke G, et al. Symptom severity of osteoarthritis of the knee: a patient-based measure developed in the veterans health study. J Gerontol A Biol Sci Med Sci 1998;53(5):M351–M360.
5. de Oliveira Silva D, Pazzinatto MF, Priore LBD, et al. Knee crepitus is prevalent in women with patellofemoral pain, but is not related with function, physical activity and pain. Phys Ther Sport 2018;33:7–11.
6. Lo GH, Strayhorn MT, Driban JB, et al. Subjective crepitus as a risk factor for incident symptomatic knee osteoarthritis: data from the osteoarthritis initiative. Arthritis Care Res (Hoboken) 2018;70(1): 53–60.
7. Natri A, Kannus P, Järvinen M. Which factors predict the long-term outcome in chronic patellofemoral pain syndrome? A 7-yr prospective follow-up study. Med Sci Sports Exerc 1998;30(11):1572–1577.
8. Pazzinatto MF, de Oliveira Silva D, Azevedo FM, et al. Knee crepitus is not associated with the occurrence of total knee replacement in knee osteoarthritis - a longitudinal study with data from the Osteoarthritis Initiative. Braz J Phys Ther 2019;23(4):329–336.
9. Pazzinatto MF, de Oliveira Silva D, Faria NC, et al. What are the clinical implications of knee crepitus to individuals with knee osteoarthritis? An observational study with data from the Osteoarthritis Initiative. Braz J Phys Ther 2019;23(6):491–496.
10. Waiteman MC, de Oliveira Silva D, Azevedo FM, et al. Women with patellofemoral pain and knee crepitus have reduced knee flexion angle during stair ascent. Phys Ther Sport 2021;48:60–66.
11. Chu CR, Williams AA, Coyle CH, et al. Early diagnosis to enable early treatment of pre-osteoarthritis. Arthritis Res Ther 2012;14(3):212.
12. Whittaker JL, Losciale JM, Juhl CB, et al. Risk factors for knee osteoarthritis after traumatic knee injury: a systematic review and meta-analysis of randomised controlled trials and cohort studies for the OPTIKNEE Consensus. Br J Sports Med 2022;56(24):1406–1421.
13. Lohmander LS, Englund PM, Dahl LL, et al. The long-term consequence of anterior cruciate ligament and meniscus injuries: osteoarthritis. Am J Sports Med 2007;35(10):1756–1769.
14. Culvenor AG, Collins NJ, Guermazi A, et al. Early patellofemoral osteoarthritis features one year after anterior cruciate ligament reconstruction: symptoms and quality of life at three years. Arthritis Care Res (Hoboken) 2016;68(6):784–792.
15. Eckstein F, Wirth W, Lohmander LS, et al. Five-year followup of knee joint cartilage thickness changes after acute rupture of the anterior cruciate ligament. Arthritis Rheumatol 2015;67(1):152–161.
16. Moses B, Orchard J, Orchard J. Systematic review: annual incidence of ACL injury and surgery in various populations. Res Sports Med 2012;20(3-4):157–179.
17. Zbrojkiewicz D, Vertullo C, Grayson JE. Increasing rates of anterior cruciate ligament reconstruction in young Australians, 2000-2015. Med J Aust 2018;208(8):354–358.
18. Campbell RJ, An V, Molnar R, et al. Rising rates of anterior cruciate ligament reconstruction in Australian adults: an analysis of Australian Medicare benefits schedule database. Knee Surg Sports Traumatol Arthrosc 2024;33(4):1–8.
19. Pollard TCB, Gwilym SE, Carr AJ. The assessment of early osteoarthritis. J Bone Joint Surg Br 2008;90(4):411–421.
20. Whittaker JL, Culvenor AG, Juhl CB, et al. OPTIKNEE 2022: consensus recommendations to optimise knee health after traumatic knee injury to prevent osteoarthritis. Br J Sports Med 2022;56(24):1393–1405.
21. Patterson BE, Culvenor AG, Barton CJ, et al. Worsening knee osteoarthritis features on magnetic resonance imaging 1 to 5 years after anterior cruciate ligament reconstruction. Am J Sports Med 2018; 46(12):2873–2883.
22. Paterno MV, Rauh MJ, Schmitt LC, et al. Incidence of second ACL injuries 2 years after primary ACL reconstruction and return to sport. Am J Sports Med 2014;42(7):1567–1573.
23. Webster KE, Feller JA, Leigh WB, et al. Younger patients are at increased risk for graft rupture and contralateral injury after anterior cruciate ligament reconstruction. Am J Sports Med 2014;42(3): 641–647.
24. Couch JL, King MG, De Oliveira Silva D, et al. Diagnostic performance of self-reported knee crepitus using a knee injury and osteoarthritis outcome score item. Osteoarthritis Cartilage 2024;33(6):764–769.
25. Hunter DJ, Guermazi A, Lo GH, et al. Evolution of semi-quantitative whole joint assessment of knee OA: MOAKS (MRI Osteoarthritis Knee Score). Osteoarthritis Cartilage 2011;19(8):990–1002.
26. Crema MD, Guermazi A, Sayre EC, et al. The association of magnetic resonance imaging (MRI)-detected structural pathology of the knee with crepitus in a population-based cohort with knee pain: the MoDEKO study. Osteoarthritis Cartilage 2011;19(12):1429–1432.
27. Runhaar J, Schiphof D, van Meer B, et al. How to define subregional osteoarthritis progression using semi-quantitative MRI osteoarthritis knee score (MOAKS). Osteoarthritis Cartilage 2014;22(10):1533–1536.
28. van Meer BL, Oei EH, Meuffels DE, et al. Degenerative changes in the knee 2 years after anterior cruciate ligament rupture and related risk

factors: a prospective observational follow-up study. Am J Sports Med 2016;44(6):1524–1533.

29. Irrgang JJ, Ho H, Harner CD, et al. Use of the International Knee Documentation Committee guidelines to assess outcome following anterior cruciate ligament reconstruction. Knee Surg Sports Traumatol Arthrosc 1998;6(2):107–114.
30. Collins NJ, Misra D, Felson DT, et al. Measures of knee function: International Knee Documentation Committee (IKDC) Subjective Knee Evaluation Form, Knee Injury and Osteoarthritis Outcome Score (KOOS), Knee Injury and Osteoarthritis Outcome Score Physical Function Short Form (KOOS-PS), Knee Outcome Survey Activities of Daily Living Scale (KOS-ADL), Lysholm Knee Scoring Scale, Oxford Knee Score (OKS), Western Ontario and Mcmaster Universities Osteoarthritis Index (WOMAC), Activity Rating Scale (ARS), and Tegner Activity Score (TAS). Arthritis Care Res (Hoboken) 2011;63 Suppl 11(0 11):S208–S228.
31. Collins NJ, Prinsen CAC, Christensen R, et al. Knee Injury and Osteoarthritis Outcome Score (KOOS): systematic review and meta-analysis of measurement properties. Osteoarthritis Cartilage 2016; 24(8):1317–1329.
32. Frobell RB, Roos HP, Roos EM, et al. Treatment for acute anterior cruciate ligament tear: five year outcome of randomised trial. BMJ 2013;346:f232.
33. Irrgang JJ, Anderson AF, Boland AL, et al; International Knee Documentation Committee. Responsiveness of the international knee documentation committee subjective knee form. Am J Sports Med 2006; 34(10):1567–1573.
34. Patterson BE, Culvenor AG, Barton CJ, et al. Patient-reported outcomes one to five years after anterior cruciate ligament reconstruction: the effect of combined injury and associations with osteoarthritis features defined on magnetic resonance imaging. Arthritis Care Res (Hoboken) 2020;72(3):412–422.
35. Øiestad BE, Holm I, Aune AK, et al. Knee function and prevalence of knee osteoarthritis after anterior cruciate ligament reconstruction: a prospective study with 10 to 15 years of follow-up. Am J Sports Med 2010;38(11):2201–2210.
36. Risberg MA, Oiestad BE, Gunderson R, et al. Changes in knee osteoarthritis, symptoms, and function after anterior cruciate ligament reconstruction: a 20-year prospective follow-up study. Am J Sports Med 2016;44(5):1215–1224.
37. Luyten FP, Bierma-Zeinstra S, Dell'Accio F, et al. Toward classification criteria for early osteoarthritis of the knee. Semin Arthritis Rheum 2018;47(4):457–463.
38. Mahmoudian A, Lohmander LS, Jafari H, et al. Towards classification criteria for early-stage knee osteoarthritis: a population-based study to enrich for progressors. Semin Arthritis Rheum 2021;51(1): 285–291.
39. Runhaar J, Kloppenburg M, Boers M, et al; the CREDO expert group. Towards developing diagnostic criteria for early knee osteoarthritis: data from the CHECK study. Rheumatology (Oxford) 2021;60(5): 2448–2455.
40. Ho-Pham LT, Lai TQ, Mai LD, et al. Prevalence of radiographic osteoarthritis of the knee and its relationship to self-reported pain. PLoS One 2014;9(4):e94563.
41. Frobell RB, Le Graverand MP, Buck R, et al. The acutely ACL injured knee assessed by MRI: changes in joint fluid, bone marrow lesions, and cartilage during the first year. Osteoarthritis Cartilage 2009; 17(2):161–167.
42. Potter HG, Jain SK, Ma Y, et al. Cartilage injury after acute, isolated anterior cruciate ligament tear: immediate and longitudinal effect with clinical/MRI follow-up. Am J Sports Med 2012;40(2):276–285.
43. Culvenor AG, Øiestad BE, Hart HF, et al. Prevalence of knee osteoarthritis features on magnetic resonance imaging in asymptomatic uninjured adults: a systematic review and meta-analysis. Br J Sports Med 2019;53(20):1268–1278.
44. van Meer BL, Oei EHG, Meuffels DE, et al. Degenerative changes in the knee 2 years after anterior cruciate ligament rupture and related risk factors: a prospective observational follow-up study. Am J Sports Med 2016;44(6):1524–1533.
45. Culvenor AG, Eckstein F, Wirth W, et al. Loss of patellofemoral cartilage thickness over 5 years following ACL injury depends on the initial treatment strategy: results from the KANON trial. Br J Sports Med 2019;53(18):1168–1173.
46. Goradia VK, Grana WA, Pearson SE. Factors associated with decreased muscle strength after anterior cruciate ligament reconstruction with hamstring tendon grafts. Arthroscopy 2006;22(1):80.
47. Robertson CJ, Hurley M, Jones F. People's beliefs about the meaning of crepitus in patellofemoral pain and the impact of these beliefs on their behaviour: a qualitative study. Musculoskelet Sci Pract 2017; 28:59–64.
48. Westman AE, Boersma K, Leppert J, et al. Fear-avoidance beliefs, catastrophizing, and distress: a longitudinal subgroup analysis on patients with musculoskeletal pain. Clin J Pain 2011;27(7):567–577.
49. Cox CL, Huston LJ, Dunn WR, et al. Are articular cartilage lesions and meniscus tears predictive of IKDC, KOOS, and Marx activity level outcomes after anterior cruciate ligament reconstruction? A 6-year multicenter cohort study. Am J Sports Med 2014;42(5):1058–1067.
50. Roos EM, Dahlberg L. Positive effects of moderate exercise on glycosaminoglycan content in knee cartilage: a four-month, randomized, controlled trial in patients at risk of osteoarthritis. Arthritis Rheum 2005;52(11):3507–3514.

Arthritis Care & Research
Vol. 78, No. 3, March 2026, pp 407–416
DOI 10.1002/acr.25621

American College of Rheumatology
Empowering Rheumatology Professionals

BRIEF REPORT

Serum Uric Acid Levels in Older Adults: Associations With Clinical Outcomes and Implications for Reference Intervals in Those Aged 70 Years and Over

Amanda J. Rickard,[1] Cammie Tran,[1] Hans G. Schneider,[2] Flavia M. Cicuttini,[1] Anita E. Wluka,[1] Ego Seeman,[3] Johannes T. Neumann,[4] Md Nazmul Karim,[1] Zhen Zhou,[1] Sultana Monira Hussain,[5] David P. Q. Clark,[1] Daniel Clayton-Chubb,[2] Andrew M. Tonkin,[1] Lawrence J. Beilin,[6] Robyn L. Woods,[1] and John J. McNeil[1]

Objective. Reports have linked both high and low serum uric acid (SUA) levels to adverse health outcomes. This study aimed to establish a reference interval for SUA in older adults and assessed its association with clinically relevant outcomes in relatively healthy, community-dwelling individuals aged ≥70 years old.

Methods. The study used data from the ASPirin in Reducing Events in the Elderly (ASPREE) trial. In Australia, 11,878 ASPREE participants had baseline SUA measurements (median age 74 years old). The study sample (n = 11,446; 55% women) comprised individuals with baseline SUA measurements, excluding those on urate-lowering medication. The reference sample (n = 10,501; 55% women) was established after further exclusion of participants with impaired renal function, defined as an estimated glomerular filtration rate <45 mL/min/1.73m^2. Reference intervals (2.5th and 97.5th percentile) were stratified by sex, and Cox proportional hazard models assessed associations between SUA levels and relevant clinical outcomes.

Results. SUA reference intervals were 0.24 to 0.54 mmol/L for men and 0.19 to 0.48 mmol/L for women. After adjusting for potential confounders, no association was observed between SUA levels and all-cause mortality, disability-free survival, cardiovascular disease, major adverse cardiovascular events, cancer incidence and mortality, or dementia in either the study or reference samples. In women, however, low SUA levels were associated with an increased risk of fractures (hazard ratio 1.23; 95% confidence interval 1.04–1.46).

Conclusion. Although previous reports have linked abnormal SUA levels to adverse health outcomes, our findings show no associations within the reference range, except for an increased fracture risk among women with low SUA levels.

INTRODUCTION

Serum uric acid (SUA) is a clinical marker measured primarily for the diagnosis of gout and renal stones. An increasing number of epidemiologic studies, however, have reported associations between both high and low SUA levels and an increased risk of chronic diseases.[1] Elevated SUA levels have been linked to the onset and progression of hypertension, cardiovascular disease (CVD), chronic kidney disease (CKD), cognitive decline, and higher all-cause mortality.[2] Conversely, low SUA levels have been

The ASPirin in Reducing Events in the Elderly (ASPREE) clinical trial was supported by the National Institute on Aging and the National Cancer Institute, NIH (U01-AG-029824 and U19-AG-062682), the National Health and Medical Research Council (NHMRC) (334047 and 1127060), Monash University, and Alfred Health. Urate testing was supported by a reagent grant from Abbott Diagnostics Australia. The ASPREE Biobank was supported by research grants from the Australian Government's Commonwealth Scientific and Industrial Research Organisation (Preventative Health Flagship 2009) and the National Cancer Institute/National Institutes of Health (5U01-AG-029824-02). Prof McNeil's work is supported by an NHMRC leadership fellow (IG1173690).

[1]Amanda J. Rickard, PhD, Cammie Tran, MPH, Flavia M. Cicuttini, PhD, Anita E. Wluka, PhD, Md Nazmul Karim, PhD, Zhen Zhou, PhD, David P. Q. Clark, MPH, Andrew M. Tonkin, MD, Robyn L. Woods, PhD, John J. McNeil, PhD: Monash University, Melbourne, Australia; [2]Hans G. Schneider, MD, Daniel Clayton-Chubb, MD: Monash University and The Alfred Hospital, Melbourne, Australia; [3]Ego Seeman, MD: University of Melbourne, Melbourne, Australia; [4]Johannes T. Neumann, PhD: Monash University, Melbourne, Australia, and University Heart & Vascular Centre, Hamburg, Germany; [5]Sultana Monira Hussain, PhD: Monash University and University of Melbourne, Melbourne, Australia; [6]Lawrence J. Beilin, MD: University of Western Australia, Perth, Australia.

Additional supplementary information cited in this article can be found online in the Supporting Information section (http://onlinelibrary.wiley.com/doi/10.1002/acr.25621).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25621.

Address correspondence via email to Amanda J. Rickard, PhD, at amanda.rickard@monash.edu.

Submitted for publication October 21, 2024; accepted in revised form July 22, 2025.

SIGNIFICANCE & INNOVATIONS

- This study established serum uric acid (SUA) reference intervals specifically for healthy adults aged ≥70 years old stratified by sex. The reference range was 0.24 to 0.54 mmol/L for men and 0.19 to 0.48 mmol/L for women.
- Despite prior evidence suggesting links between SUA levels and health risks, this study found no association between SUA levels and all-cause mortality, disability-free survival, cardiovascular events, cancer incidence and death, or dementia.
- A novel finding was that women with low SUA levels had a significantly higher risk of fractures, highlighting a sex-specific health risk that may require further investigation.

associated with Alzheimer and Parkinson disease and increased mortality rate from all causes and CVD.[2,3] Given the growing clinical significance of SUA levels, identifying an optimal SUA range is increasingly relevant to clinical practice.

In reference intervals proposed over the past two decades, there has been notable variation in both the upper and lower cutoff values (summarized in Supplementary Table 1). For men, the upper limits have ranged from 0.41 to 0.55 mmol/L, and the lower limits have ranged from 0.16 to 0.24 mmol/L.[4,5] In women, the upper limits have ranged from 0.35 to 0.44 mmol/L, and the lower limits have ranged from 0.13 to 0.17 mmol/L.[4–6] This variability is likely influenced by factors such as difference in the age groups studied, the ethnicity of the populations examined, the prevalence of comorbidities, varying exclusion criteria, and differences in analytical methods across platforms.[7]

The purpose of this study was to establish a reference interval for SUA in a healthy, community-dwelling population aged ≥70 years old, defined as individuals free of dementia, disability, and cardiovascular events. Additionally, we examined the association between SUA levels and the risk of death and a range of morbidity outcomes relevant to older adults. By focusing on an exclusively older, community-dwelling population in which SUA measurements are common and chronic disease incidence is high, this study addresses an important gap in the literature. Data were sourced from the ASPirin in Reducing Events in the Elderly (ASPREE) study, which included an exclusively elderly cohort with extensive phenotyping and comprehensive long-term follow-up extending up to 11 years.

METHODS

Data source. The ASPREE study was a randomized clinical trial of 19,114 participants from Australia and the United States. The protocol and primary results of the study have been published elsewhere.[8] In brief, community-dwelling individuals aged ≥70 years old were recruited between March 2010 and December 2014. Exclusion criteria included a history of CVD events, current CVD events, dementia, physical disability, or a chronic illness likely to cause death within the next five years. Details of the recruitment process, along with a full list of the inclusion and exclusion criteria, are provided in the Supplementary Methods. In Australia, 11,878 ASPREE participants provided nonfasting blood samples for the analysis of a range of biomarkers including SUA (Figure 1).

Biospecimen collection and analysis. Blood samples were processed within four hours of collection, and serum was stored at −80°C for future analyses. SUA level was analyzed using a uricase/peroxidase colorimetric method (Abbott Alinity ci) with a coefficient of variation of 0.38% at 0.23 mmol/L and 1.2% at 0.57 mmol/L. Samples were analyzed at the Alfred Pathology Service, which is accredited by the National Association of Testing Authorities, adhering to the standards outlined in International Organization for Standardization 15189.[9]

Study sample and reference sample. The initial study population (study sample) comprised all Australian ASPREE participants with a SUA measurement at baseline, excluding individuals taking urate-lowering medication (Figure 1). The exclusion of individuals on allopurinol removed a group with pharmacologically lowered SUA levels, ensuring the remainder were reflective of physiologic SUA levels. The reference sample was established after further exclusion of participants with impaired renal function, defined as an estimated glomerular filtration rate (eGFR) <45 mL/min/1.73m^2 (Supplementary Table 2 and Supplementary Figure 1). The reference interval was defined as the values falling between the 2.5th and 97.5th percentiles of the SUA distribution within the reference sample.

Longitudinal outcomes. Longitudinal outcomes were determined via annual in-person visits, medical record reviews, and six monthly phone calls. These included all-cause mortality, disability-free survival (DFS), CVD, major adverse cardiac events (MACE), cancer incidence and death, dementia, and fractures. DFS, a primary outcome of the ASPREE trial, is a composite endpoint defined by the first occurrence of death, dementia, or persistent physical disability.[8] Death, dementia, and cancer were prespecified outcomes confirmed by adjudication committees comprised of specialist clinicians.[8,10,11] CVD encompassed fatal coronary heart disease, nonfatal myocardial infarction, fatal or nonfatal stroke, or hospitalization for heart failure. MACE comprised a composite of fatal and nonfatal ischemic stroke, nonfatal myocardial infarction, or fatal coronary heart disease. A diagnosis of dementia was based on *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* criteria,[12] which included the presence of memory impairment, specific domains of cognitive disturbance, and an impact on functioning. Fractures included

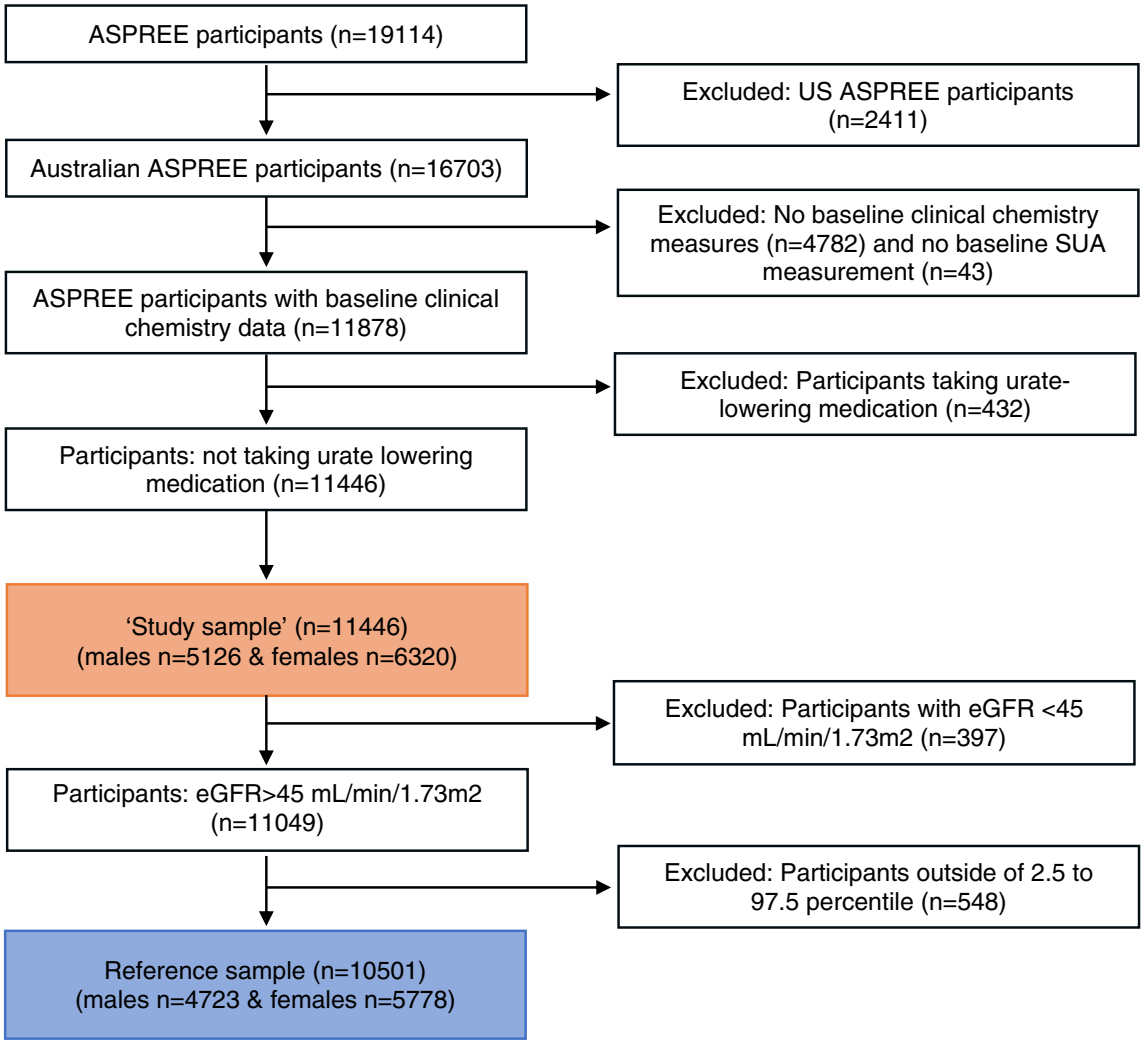

Figure 1. Consort flow-chart

vertebral, hip, and other fractures that were confirmed by medical imaging.

Analytical approach. To explore the cross-sectional associations between SUA values and health measures, SUA levels in both the study sample and the reference sample were divided into quintiles. Quintile one represented the lowest 20% of SUA levels, quintile five represented the highest 20%, and quintile two to four were combined to represent the central 60%. Baseline characteristics, including self-reported health issues, were assessed across these quintiles. Continuous variables were presented as median (interquartile range [IQR]) or mean (SD) and compared using the Kruskal–Wallis or analysis of variance test as appropriate. Categorical variables were expressed as proportions and compared using a chi-square test.

Key health measures assessed included smoking status (current and former), alcohol use, body mass index (BMI), physical activity (self-reported Lifestyle Interventions and Independence for Elders (LIFE) Disability questionnaire with low physical activity defined as "No walking outside the home or walked outside home but longest amount of time walked without sitting down to rest was less than 10 minutes"), and frailty status defined from five measures adapted from the Fried frailty domains (body weight, strength, exhaustion, walking speed, and physical activity).[13] Frailty status was stratified according to the number of domains affected, that is, prefrail (one or two domains) and frail (three, four, or five domains). Hypertension was defined as a systolic blood pressure ≥140 mm Hg, a diastolic blood pressure ≥90 mm Hg, or a current use of antihypertensive medication. Other covariates included hyperlipidaemia, diabetes, albuminuria (urinary albumin/creatinine ratios), eGFR (CKD-epidemiology estimating equation derived from serum creatinine) and abnormal liver enzymes (gamma-glutamyl transpeptidase [GGT]).

To examine the association between SUA levels and clinical outcomes, Cox proportional hazards models were used to estimate cause-specific hazard ratios (HR) and 95% confidence

intervals (CI). Quintiles two to four were the reference. ASPREE clinical outcomes included all-cause mortality, DFS, CVD, MACE, stroke, cancer incidence and death, dementia, and fractures. Analyses were conducted using both unadjusted and fully adjusted models. The fully adjusted model accounted for age (continuous), BMI (continuous), smoking status (categorical: current or former), current alcohol use (categorical), physical activity level (categorical), hypertension (categorical), diabetes (categorical), high-density lipoprotein cholesterol (HDL-C; continuous), non–HDL-C (continuous), triglycerides (continuous), eGFR (continuous), and GGT (categorical: greater than 50 U/L). The statistical analyses were performed using Stata software version 17.0 (StataCorp LLC).

Ethics approval for the principal ASPREE trial was obtained from the Monash University Human Research Ethics committee (2006/745MC and CF11/1100) and for the ASPREE Healthy Ageing Biobank from the Alfred Hospital Ethics Committee (18/08). Written consent was obtained from all participants.

RESULTS

Study population. The SUA reference intervals for this older White population were established as 0.24 to 0.54 mmol/L for men and 0.19 to 0.48 mmol/L for women. A comparison of SUA levels from men and women using the Tukey-Kramer test showed that men had significantly higher mean SUA levels (0.38 mmol/L; SD 0.08) than women (0.32 mmol/L; SD 0.08; Supplementary Table 3 and Supplementary Figure 2). Given this sex difference, all analyses were conducted separately for men and women. Of the 11,878 Australian participants who provided a blood sample at baseline, 432 individuals on urate-lowering medication were excluded. The resulting study sample included 11,446 participants. The reference sample was further refined by excluding 397 participants with an eGFR below 45 mL/min/1.73m^2, yielding a final sample of 10,501 participants.

The study sample (n = 11,446) had a median age of 74 years and consisted of 6,320 (55%) women (Table 1). Median SUA concentrations were 0.38 mmol/L (IQR 0.33–0.43) for men and 0.31 mmol/L (IQR 0.26–0.36) for women. The reference sample (n = 10,501) had a median age of 74 years and consisted of 5,778 (55%) women (Table 1). Median SUA concentrations were 0.37 mmol/L (IQR 0.33–0.42) for men and 0.31 mmol/L (IQR 0.26–0.35) for women. Among men, 2,605 participants (55.2%) were identified as current or former smokers, and 4,020 (85%) reported current alcohol consumption. In contrast, among women, 2,009 participants (35%) were current or former smokers, and 4,369 (76%) reported current alcohol consumption. The mean BMI was 28 for both men and women, and the waist-to-height ratio was 0.59 for men and 0.58 for women.

Baseline cross-sectional analysis. Kernel density curves for SUA concentrations in the study and reference samples were analyzed by age, diuretics use, BMI, and alcohol consumption separately in men (Supplementary Figure 3) and women (Supplementary Figure 4). Although significant shifts in SUA distributions were observed when stratified by diuretic use, BMI, and alcohol intake, these differences were deemed insufficient to justify establishing separate reference intervals.

The baseline characteristics of the participants in the study and reference samples, stratified by SUA quintiles, are shown in Table 1. In both samples, male and female participants with higher SUA levels (quintile 5 [Q5]) were characterized by a higher BMI, greater waist-to-height ratio, increased total cholesterol–to–HDL-C ratio, higher prevalence of hypertension, and greater frequency of impaired renal function (eGFR <60 mL/min/1.73m^2) and elevated liver enzymes (alanine aminotransaminase [ALT], aspartate aminotransferase [AST], and GGT). Among women, higher SUA levels were also associated with a greater prevalence of diabetes.

SUA levels and longitudinal outcomes. In the study and reference sample, over a median follow-up ranging from 8.3 (IQR 6.9–9.4) to 8.6 (IQR 7.5–10.1) years, no significant relationships were identified between SUA quintiles and all-cause mortality, CVD, MACE, or cancer mortality in the fully adjusted model (Table 2). Additionally, during a median follow-up ranging from 6.0 (IQR 4.9–6.9) to 6.4 (IQR 5.3–7.5) years, no associations were observed between SUA quintiles and DFS, or dementia in either men or women. In women, low SUA levels (Q1) were associated with a higher risk of fractures in the fully adjusted model after a median follow-up of 4.0 years (IQR 2.9–5.0; study sample: HR 1.23, 95% CI 1.04–1.46; reference sample: HR 1.22, 95% CI 1.02–1.46). Restricted spline curves demonstrated a nonlinear trend between SUA levels and risk of fractures in men and women (Supplementary Figure 5).

DISCUSSION

In this study, we established SUA reference intervals based on data from a relatively healthy population of individuals aged ≥70 years old. After adjusting for potential confounders, SUA levels in both the study and reference populations were not associated with a range of clinical events including dementia. In women only, however, low SUA levels were associated with a higher risk of fractures.

Existing reference intervals for SUA have been established primarily using data from middle-aged populations across various ethnic groups (Supplementary Table 1). For instance, the Brazilian Longitudinal Study of Adult Health recently proposed SUA reference intervals in healthy adults aged 35 to 74 years old to range from 0.24 to 0.55 mmol/L for men and 0.17 to 0.41 mmol/L for women.[4] Similarly, SUA reference intervals were established in a

Table 1. Baseline participant characteristics according to SUA quintiles in both study and reference sample, stratified by sex*

	Study sample				Reference sample			
	Total	SUA quintiles			Total	SUA quintiles		
		Q1	Q2–4	Q5		Q1	Q2–4	Q5
Men[a]								
SUA, median (IQR), mmol/L	0.38 (0.33–0.43)	0.28 (0.26–0.30)	0.38 (0.35–0.41)	0.49 (0.46–0.52)	0.37 (0.33–0.42)	0.29 (0.27–0.31)	0.38 (0.35–0.40)	0.47 (0.45–0.50)
Age, median (IQR), y	73.8 (71.6–77.2)	73.8 (71.7–77.6)	73.8 (71.6–77.2)	73.8 (71.8–76.9)	73.7 (71.6–77.1)	73.9 (71.7–77.0)	73.7 (71.6–76.5)	73.6 (71.6–76.5)
Current or former smoker, n (%)	2,842 (55.4)	536 (51.6)	1,704 (55.3)	602 (59.9)	2,605 (55.2)	590 (52.6)	1,453 (54.5)	562 (60.1)
Current alcohol use, n (%)	4,356 (85.0)	868 (83.5)	2,607 (84.6)	881 (87.7)	4,020 (85.1)	929 (82.8)	2,264 (84.9)	827 (88.4)
Vegetarian diet, n (%)	63 (1.34)	22 (2.33)	36 (1.27)	5 (0.54)	59 (1.36)	20 (1.96)	33 (1.35)	6 (0.70)
BMI, mean (SD)	27.8 (3.73)	26.7 (3.44)	27.8 (3.59)	29.2 (3.97)	27.8 (3.7)	26.8 (3.5)	27.7 (3.5)	29.1 (3.9)
BMI, n (%)								
<21	80 (1.57)	32 (3.09)	40 (1.30)	7 (0.80)	69 (1.47)	24 (2.63)	38 (1.26)	7 (0.89)
21–29	3,792 (74.21)	850 (81.97)	2,319 (75.54)	623 (62.11)	3,523 (74.83)	742 (82.55)	2,274 (75.62)	497 (62.91)
≥30	1,238 (24.23)	155 (14.95)	711 (23.16)	372 (37.09)	1,116 (23.70)	135 (14.82)	695 (23.11)	286 (36.20)
Waist to height ratio, mean (SD)	0.59 (0.06)	0.57 (0.06)	0.59 (0.06)	0.61 (0.06)	0.59 (0.06)	0.57 (0.06)	0.59 (0.06)	0.61 (0.06)
Moderate to high physical activity, n (%)	3,467 (67.6)	714 (68.7)	2,141 (69.5)	612 (60.9)	3,216 (68.1)	774 (69.0)	1,854 (69.5)	588 (63.0)
Frailty status, n (%)								
Not frail	3,320 (64.8)	683 (65.7)	2,012 (65.3)	625 (62.2)	3,093 (65.5)	747 (66.6)	1,748 (65.6)	598 (64.0)
Prefrail or frail	1,806 (35.2)	356 (34.3)	1,070 (34.7)	380 (37.8)	1,630 (34.5)	375 (33.4)	918 (34.4)	337 (36.0)
Hypertension, n (%)	3,849 (75.1)	718 (69.1)	2,283 (74.01)	848 (84.4)	3,511 (74.3)	766 (68.3)	1,984 (74.4)	761 (81.4)
Antihypertensive medication, n (%)	2,437 (51.0)	388 (40.5)	1,393 (48.8)	656 (67.6)	2,179 (49.6)	416 (40.2)	1,195 (48.5)	568 (63.3)
Diuretics, n (%)	745 (14.5)	66 (6.35)	402 (13.0)	277 (27.6)	637 (13.5)	73 (6.51)	352 (13.2)	212 (22.7)
Non–HDL-C, mean (SD), mmol/L	3.64 (0.89)	3.53 (0.93)	3.65 (0.88)	3.72 (0.87)	3.64 (0.89)	3.55 (0.91)	3.65 (0.87)	3.72 (0.88)
HDL-C, mean (SD), mmol/L	1.41 (0.39)	1.51 (0.40)	1.40 (0.38)	1.33 (0.37)	1.41 (0.39)	1.50 (0.40)	1.39 (0.37)	1.35 (0.38)
Triglyceride, mean (SD), mmol/L	1.31 (0.67)	1.12 (0.55)	1.30 (0.63)	1.54 (0.81)	1.30 (0.66)	1.14 (0.56)	1.30 (0.63)	1.51 (0.78)
Diabetes, n (%)	584 (11.4)	125 (12.0)	340 (11.0)	119 (11.8)	512 (10.8)	122 (10.9)	284 (10.7)	106 (11.3)
Glucose, mean (SD), mmol/L	6.05 (1.71)	6.04 (2.00)	6.00 (1.60)	6.22 (1.69)	6.03 (1.66)	5.99 (1.84)	5.97 (1.54)	6.23 (1.74)
Chronic kidney disease, n (%)	778 (15.6)	72 (7.15)	408 (13.7)	298 (30.4)	592 (12.9)	74 (6.8)	326 (12.6)	192 (21.1)
eGFR, mean (SD), mL/min/1.73m^2	73.3 (12.8)	78.6 (10.8)	73.5 (12.2)	66.9 (13.8)	74.2 (11.8)	78.0 (10.8)	74.0 (11.6)	70.1 (11.9)
uACR, median (IQR)	0.7 (0.4–1.7)	0.7 (0.4–1.4)	0.7 (0.4–1.4)	0.7 (0.4–1.7)	0.7 (0.4–1.3)	0.6 (0.4–1.3)	0.7 (0.4–1.3)	0.7 (0.4–1.6)
GGT, median (IQR), U/L	24 (18–34)	22 (16–30.5)	23 (17–33)	29 (21–43)	24 (18–34)	22 (17–31)	23 (17–33)	28 (21–41)
GGT, n (%)								
<50	4,542 (88.8)	953 (92.0)	2,770 (90.0)	819 (81.8)	4,197 (89.1)	1,024 (91.4)	2,404 (90.3)	769 (82.7)
>50	572 (11.2)	83 (8.0)	307 (10.0)	182 (18.2)	516 (11.0)	96 (8.6)	259 (9.7)	161 (17.3)
ALT, mean (SD), U/L	21.9 (11.5)	20.5 (13.6)	21.8 (10.2)	24.1 (12.6)	21.9 (11.3)	20.7 (13.4)	21.7 (10.4)	23.9 (10.9)
AST, mean (SD), U/L	22.4 (8.08)	22.1 (9.15)	22.2 (6.94)	23.5 (9.88)	22.4 (7.6)	22.0 (8.9)	22.2 (7.1)	23.3 (7.1)
Women[b]								
SUA, median (IQR), mmol/L	0.31 (0.26–0.36)	0.23 (0.21–0.24)	0.31 (0.28–0.34)	0.43 (0.40–0.46)	0.31 (0.26–0.35)	0.23 (0.21–0.24)	0.31 (0.28–0.34)	0.41 (0.39–0.44)
Age, median (IQR), y	73.9 (71.7–77.5)	73.6 (71.5–77.0)	73.8 (71.7–77.4)	74.6 (71.9–78.4)	73.8 (71.6–77.2)	73.6 (71.5–77.0)	73.7 (71.6–77.3)	74.3 (71.8–77.6)
Current or former smoker, n (%)	2,185 (34.6)	479 (35.0)	1,300 (34.3)	406 (35.1)	2,009 (34.8)	434 (35.7)	1,218 (34.4)	357 (35.1)
Current alcohol use, n (%)	4,731 (74.9)	1,010 (73.7)	3,892 (76.3)	829 (71.7)	4,369 (75.6)	904 (74.3)	2,709 (76.4)	756 (74.4)
Vegetarian diet, n (%)	166 (2.88)	49 (3.88)	99 (2.87)	18 (1.72)	150 (2.85)	41 (3.65)	95 (2.95)	14 (1.52)
BMI, mean (SD)	28.0 (5.04)	25.7 (4.18)	28.0 (4.76)	30.8 (5.46)	27.9 (4.9)	25.8 (4.2)	27.9 (4.8)	30.5 (5.3)
BMI, n (%)								
<21	313 (4.98)	140 (10.27)	154 (4.07)	19 (1.66)	275 (4.78)	114 (9.41)	151 (4.08)	10 (1.19)
21–29	4,045 (64.31)	1,035 (75.94)	2,494 (65.98)	516 (44.99)	3,762 (65.38)	923 (76.22)	2,443 (65.96)	396 (47.20)

(Continued)

Table 1. (*Cont'd*)

	Study sample				Reference sample			
	Total	SUA quintiles			Total	SUA quintiles		
		Q1	Q2–4	Q5		Q1	Q2–4	Q5
≥30	1,932 (30.72)	188 (13.79)	1,132 (29.95)	612 (53.36)	1,717 (29.84)	174 (14.37)	1,110 (29.97)	433 (51.61)
Waist to height ratio, mean (SD)	0.58 (0.08)	0.55 (0.07)	0.58 (0.08)	0.63 (0.08)	0.58 (0.09)	0.55 (0.07)	0.58 (0.08)	0.62 (0.08)
Moderate to high physical activity, n (%)	3,827 (60.6)	920 (67.2)	2,305 (60.8)	602 (52.0)	3,525 (61.0)	808 (66.5)	2,151 (60.7)	599 (55.7)
Frailty status, n (%)								
Not frail	4,000 (63.3)	896 (65.4)	2,455 (64.7)	649 (56.1)	3,712 (64.2)	797 (67.1)	2,312 (65.2)	603 (59.4)
Prefrail or frail	2,320 (36.7)	474 (34.6)	1,338 (35.3)	508 (43.9)	2,066 (35.8)	419 (34.5)	1,234 (34.8)	413 (40.6)
Hypertension, n (%)	4,631 (73.3)	861 (62.9)	2,777 (73.2)	993 (85.8)	4,185 (72.4)	770 (63.3)	2,569 (72.5)	846 (83.3)
Antihypertensive medication, n (%)	3,416 (55.9)	571 (43.7)	1,983 (54.1)	862 (75.9)	3,028 (54.3)	509 (43.8)	1,810 (52.9)	709 (71.3)
Diuretics, n (%)	1,344 (21.3)	157 (11.5)	696 (18.7)	666 (57.6)	1,140 (19.7)	140 (11.5)	626 (17.7)	374 (36.8)
Non–HDL-C, mean (SD), mmol/L	3.73 (0.96)	3.62 (0.92)	3.74 (0.96)	3.83 (1.00)	3.73 (0.95)	3.64 (0.92)	3.74 (0.96)	3.80 (0.97)
HDL-C, mean (SD), mmol/L	1.74 (0.46)	1.86 (0.46)	1.74 (0.45)	1.57 (0.42)	1.74 (0.46)	1.86 (0.46)	1.75 (0.45)	1.59 (0.42)
Triglyceride, mean (SD), mmol/L	1.31 (0.61)	1.09 (0.49)	1.29 (0.57)	1.65 (0.70)	1.29 (0.58)	1.09 (0.49)	1.28 (0.56)	1.58 (0.64)
Diabetes, n (%)	491 (7.77)	71 (5.18)	269 (7.09)	151 (13.1)	445 (7.7)	63 (5.2)	248 (7.0)	134 (13.2)
Glucose, mean (SD), mmol/L	5.75 (1.43)	5.65 (1.62)	5.70 (1.32)	6.00 (1.54)	5.74 (1.43)	5.65 (1.61)	5.70 (1.33)	5.96 (1.53)
Chronic kidney disease, n (%)	1,125 (18.3)	88 (6.59)	574 (15.6)	463 (40.8)	828 (14.8)	77 (6.5)	475 (13.8)	276 (27.9)
eGFR, mean (SD), mL/min/1.73m^2	72.9 (13.6)	78.8 (10.8)	73.4 (12.7)	64.4 (15.2)	74.1 (12.4)	78.6 (10.7)	74.0 (12.1)	68.9 (13.1)
uACR, median (IQR)	1.0 (0.6–1.7)	1.0 (0.6–1.9)	0.9 (0.6–1.7)	1.0 (0.6–1.9)	0.9 (0.6–1.7)	1.0 (0.6–1.8)	0.9 (0.6–1.6)	1.0 (0.5–1.8)
GGT, median (IQR), U/L	19 (17–34)	17 (13–24)	19 (14–28)	23 (17–34)	19 (14–28)	17 (13–24)	19 (14–28)	23 (17–34)
GGT, n (%)								
<50	5,816 (92.0)	1,304 (95.2)	3,485 (91.9)	1,027 (88.8)	5,319 (92.1)	1,152 (94.7)	3,255 (91.8)	912 (89.8)
>50	504 (8.0)	66 (4.8)	308 (8.1)	130 (11.2)	459 (7.9)	64 (5.3)	291 (8.2)	104 (10.2)
ALT, mean (SD), U/L	18.8 (9.84)	17.4 (6.95)	18.8 (10.1)	20.4 (11.6)	18.8 (9.5)	17.4 (7.1)	18.8 (10.1)	20.2 (9.5)
AST, mean (SD), U/L	21.3 (7.44)	21.1 (5.76)	21.3 (7.23)	21.7 (9.6)	21.3 (7.0)	21.1 (5.9)	21.3 (7.2)	21.4 (7.4)

* ALT, alanine aminotransaminase; AST, aspartate aminotransferase; BMI, body mass index; eGFR, estimated glomerular filtration rate; GGT, gamma-glutamyl transferase; HDL-C, high-density lipoprotein cholesterol; IQR, interquartile range; Q, quintile; SUA, serum uric acid; uACR, urine albumin to creatinine ratio.
[a] Study sample: total n = 5,126, Q1 n = 1,039, Q2–4 n = 3,082, Q5 = 1,005; reference sample: total n = 4,723, Q1 n = 1,122, Q2–4 n = 2,666, Q5 n = 935.
[b] Study sample: total n = 6,320, Q1 n = 1,370, Q2–4 n = 3,793, Q5 n = 1,157; reference sample: total n = 5,778, Q1 n = 1,216, Q2–4 n = 3,546, Q5 n = 1,016.

Table 2. Unadjusted and adjusted hazard ratios for association between SUA quintiles and clinical outcomes in both study and reference sample, stratified by sex*

	Study sample				Reference sample			
	Total	SUA quintiles			Total	SUA quintiles		
		Q1	Q2–Q4	Q5		Q1	Q2–Q4	Q5
Men[a]								
SUA, median (IQR), mmol/L	0.38 (0.33–0.43)	0.28 (0.26–0.30)	0.38 (0.35–0.41)	0.49 (0.46–0.52)	0.37 (0.33–0.42)	0.29 (0.27–0.31)	0.38 (0.35–0.40)	0.47 (0.45–0.50)
All-cause mortality								
Number (rate per 1,000 PY)	881 (20.3)	171 (19.3)	519 (19.9)	191 (22.7)	788 (19.7)	189 (19.8)	434 (19.2)	165 (21.0)
Age adjusted HR (95% CI)		0.89 (0.75–1.06)	1.00 (ref)	1.18 (1.00–1.39)		0.97 (0.82–1.16)	1.00 (ref)	1.15 (0.96–1.37)
Fully adjusted HR (95% CI)[b]		0.85 (0.71–1.03)	1.00 (ref)	1.16 (0.97–1.39)		0.93 (0.78–1.12)	1.00 (ref)	1.13 (0.94–1.37)
Disability-free survival[c]								
Number (rate per 1,000 PY)	812 (25.1)	156 (23.7)	484 (24.8)	172 (27.2)	740 (24.8)	169 (23.8)	417 (24.7)	154 (26.1)
Age adjusted HR (95% CI)		0.90 (0.75–1.08)	1.00 (ref)	1.12 (0.94–1.34)		0.92 (0.77–1.10)	1.00 (ref)	1.11 (0.92–1.33)
Fully adjusted HR (95% CI)[b]		0.88 (0.72–1.07)	1.00 (ref)	1.10 (0.91–1.33)		0.90 (0.74–1.08)	1.00 (ref)	1.07 (0.88–1.30)
Cardiovascular disease								
Number (rate per 1,000 PY)	624 (15.5)	120 (14.6)	363 (14.9)	141 (18.2)	568 (15.3)	130 (14.7)	310 (14.7)	128 (17.6)
Age adjusted HR (95% CI)		0.94 (0.77–1.16)	1.00 (ref)	1.25 (1.03–1.52)		0.98 (0.80–1.20)	1.00 (ref)	1.24 (1.01–1.52)
Fully adjusted HR (95% CI)[b]		1.07 (0.86–1.28)	1.00 (ref)	1.03 (0.83–1.28)		1.11 (0.90–1.38)	1.00 (ref)	1.09 (0.88–1.35)
MACE								
Number (rate per 1,000 PY)	505 (12.5)	94 (11.4)	295 (12.1)	116 (14.8)	463 (12.4)	105 (11.8)	253 (11.8)	105 (14.4)
Age adjustedHR (95% CI)		0.91 (0.72–1.15)	1.00 (ref)	1.26 (1.01–1.56)		0.97 (0.77–1.21)	1.00 (ref)	1.24 (0.99–1.56)
Fully adjusted HR (95% CI)[b],		1.04 (0.82–1.33)	1.00 (ref)	1.04 (0.82–1.32)		1.10 (0.87–1.40)	1.00 (ref)	1.10 (0.87–1.41)
Dementia[c]								
Number (rate per 1,000 PY)	327 (7.97)	75 (8.98)	200 (8.10)	52 (6.50)	307 (8.11)	80 (8.86)	177 (8.29)	50 (5.07)
Age adjusted HR (95% CI)		1.05 (0.81–1.37)	1.00 (ref)	0.82 (0.61–1.11)		1.04 (0.80–1.35)	1.00 (ref)	0.85 (0.62–1.16)
Fully adjusted HR (95% CI)[b]		0.95 (0.71–1.27)	1.00 (ref)	0.93 (0.67–1.29)		0.98 (0.74–1.30)	1.00 (ref)	0.89 (0.64–1.25)
Cancer incidence								
Number (rate per 1,000 PY)	1,214 (32.1)	209 (26.9)	749 (33.0)	256 (34.8)	1,115 (31.9)	222 (26.4)	653 (33.3)	240 (34.8)
Age adjusted HR (95% CI)		0.81 (0.69–0.94)	1.00 (ref)	1.06 (0.92–1.22)		0.79 (0.68–0.92)	1.00 (ref)	1.06 (0.91–1.23)
Fully adjusted HR (95% CI)[b]		0.81 (0.70–0.96)	1.00 (ref)	1.06 (0.90–1.23)		0.80 (0.68–0.93)	1.00 (ref)	1.06 (0.90–1.24)
Cancer mortality								
Number (rate per 1,000 PY)	397 (9.16)	70 (7.9)	234 (8.97)	93 (11.1)	358 (8.9)	75 (7.9)	203 (9.0)	80 (10.2)
Age adjusted HR (95% CI)		0.83 (0.64–1.09)	1.00 (ref)	1.26 (0.99–1.56)		0.85 (0.64–1.10)	1.00 (ref)	1.17 (0.90–1.50)
Fully adjusted HR (95% CI)[b]		0.84 (0.63–1.11)	1.00 (ref)	1.26 (0.97–1.64)		0.86 (0.65–1.13)	1.00 (ref)	1.14 (0.87–1.50)
Fractures[d]								
Number (rate per 1,000 PY)	308 (14.3)	61 (13.8)	196 (15.2)	51 (12.0)	287 (14.4)	67 (14.1)	173 (15.5)	47 (11.9)
Age adjusted HR (95% CI)		0.89 (0.67–1.19)	1.00 (ref)	0.80 (0.59–1.09)		0.90 (0.68–1.19)	1.00 (ref)	0.78 (0.56–1.07)
Fully adjusted HR (95% CI)[b]		0.76 (0.57–1.04)	1.00 (ref)	0.98 (0.70–1.36)		0.78 (0.58–1.06)	1.00 (ref)	0.91 (0.65–1.28)
Women[e]								
SUA, median (IQR), mmol/L	0.31 (0.26–0.36)	0.23 (0.21–0.24)	0.31 (0.28–0.34)	0.43 (0.40–0.46)	0.31 (0.26–0.35)	0.23 (0.21–0.24)	0.31 (0.28–0.34)	0.41 (0.39–0.44)
All-cause mortality								
Number (rate per 1,000 PY)	756 (13.7)	154 (12.9)	428 (12.9)	174 (17.5)	658 (13.1)	140 (13.2)	385 (12.4)	133 (15.1)
Age adjusted HR (95% CI)		1.07 (0.89–1.29)	1.00 (ref)	1.27 (1.06–1.51)		1.13 (0.93–1.37)	1.00 (ref)	1.20 (0.99–1.46)
Fully adjusted HR (95% CI)[b]		1.07 (0.88–1.30)	1.00 (ref)	1.10 (0.90–1.34)		1.10 (0.90–1.36)	1.00 (ref)	1.12 (0.91–1.39)
Disability-free survival[c]								
Number (rate per 1,000 PY)	789 (19.4)	162 (18.4)	453 (18.5)	174 (23.8)	684 (18.4)	142 (18.2)	421 (18.4)	121 (18.6)
Age adjusted HR (95% CI)		1.05 (0.88–1.26)	1.00 (ref)	1.19 (1.00–1.42)		1.03 (0.85–1.25)	1.00 (ref)	0.98 (0.80–1.20)

(Continued)

Table 2. (*Cont'd*)

	Study sample				Reference sample			
	Total	SUA quintiles			Total	SUA quintiles		
		Q1	Q2–Q4	Q5		Q1	Q2–Q4	Q5
Fully adjusted HR (95% CI)[b]		1.14 (0.94–1.37)	1.00 (ref)	0.99 (0.82–1.21)		1.10 (0.90–1.35)	1.00 (ref)	0.85 (0.69–1.05)
Cardiovascular disease								
Number (rate per 1,000 PY)	509 (9.92)	100 (8.96)	304 (9.84)	105 (11.3)	455 (9.65)	89 (8.97)	277 (9.56)	89 (10.8)
Age adjusted HR (95% CI)		0.96 (0.76–1.20)	1.00 (ref)	1.06 (0.85–1.32)		0.99 (0.78–1.25)	1.00 (ref)	1.10 (0.87–1.39)
Fully adjusted HR (95% CI)[b]		1.04 (0.82–1.33)	1.00 (ref)	0.92 (0.72–1.19)		1.04 (0.81–1.34)	1.00 (ref)	1.03 (0.80–1.33)
MACE								
Number (rate per 1,000 PY)	363 (7.03)	68 (6.05)	223 (7.19)	72 (7.71)	324 (6.83)	62 (6.21)	204 (7.01)	58 (7.00)
Age adjusted HR (95% CI)		0.87 (0.67–1.15)	1.00 (ref)	1.00 (0.67–1.15)		0.92 (0.69–1.22)	1.00 (ref)	0.97 (0.73–1.30)
Fully adjusted HR (95% CI)[b]		0.96 (0.72–1.29)	1.00 (ref)	0.83 (0.61–1.11)		0.97 (0.72–1.32)	1.00 (ref)	0.88 (0.64–1.20)
Dementia[c]								
Number (rate per 1,000 PY)	385 (7.44)	97 (8.67)	222 (7.11)	66 (7.04)	345 (7.26)	86 (8.65)	210 (7.19)	49 (5.85)
Age adjusted HR (95% CI)		1.29 (1.02–1.64)	1.00 (ref)	0.94 (0.71–1.23)		1.26 (0.98–1.62)	1.00 (ref)	0.80 (0.59–1.09)
Fully adjusted HR (95% CI)[b]		1.13 (0.87–1.46)	1.00 (ref)	1.05 (0.77–1.43)		1.13 (0.87–1.48)	1.00 (ref)	0.91 (0.66–1.27)
Cancer incidence								
Number (rate per 1,000 PY)	943 (19.0)	186 (17.2)	571 (19.1)	186 (20.9)	851 (18.7)	165 (17.2)	527 (18.8)	159 (20.0)
Age adjusted HR (95% CI)		0.91 (0.77–1.07)	1.00 (ref)	1.08 (0.92–1.28)		0.92 (0.77–1.10)	1.00 (ref)	1.06 (0.89–1.27)
Fully adjusted HR (95% CI)[b]		0.96 (0.80–1.14)	1.00 (ref)	0.98 (0.82–1.18)		0.95 (0.79–1.14)	1.00 (ref)	1.01 (0.84–1.23)
Cancer mortality								
Number (rate per 1,000 PY)	331 (6.02)	55 (4.61)	203 (6.13)	73 (7.34)	286 (5.67)	49 (4.63)	180 (5.81)	57 (6.46)
Age adjusted HR (95% CI)		0.78 (0.58–1.05)	1.00 (ref)	1.15 (0.88–1.51)		0.82 (0.60–1.48)	1.00 (ref)	1.10 (0.81–1.48)
Fully adjusted HR (95% CI)[b]		0.81 (0.59–1.11)	1.00 (ref)	0.95 (0.70–1.29)		0.79 (0.57–1.11)	1.00 (ref)	1.03 (0.74–1.42)
Fractures[d]								
Number (rate per 1,000 PY)	835 (31.3)	214 (37.8)	500 (31.2)	121 (24.4)	753 (30.9)	188 (37.5)	460 (30.7)	105 (24.2)
Age adjusted HR (95% CI)		1.25 (1.07–1.47)	1.00 (ref)	0.75 (0.61–0.91)		1.26 (1.06–1.49)	1.00 (ref)	0.78 (0.63–0.97)
Fully adjusted HR (95% CI)[b]		1.23 (1.04–1.46)	1.00 (ref)	0.77 (0.62–0.96)		1.22 (1.02–1.46)	1.00 (ref)	0.80 (0.64–1.00)

* IQR, interquartile range; HR, hazard ratio; CI, confidence interval; MACE, major adverse cardiovascular events; PY, person-years; Q, quintile; ref, reference; SUA, serum uric acid.
[a] Study sample: total n = 5,126, Q1 n = 1,039, Q2–Q4 n = 3,082, Q5 n = 1,005; reference sample: total n = 4,723, Q n = 1,122, Q2–Q4 n = 2,666, Q5 n = 935.
[b] Fully adjusted model: adjusted for age, body mass index, smoking status (current or former), current alcohol use, physical activity level, hypertension, diabetes, HDL-C, non–HDL-C, triglycerides, estimated glomerular filtration rate, and gamma-glutamyl transpeptidase (>50 U/L).
[c] Data from up to nine years of follow-up.
[d] Data for clinical trial only.
[e] Study sample: total n = 6,320, Q1 n = 1,370, Q2–Q4 n = 3,793, Q5 n = 1,157; reference sample: total n = 5,778, Q1 n = 1,216, Q2–Q4 n = 3,546, Q5 n = 1,016.

healthy Ethiopian population aged 18 to 60 years old that ranged from 0.16 to 0.41 mmol/L for men and 0.13 to 0.35 mmol/L for women.[5] Only one study reported reference intervals in an exclusively older healthy geriatric Chinese population (aged 60–96 years), which were lower than those found in the present study: 0.18 to 0.46 mmol/L for men and 0.13 to 0.44 mmol/L for women.[6] Variations in factors such as ethnicity, medication use, diet, and the presence of comorbidities may explain these differences. The revised SUA reference intervals presented in the current study, derived from a predominantly White population, are higher for both men and women than previously published intervals.

The baseline cross-sectional analysis revealed associations between SUA levels and several baseline characteristics that may have confounded an unadjusted analysis. Participants in the highest quintile of SUA (Q5) exhibited a higher BMI, waist-to-height ratio, total cholesterol–to–HDL-C ratio, and a greater prevalence of CKD, hypertension, antihypertensive medication use, and elevated liver enzymes (ALT, AST, and GGT). This data support previous research linking elevated SUA levels to an adverse cardiometabolic profile.[14]

In both the study and reference samples, after adjusting for cardiovascular risk factors, SUA levels were not significantly associated with clinical outcomes, including all-cause mortality, DFS, CVD, MACE, cancer incidence or mortality, or dementia. In women, however, low SUA levels were associated with an increased risk of fractures. Although early studies on the relationship between SUA and fracture risk have produced inconsistent findings, more recent research supports a protective role of higher SUA levels. For example, Romero et al reported that in a cohort of women who were menopausal, higher SUA levels were associated with increased bone mineral density and a lower incidence of vertebral fractures, which is consistent with our findings.[15] Although SUA is known to influence several biochemical pathways, including acting as an antioxidant at physiologic concentrations, the mechanisms linking SUA to bone strength and fracture risk remain to be elucidated.

Our findings align with recent research using an innovative approach (truncated minimum chi-square) that continuously assesses reference intervals based on age, avoiding abrupt shifts between age groups.[16] Their study confirmed that SUA levels increase with age and, consistent with our results, reported that average levels in women are lower than in men, although they converge at older ages. Additionally, they observed diurnal variation in SUA levels and highlighted differences in measurements across various analytical systems. These factors underscore the need for context-specific interpretation of SUA levels in clinical practice.

The strengths of this study include the extensive characterization of the cohort at baseline, prolonged follow-up, and adjudication of longitudinal outcomes by expert panels. The principal limitation is the lack of data on recurrent gout or urate stones, which are key clinical targets for treatment. In clinical practice, the SUA level targeted in treating gout is within the reference interval. The main contribution of this report is the reassurance it provides, that SUA levels within and near the proposed reference intervals are not significantly associated with increased morbidity or mortality. An important exception is the potential association between low SUA levels and an increased risk of fractures in women. This finding may have clinical relevance when treating older women who already have an elevated risk of fracture because of osteoporosis or osteopenia. Although the findings presented here are specific to individuals aged 70 years and older, limiting generalizability to younger populations, they provide valuable, age-specific reference data for older adults. Additionally, the predominantly Caucasian population of European ancestry may limit the applicability of these findings to other ethnic groups.

In summary, the proposed reference interval for SUA in a predominantly White population aged ≥70 years old are 0.24 to 0.54 mmol/L for men and 0.19 to 0.48 mmol/L for women. Despite previous reports linking SUA levels with an increased incidence of various chronic diseases, our analysis identified no such relationships after adjusting for cardiovascular risk factors, except for an increased fracture risk in women with low SUA.

ACKNOWLEDGMENTS

Open access publishing facilitated by Monash University, as part of the Wiley - Monash University agreement via the Council of Australian University Librarians.

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 Rickard 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. Crawley WT, Jungels CG, Stenmark KR, et al. U-shaped association of uric acid to overall-cause mortality and its impact on clinical management of hyperuricemia. Redox Biol 2022;51:102271.
2. Wen S, Arakawa H, Tamai I. Uric acid in health and disease: from physiologic functions to pathogenic mechanisms. Pharmacol Ther 2024;256:108615.
3. Tseng WC, Chen YT, Ou SM, et al; Taiwan Geriatric Kidney Disease (TGKD) Research Group. U-shaped association between serum uric acid levels with cardiovascular and all-cause mortality in the elderly: the role of malnourishment. J Am Heart Assoc 2018;7(4):e007523.
4. Dório M, Benseñor IM, Lotufo P, et al. Reference range of serum uric acid and prevalence of hyperuricemia: a cross-sectional study from baseline data of ELSA-Brasil cohort. Adv Rheumatol 2022;62(1):15.

5. Abebe M, Melku M, Enawgaw B, et al. Reference intervals of routine clinical chemistry parameters among apparently healthy young adults in Amhara National Regional State, Ethiopia. PLoS One 2018;13(8): e0201782.
6. Yang Y, Jiang H, Tang A, et al. Reference intervals for serum bilirubin, urea, and uric acid in healthy Chinese geriatric population. J Clin Lab Anal 2018;32(3):e22318.
7. Guo K, Han Y, Liu S, et al. Prevalence of and trends in hyperuricemia by race and ethnicity among US adolescents, 1999-2018. Arthritis Res Ther 2024;26(1):193.
8. McNeil JJ, Woods RL, Nelson MR, et al; ASPREE Investigator Group. Effect of aspirin on disability-free survival in the healthy elderly. N Engl J Med 2018;379(16):1499–1508.
9. International Organization for Standardization. Medical laboratories — requirements for quality and competence. December 2022. Accessed November 3, 2025. https://www.iso.org/standard/76677.html
10. McNeil JJ, Wolfe R, Woods RL, et al; ASPREE Investigator Group. Effect of aspirin on cardiovascular events and bleeding in the healthy elderly. N Engl J Med 2018;379(16):1509–1518.
11. McNeil JJ, Nelson MR, Woods RL, et al; ASPREE Investigator Group. Effect of aspirin on all-cause mortality in the healthy elderly. N Engl J Med 2018;379(16):1519–1528.
12. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 4th ed. American Psychiatric Publishing, Inc.; 1994.
13. Fried LP, Tangen CM, Walston J, et al; Cardiovascular Health Study Collaborative Research Group. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001;56(3):M146–M156.
14. Lee SJ, Oh BK, Sung KC. Uric acid and cardiometabolic diseases. Clin Hypertens 2020;26(1):13.
15. Gómez-de-Tejada-Romero MJ, Murias-Henríquez C, Saavedra-Santana P, et al. Influence of serum uric acid on bone and fracture risk in postmenopausal women. Aging Clin Exp Res 2024;36(1):156.
16. Haeckel RW, Wosniok W, Torge A, et al. Age- and sex-dependent reference intervals for uric acid estimated by the truncated minimum chi-square (TMC) approach, a new indirect method. J Lab Med 2020; 44(3):157–163.

Arthritis Care & Research
Vol. 78, No. 3, March 2026, pp 417–423
DOI 10.1002/acr.25639

Perceptions About Asymptomatic Hyperuricemia and Views About Urate-Lowering Therapy in People With Asymptomatic Hyperuricemia

Nicola Dalbeth,[1] Sarah Stewart,[1] Gregory D. Gamble,[1] Borislav Mihov,[1] Lisa K. Stamp,[2] Janine Haslett,[2] William J. Taylor,[3] Tony R. Merriman,[4] Adwoa Dansoa Tabi-Amponsah,[1] Anne Horne,[1] Tuhina Neogi,[5] Tristan Pascart,[6] Mariano Andrés,[7] Maria-Luisa Peral-Garrido,[8] Eleonora Norkuviene,[9] Janitzia Vazquez-Mellado,[10] Till Uhlig,[11] Mingshu Sun,[12] Changgui Li,[12] and Keith J. Petrie[1]

Objective. Asymptomatic hyperuricemia is a precursor of gout and is also associated with cardiovascular disease and chronic kidney disease. The aim of this study was to understand perceptions about asymptomatic hyperuricemia and views about urate-lowering therapy in people with asymptomatic hyperuricemia.

Methods. Participants in a multinational study of asymptomatic hyperuricemia completed questionnaires about their perceptions of hyperuricemia, concern about hyperuricemia-associated health conditions, and willingness to take urate-lowering medication. All had a screening serum urate of ≥0.48 mmol/L (8 mg/dL) and no current or previous symptoms of gout.

Results. Overall, participants perceived that hyperuricemia had no or very few consequences on their life. Dietary factors were the most reported cause, whereas 37% did not know the cause. Participants reported a wide range in concern about hyperuricemia and the risk of developing gout. Concern about the risk of developing kidney disease or cardiovascular disease was also highly variable but was higher than concern about elevated serum urate ($P < 0.001$). Most did not think a urate-lowering medication was necessary, and there was moderate concern about the long-term use of urate-lowering medication. Medication necessity beliefs were most strongly associated with whether participants were willing to take urate-lowering medication (partial $R^2 = 0.27$, $P < 0.001$).

Conclusion. Most participants perceived minimal consequences of asymptomatic hyperuricemia. Hyperuricemia was not well understood by participants, and biologic causes were generally underrecognized. Despite a range of concerns about hyperuricemia-associated conditions, particularly kidney disease and cardiovascular disease, a urate-lowering medication for asymptomatic hyperuricemia was not considered necessary, which aligns with most current management guidelines.

INTRODUCTION

Hyperuricemia (elevated serum urate level) is the necessary precursor for the development of gout[1,2] and is also associated with cardiovascular disease and chronic kidney disease.[3,4] Many factors contribute to serum urate concentrations, including genetic factors, kidney function, body mass index, dietary factors, and medications.[5,6] There is a strong concentration-dependent

Supported by the Health Research Council of New Zealand (grant 19/232). Dr Stewarts's work was supported by the Auckland Medical Research Foundation Postdoctoral Fellowship (grant 1318001). Dr Neogi's work was supported by National Institute of Arthritis and Musculoskeletal and Skin Diseases, NIH (grants K24-AR-070892 and P30-AR-072571). Dr Pascart's work was supported by funding from Horizon Pharmaceuticals (Amgen) for the French site. Dr Sun's work was supported by an Excellent Physician Scholarship from Simcere Pharma, Nanjing, China. Dr Li's work was supported by the National Key Research and Development Program of China (2022YFC2503300).

[1]Nicola Dalbeth, MD, Sarah Stewart, PhD, Gregory D. Gamble, MSc, Borislav Mihov, BPhty, Adwoa Dansoa Tabi-Amponsah BBiomedSc(Hons), Anne Horne, MBChB, Keith J. Petrie PhD: University of Auckland, Auckland, New Zealand; [2]Lisa K. Stamp, PhD, Janine Haslett, RN: University of Otago Christchurch and Health New Zealand, Christchurch, New Zealand; [3]William J. Taylor, PhD: University of Otago Wellington, Wellington, New Zealand; [4]Tony R. Merriman, PhD: University of Otago Dunedin, Dunedin, New Zealand, and University of Alabama at Birmingham; [5]Tuhina Neogi, PhD: Boston University School of Medicine, Boston, Massachusetts; [6]Tristan Pascart, PhD: Lille Catholic University, Lille, France; [7]Mariano Andrés, PhD: Hospital General Universitario de Alicante, Alicante, Spain; [8]Maria-Luisa Peral-Garrido, MD: Vinalopo University Hospital, Elche, Spain; [9]Eleonora Norkuviene, MD: Lithuanian University of Health Sciences, Kaunas, Lithuania; [10]Janitzia Vazquez-Mellado, PhD: Hospital General de México, Mexico City, Mexico; [11]Till Uhlig, PhD: Diakonhjemmet Hospital, Oslo, Norway; [12]Mingshu Sun, MD, Changgui Li, MD: Qingdao University, Qingdao, China.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25639.

Address correspondence via email to Nicola Dalbeth, MD, at n.dalbeth@auckland.ac.nz.

Submitted for publication April 8, 2025; accepted in revised form August 13, 2025.

SIGNIFICANCE & INNOVATIONS

- Illness perceptions are key determinants of patients' behavior for managing health conditions.
- Asymptomatic hyperuricemia is common and is a precursor to gout.
- This study provides new insights into perceptions about hyperuricemia and views about urate-lowering therapy in people with asymptomatic hyperuricemia.

relationship between serum urate and development of gout, with a hazard ratio of >60 for the highest concentrations of serum urate.[2] Currently, urate-lowering therapy is not recommended for asymptomatic hyperuricemia,[7] and most people with hyperuricemia have no symptoms and do not develop gout.[2]

Illness perceptions are key determinants of patients' behavior for managing health conditions and have been associated with important outcomes in chronic diseases.[8] When faced with a new health threat, individuals will actively construct a mental representation of the illness and this guides their behavior to manage the condition, such as engagement with treatment.[9] Research suggests that it is difficult for individuals to conceive of illnesses without symptoms and consider such symptom-less illnesses to be a serious threat to health.[10,11]

Perceptions of the benefits and concerns about medication are also important factors in patients' acceptance and adherence to new treatment regimens. Perceptions of the necessity of the medicine, as well as concerns about potential adverse effects, strongly influence a person's willingness to initiate and maintain a prescribed therapy.[12] Studies suggest that individuals who view medications as essential for managing their condition are more likely to adhere, whereas those with concerns about dependency or side effects exhibit lower adherence and persistence with the treatment.[13]

Although many studies have examined patient perceptions about gout, perceptions of asymptomatic hyperuricemia have not been reported. This is a relevant question because asymptomatic hyperuricemia is common[14] and can progress to gout. The aim of this study was to understand perceptions about asymptomatic hyperuricemia and views about urate-lowering therapy in people with asymptomatic hyperuricemia.

METHODS

This was a prespecified analysis of baseline data from the Transitions in Gout Research (TIGER) study. The TIGER study is a five-year international multicenter prospective cohort study designed to understand risk factors for development of gout in people with asymptomatic hyperuricemia. The study has been registered (ACTRN12619000915156), and the full study protocol has been published.[15]

In brief, TIGER study participants were recruited from June 2019 to August 2024 at study sites in New Zealand (three sites), France, Spain, Lithuania, United States, China, and Mexico. Participants were recruited from primary and secondary care settings, by public advertising, and through community pathology collection centers. All participants had a screening serum urate of ≥0.48 mmol/L (8 mg/dL), had no current or previous clinical symptoms of gout, and were aged between 18 and 80 years. The key exclusion criteria were as follows: CKD4 or worse; other forms of inflammatory arthritis; serious illness with poor prognosis less than five years; previous synovial fluid analysis showing monosodium urate crystals; clinically evident tophi; taking urate-lowering therapy, canakinumab, or colchicine. This study was approved by the New Zealand Ministry of Health Southern Health and Disability Ethics Committee (MEC/05/10/130AM16) and by the local institutional review board for each participating center. All participants provided written informed consent.

All participants completed a baseline visit at the study site. The baseline visit included recording of demographic data, clinical risk factor assessment, and physical examination including body mass index and waist circumference. Current medications and health conditions were recorded, including those documented in the modified Rheumatic Disease Comorbidity Index (mRDCI).[16] Participants also completed 100 mm pain visual analog scales for pain over the last week and global assessment of health. Laboratory tests were taken for measurement of serum creatinine, and estimated glomerular filtration rate (eGFR) was calculated according to the 2021 Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation.[17]

At the baseline visit, participants also completed a hyperuricemia-specific Brief Illness Perceptions Questionnaire (BIPQ)[18]; because urate-lowering therapy use was an exclusion, the treatment control item in the BIPQ was not included. Participants also were asked about their concern about hyperuricemia-associated health conditions (gout, kidney disease, and cardiovascular disease; 0–10 Likert scale) and completed the Beliefs About Medicines Questionnaire[19] subscales, which assessed their views about medicines in general, and two questions about their specific views about taking medications to reduce serum urate (0–10 Likert scale): "How much do you feel you need medication to reduce your urate level?" and "How concerned are you about the long-term use of a medicine to reduce your urate level?" To assess how willing the person would be to take a medication to reduce their serum urate we asked, "How agreeable would you be to taking a medication to reduce your urate level?" (0–10 Likert scale, with 0 being not at all agreeable and 10 being extremely agreeable). All questionnaires were translated and completed in the local language.

The current analysis was a descriptive and exploratory (hypothesis generating) cross-sectional study because a specific hypothesis was not tested and inferential analysis was not the main objective. Given the low missing data rate, complete case

analysis/listwise deletion was used throughout. Data from all participating sites were entered into a central database at the coordinating site at the University of Auckland, Auckland, New Zealand. Means with SDs and percentages were used to describe the demographic and clinical characteristics of participants. Medians with interquartile ranges (IQRs) and percentages were used to describe the results of the questionnaires. "Agreeable to take a medication to reduce serum urate" scores were modeled. Variables were selected for inclusion in the model using external clinical judgment based upon previous experience with each of the instruments to reduce the risk of overfitting and to avoid a completely *P* value–driven model selection process. A further screening process to determine those variables with significant Spearman correlation with the prespecified dependent variable was performed to reduce the number of variables in the model, and iterative stepwise regression was then performed to optimize the order of inclusion of variables in the model because some collinearity was anticipated. Diagnostic plots, including plots of residual, normal probability plots and leverage plots were produced using the REG procedure of SAS and inspected to establish the goodness of fit of the model and that the residuals were normally distributed. Multicollinearity was assessed using the variance inflation factor. A variance inflation factor >5 indicated multicollinearity.[20] The final model was reviewed for clinical and scientific relevance to determine whether post hoc inclusion of additional variables in the model was required, and the final model was selected on the basis of clinical relevance, parsimony, and goodness of fit. Standardized beta coefficients were tabulated to enable comparison between variables. General linear modeling was used to compare concern between hyperuricemia-associated disease conditions. A significant main effect was reduced to pairwise comparisons with *P* adjusted using the method of Tukey to preserve an overall 5% significance level. The procedures of SAS (version 9.04; SAS Institute Inc) were used for these analyses.

Table 1. Clinical features of participants*

Clinical feature (n = 268)	Mean (±SD) or n (%)
Age, y	48 (±18)
Male	218 (81)
Ethnicity	
African/Black	5 (1.9)
East Asian	65 (24.3)
Hispanic	10 (3.7)
Māori	11 (4.1)
Pacific Islander	19 (7.1)
South Asian	36 (13.4)
White	112 (41.8)
Other Indigenous	1 (0.4)
Other/unknown	9 (3.4)
Country of residence	
New Zealand	144 (53.7)
China	64 (23.9)
France	21 (7.8)
Spain	21 (7.8)
Mexico	10 (3.7)
Lithuania	8 (3.0)
Hypertension	109 (40.7)
Cardiovascular disease[a]	58 (22)
Type 2 diabetes	34 (12.7)
eGFR < 60 mL/minute/1.73 m^2	38 (15)
eGFR, mL/minute/1.73 m^2	90.4 (±25.2)
Modified Rheumatic Disease Comorbidity Index score	2.38 (±2.20)
Units of alcohol/day	4.2 (±7.6)
Units of beer/day	1.9 (±5.2)
Aspirin use	23 (8.6)
Diuretic use	44 (16.4)
Body mass index, kg/m^2	30.5 (±6.5)
Waist circumference, cm	103.3 (±15.6)
Pain score, 100-mm VAS	12.0 (±21.4)
Patient global assessment of health, 100-mm VAS	81.1 (±16.3)
Serum urate, mmol/L	0.52 (±0.04)
Serum creatinine, μmol/L	91.4 (±26.3)

* eGFR, estimated glomerular filatration rate; VAS, visual analog scale.
[a] Included are chronic heart failure, peripheral vascular disease, myocardial revascularization, peripheral artery surgery revascularization (including aortic aneurysm), cardiac arrhythmia (atrial fibrillation/flutter), angina, myocardial infarction, and stroke.

RESULTS

Participant characteristics. The cohort included 268 people with asymptomatic hyperuricemia who completed the BIPQ. None of the participants had a history of arthritis or clinical evidence of synovitis or subcutaneous tophi at the time of the baseline study visit. The clinical features of these individuals are shown in Table 1. Participants were predominantly men, with mean age 48 years and mean body mass index 30.5 kg/m^2. Pain scores were low, and global assessment of health was high. Mean serum urate concentration was 0.52 mmol/L (8.7 mg/dL).

Perceptions about hyperuricemia. Overall, participants perceived that hyperuricemia had no or very little consequences on their life, and the presence of hyperuricemia had little emotional impact (Table 2). There was wide variation in participants' perceptions of chronicity and personal control and concern about elevated serum urate levels. Participants tended to report low understanding about hyperuricemia. Dietary factors were the most reported factor believed to cause elevated serum urate levels, whereas 37% described that the cause of their hyperuricemia as unknown. Genetic factors were believed to be the most important cause in 8.9% of participants.

Consistent with concern about elevated serum urate levels, participants reported a wide range in concern about the risk of developing gout (Table 2). There was similar concern about hyperuricemia and the risk of developing gout (post hoc Tukey $P = 0.11$). Concern about the risk of developing kidney disease or cardiovascular disease was also highly variable but was higher than concern about hyperuricemia (post hoc Tukey $P < 0.001$ for both kidney disease and cardiovascular disease).

Table 2. Perceptions about hyperuricemia, beliefs about medicines in general, and views about medications to reduce serum urate*

Characteristic	Median (25th–75th percentile range)
Brief Hyperuricemia Perception Questionnaire	
Consequences (10 = severely affects life)	0 (0–2)
Timeline (10 = will continue forever)	4 (2–7)
Personal control (10 = extreme amount)	5 (2–6)
Symptoms (10 = severe symptoms)	0 (0–1)
Concern (10 = extremely concerned)	5 (2–7)
Understanding (10 = very clearly)	4 (1–7)
Emotional response (10 = extremely affected)	1 (0–3)
Most important factor believed to cause hyperuricemia	
Dietary factors, n (%)	102 (38)
Unknown, n (%)	100 (37)
Genetic factors, n (%)	24 (8.9)
Overweight, n (%)	18 (6.7)
Medications, n (%)	7 (2.6)
Kidney disease, n (%)	7 (2.6)
Other health condition, n (%)	7 (2.6)
Lack of exercise, n (%)	3 (1.1)
Other, n (%)	2 (0.7)
Concern about hyperuricemia-associated health conditions	
Concern about risk of developing gout (10 = extremely concerned)	5 (3–8)
Concern about risk of developing kidney disease (10 = extremely concerned)	6 (3–9)
Concern about risk of developing cardiovascular disease (10 = extremely concerned)	6 (3–9)
Beliefs About Medicines Questionnaire—general subscales	
Harm subscale (maximum score 20)	9.5 (9–12)
Overuse subscale (maximum score 20)	12 (10–14)
Benefit subscale (maximum score 20)[a]	16 (15–18)
Views about using medication to reduce serum urate	
Need to take medication to reduce serum urate (10 = essential)	2 (0–5)
Agreeable to take a medication to reduce serum urate (10 = extremely agreeable)[b]	5 (2–8)
Concern about the long-term use of a medication to reduce serum urate (10 = extremely concerned)	6 (2–8)

* [a] The benefit subscale is missing beliefs about medicines benefit questions from one site (n = 247 for this variable).
[b] Responses were available from 258 participants.

Views about medications to reduce serum urate levels. Generally, participants viewed medications positively (Table 2). Most did not think they needed medication to reduce their serum urate, and there was moderate concern about long-term medication use. There was a highly variable response about whether they were agreeable to take a medication to reduce the serum urate with median (IQR) score 5 (2–8; Figure 1).

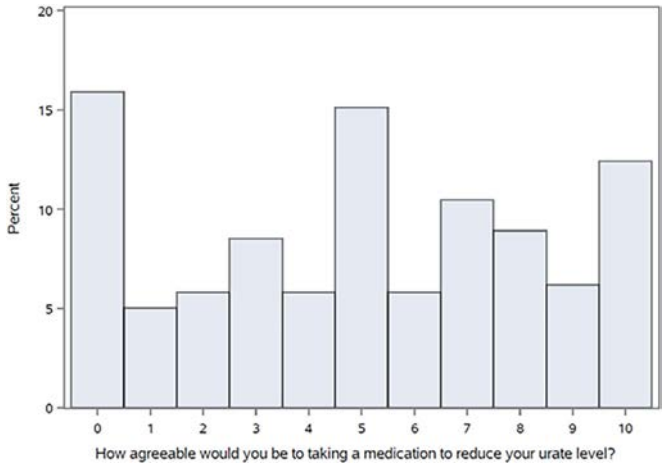

Figure 1. Distribution of agreeable to taking a medication to reduce serum urate scores (0 being not at all agreeable and 10 being extremely agreeable).

In bivariable analysis, willingness to take a medication to reduce the serum urate associated with concern about hyperuricemia and risk of developing hyperuricemia-associated health conditions (gout, kidney disease, and cardiovascular disease), perceived need to take a medication to reduce the serum urate, beliefs about medicines, and comorbidities (Table 3). In multivariable analysis, the perceived need for medication to reduce serum urate was most strongly associated with whether participants were agreeable to take urate-lowering medication (R^2 = 0.27, $P < 0.001$; Table 4). In addition, beliefs about the harms and benefits of medicines, the mRDCI, and concern about hyperuricemia were significantly associated (model cumulative R^2 = 0.38, $P < 0.001$). Sex or geographic region was not associated with willingness to take urate-lowering medications.

DISCUSSION

This study provides novel insights into the views and experience of people with asymptomatic hyperuricemia. Consistent with an asymptomatic condition, participants reported low scores for the consequences, symptoms, and emotional response domains of the hyperuricemia-specific BIPQ. There was a broad range of concern about hyperuricemia and hyperuricemia-associated conditions including gout, kidney disease, and cardiovascular disease. Participants did not consider that a urate-lowering medication was needed, which is consistent with most current international management guidelines.

As reported in our prior studies of gout, dietary factors were considered the most important factor causing asymptomatic hyperuricemia.[21–24] More than one-third of participants did not know the cause, and fewer than 10% of participants reported that genetic factors were the most important factor. Biologic causes such as genetic variants, kidney function, and body composition

Table 3. Correlations between being agreeable to take a medication to reduce serum urate scores and other variables*

Variables	Spearman r (95% CI)	*P* value
Age, y	0.08 (−0.04 to 0.20)	0.18
Male sex	0.04 (−0.09 to 0.16)	0.56
Serum urate level	0.03 (−0.15 to 0.10)	0.68
Modified Rheumatic Disease Comorbidity Index	0.24 (0.12 to 0.35)	<0.0001
Brief Hyperuricemia Perception Questionnaire		
Consequences (10 = severely affects life)	0.02 (−0.10 to 0.14)	0.74
Timeline (10 = will continue forever)	0.06 (−0.07 to 0.18)	0.37
Personal control (10 = extreme amount)	−0.002 (−0.13 to 0.12)	0.98
Symptoms (10 = severe symptoms)	0.05 (−0.07 to 0.18)	0.37
Concern (10 = extremely concerned)	0.17 (0.05 to 0.29)	0.0061
Understanding (10 = very clearly)	0.08 (−0.04 to 0.20)	0.21
Emotional response (10 = extremely affected)	0.04 (−0.08 to 0.16)	0.51
Concern about hyperuricemia-associated health conditions		
Concern about risk of developing gout (10 = extremely concerned)	0.22 (0.09 to 0.33)	0.0006
Concern about risk of developing kidney disease (10 = extremely concerned)	0.22 (0.11 to 0.34)	0.0003
Concern about risk of developing cardiovascular disease (10 = extremely concerned)	0.24 (0.12 to 0.35)	0.0001
Beliefs About Medicines Questionnaire—general subscales		
Harm subscale (maximum score 20)	−0.28 (−0.39 to −0.16)	<0.0001
Overuse subscale (maximum score 20)	−0.15 (−0.28 to −0.03)	0.015
Benefit subscale (maximum score 20)	0.30 (0.18 to 0.41)	<0.0001
Views about medications to reduce serum urate		
Need to take medication to reduce serum urate (10 = essential)	0.45 (0.34 to 0.54)	<0.0001
Concern about the long-term use of a medication to reduce serum urate (10 = extremely concerned)	0.02 (−0.11 to 0.14)	0.80

* CI, confidence interval.

were generally underrecognized in contrast to the scientific literature, which shows that these factors play a central role in serum urate regulation.[5,6,25,26] These findings highlight the ongoing gap of translating scientific discoveries to community understanding about hyperuricemia and gout.

The study provides insights into personal decisions about taking medications, particularly for an asymptomatic laboratory finding. Our findings indicate the importance of patient perceptions about the need for medication. In the context of asymptomatic hyperuricemia, there is some evidence that incident gout can be prevented with urate-lowering therapy,[27] but there is no compelling evidence for prevention of kidney disease progression[28–30] or cardiovascular disease.[27] Study participants reported low necessity beliefs about urate-lowering medication, which is consistent with current evidence and management guidelines. Collectively, these data indicate that much more compelling evidence would be needed for people with asymptomatic hyperuricemia to adopt urate-lowering medication.

Table 4. Stepwise linear regression model to determine variables associated with being agreeable to taking a medication to reduce serum urate scores*

Variable	ß coefficient	SE	Standardized ß coefficient	95% confidence intervals	t value	*P* value	Partial R^2	Cumulative R^2	VIF
Perceived need for medication to reduce urate	0.497	0.066	0.419	0.366 to 0.627	7.50	<0.0001	0.270	0.270	1.152
BMQ-harm	−0.227	0.064	−0.202	−0.353 to −0.101	−3.55	0.0005	0.059	0.329	1.193
Modified Rheumatic Disease Comorbidity Index	0.205	0.084	0.132	0.043 to 0.370	2.45	0.015	0.017	0.346	1.069
BMQ-benefit	0.190	0.081	0.137	0.030 to 0.350	2.35	0.020	0.012	0.358	1.250
Brief Hyperuricemia Perception Questionnaire-concern	0.157	0.058	0.149	0.041 to 0.272	2.68	0.008	0.019	0.377	1.145

* All variables with $P < 0.05$ in the bivariate correlation analysis shown in Table 3 were included in the model. The following variables were included in this model: modified Rheumatic Disease Comorbidity Index, Brief Hyperuricemia Perception Questionnaire-concern, concern about gout risk, concern about kidney disease risk, concern about heart disease risk, BMQ-harm, BMQ-overuse, BMQ-benefit, and perceived need for medication to reduce urate. BMQ, Beliefs About Medicines Questionnaire. VIF, variance inflation factor.

Our findings about willingness to take urate-lowering therapy for asymptomatic hyperuricemia align with numerous studies that reported that beliefs about the necessity of treatment is a strong predictor of medication adherence in long-term symptomatic health conditions.[13] The results suggest that it would be difficult for individuals to take long-term medication for asymptomatic hyperuricemia because the condition elicits a low level of concern and perceived consequences. Moreover, the treatment is seen as not essential, and people with asymptomatic hyperuricemia have moderate concern for taking medication for the condition long term.

Study strengths include the well-characterized multinational cohort, with consistent phenotyping and data collection. It should be noted that the study findings may not be generalizable to all people with asymptomatic hyperuricemia because the study participants were volunteers in a cohort study of asymptomatic hyperuricemia and may have had different perceptions and health concerns to those not volunteering for the study. Additionally, this was a cross-sectional study, and hypothetical views about taking a medication in the context of a research questionnaire may not translate to medication adherence in real life.

In summary, most people with asymptomatic hyperuricemia report no or very few self-identified consequences of this laboratory finding. Asymptomatic hyperuricemia is not well understood, and biologic causes are generally underrecognized by people with asymptomatic hyperuricemia. Despite a range of concerns about hyperuricemia-associated conditions, particularly kidney disease and cardiovascular disease, a urate-lowering medication was not considered necessary by individuals with asymptomatic hyperuricemia, which is in line with most of the current management guidelines regarding asymptomatic hyperuricemia.

ACKNOWLEDGMENTS

Open access publishing facilitated by The University of Auckland, as part of the Wiley - The University of Auckland agreement via the Council of Australian University Librarians.

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 Dalbeth 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

The Health Research Council of New Zealand had no role in the study design or in the collection, analysis, or interpretation of the data or the decision to submit the manuscript for publication. Publication of this article was not contingent upon approval by Health Research Council of New Zealand.

REFERENCES

1. Campion EW, Glynn RJ, DeLabry LO. Asymptomatic hyperuricemia. Risks and consequences in the Normative Aging Study. Am J Med 1987;82:421–426.
2. Dalbeth N, Phipps-Green A, Frampton C, et al. Relationship between serum urate concentration and clinically evident incident gout: an individual participant data analysis. Ann Rheum Dis 2018;77:1048–1052.
3. Li L, Yang C, Zhao Y, et al. Is hyperuricemia an independent risk factor for new-onset chronic kidney disease?: a systematic review and meta-analysis based on observational cohort studies. BMC Nephrol 2014;15:122.
4. Li M, Hu X, Fan Y, et al. Hyperuricemia and the risk for coronary heart disease morbidity and mortality a systematic review and dose-response meta-analysis. Sci Rep 2016;6:19520.
5. Major TJ, Topless RK, Dalbeth N, et al. Evaluation of the diet wide contribution to serum urate levels: meta-analysis of population based cohorts. BMJ 2018;363:k3951.
6. Topless RKG, Major TJ, Florez JC, et al. The comparative effect of exposure to various risk factors on the risk of hyperuricaemia: diet has a weak causal effect. Arthritis Res Ther 2021;23:75.
7. FitzGerald JD, Dalbeth N, Mikuls T, et al. 2020 American College of Rheumatology Guideline for the Management of Gout. Arthritis Rheumatol 2020;72:879–895. Erratum in: *Arthritis Rheumatol* 2021; 73:413.
8. Petrie KJ, Weinman J. Patients' perceptions of their illness: the dynamo of volition in health care. Curr Dir Psychol Sci 2012;21:60–65.
9. Leventhal H, Meyer D, Nerenz D. The common sense representation of illness danger. In: Rachman S, ed. Contributions to Medical Psychology. Pergamon Press; 1980:17–30.
10. Halm EA, Mora P, Leventhal H. No symptoms, no asthma: the acute episodic disease belief is associated with poor self-management among inner-city adults with persistent asthma. Chest 2006;129: 573–580.
11. Petrie KJ, Weinman J. Why illness perceptions matter. Clin Med (Lond) 2006;6:536–539.
12. Horne R, Weinman J. Patients' beliefs about prescribed medicines and their role in adherence to treatment in chronic physical illness. J Psychosom Res 1999;47:555–567.
13. Horne R, Chapman SC, Parham R, et al. Understanding patients' adherence-related beliefs about medicines prescribed for long-term conditions: a meta-analytic review of the Necessity-Concerns Framework. PLoS One 2013;8:e80633.
14. Chen-Xu M, Yokose C, Rai SK, et al. Contemporary prevalence of gout and hyperuricemia in the United States and decadal trends: the National Health and Nutrition Examination Survey, 2007-2016. Arthritis Rheumatol 2019;71:991–999.
15. Stewart S, Gamble G, Taylor W, et al. Development of gout in people with asymptomatic hyperuricemia: study protocol for a 5-year prospective cohort. BMJ Open 2024;14:e090415.
16. Spaetgens B, Wijnands JMA, van Durme C, et al. Content and construct validity of the Rheumatic Diseases Comorbidity Index in patients with gout. Rheumatology (Oxford) 2015;54:1659–1663.
17. Inker LA, Eneanya ND, Coresh J, et al; Chronic Kidney Disease Epidemiology Collaboration. New creatinine- and cystatin C-based equations to estimate GFR without race. N Engl J Med 2021;385:1737–1749.
18. Broadbent E, Petrie KJ, Main J, et al. The brief illness perception questionnaire. J Psychosom Res 2006;60:631–637.

19. Horne R, Weinman J, Hankins M. The Beliefs About Medicines Questionnaire: the development and evaluation of a new method for assessing the cognitive representation of medication. Psychol Health 1999;14:1–24.
20. Sheather S. A Modern Approach to Regression with R. Springer Nature; 2009.
21. Dalbeth N, Petrie KJ, House M, et al. Illness perceptions in patients with gout and the relationship with progression of musculoskeletal disability. Arthritis Care Res (Hoboken) 2011;63:1605–1612.
22. Douglas M, Yielder R, Kleinstäuber M, et al. The big picture: patient drawings of gout and their relationship to illness perceptions and stigma. J Rheumatol 2024;51:203–205.
23. Murdoch R, Mihov B, Horne AM, et al. Impact of television depictions of gout on perceptions of illness: a randomized controlled trial. Arthritis Care Res (Hoboken) 2023;75:2151–2157.
24. Petrie KJ, MacKrill K, Derksen C, et al. An illness by any other name: the effect of renaming gout on illness and treatment perceptions. Health Psychol 2018;37:37–41.
25. Tin A, Marten J, Halperin Kuhns VL, et al. Target genes, variants, tissues and transcriptional pathways influencing human serum urate levels. Nat Genet 2019;51:1459–1474.
26. Tin A, Schlosser P, Matias-Garcia PR, et al; Estonian Biobank Research Team; Genetics of DNA Methylation Consortium. Epigenome-wide association study of serum urate reveals insights into urate co-regulation and the SLC2A9 locus. Nat Commun 2021;12:7173.
27. Mackenzie IS, Hawkey CJ, Ford I, et al; ALL-HEART Study Group. Allopurinol versus usual care in UK patients with ischaemic heart disease (ALL-HEART): a multicentre, prospective, randomised, open-label, blinded-endpoint trial. Lancet 2022;400:1195–1205.
28. Heerspink HJL, Stack AG, Terkeltaub R, et al; SAPPHIRE Investigators. Combination treatment with verinurad and allopurinol in CKD: a randomized placebo and active controlled trial. J Am Soc Nephrol 2024;35:594–606.
29. Doria A, Galecki AT, Spino C, et al; PERL Study Group. Serum urate lowering with allopurinol and kidney function in type 1 diabetes. N Engl J Med 2020;382:2493–2503.
30. Badve SV, Pascoe EM, Tiku A, et al; CKD-FIX Study Investigators. Effects of allopurinol on the progression of chronic kidney disease. N Engl J Med 2020;382:2504–2513.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 78, No. 3, March 2026, pp 424–426

LETTER

DOI 10.1002/acr.25646

Reframing pain modulation and disease-modifying antirheumatic drugs de-escalation: lessons from a lifestyle intervention in rheumatoid arthritis and osteoarthritis. Comment on the article by Wagenaar et al

To the Editor:

We read with great interest the study by Wagenaar et al reporting the two-year follow-up of the Plants for Joints (PFJ) lifestyle intervention for patients with rheumatoid arthritis (RA) and metabolic syndrome–associated osteoarthritis (MSOA).[1] Their findings of sustained improvements in disease activity, symptom burden, and metabolic markers following a structured, plant-based, multidisciplinary intervention are timely and highly relevant to clinicians managing chronic musculoskeletal pain. From the vantage point of a physician routinely caring for patients with persistent pain, we would like to offer several reflections on the clinical and mechanistic implications of these findings, particularly in the context of long-term pain modulation and treatment sustainability.

First, the role of central pain sensitization and its potential modulation through lifestyle interventions warrants deeper attention. The observed long-term reduction in Western Ontario and McMaster Universities Osteoarthritis Index pain scores in participants with MSOA, maintained even with decreased use of analgesics, raises the question of whether the PFJ intervention achieved more than peripheral anti-inflammatory or biomechanical benefits. Chronic osteoarthritis pain is increasingly recognized to involve central sensitization, with overlapping mechanisms seen in fibromyalgia and other chronic pain states.[2–4] Given the intervention's inclusion of stress reduction and physical activity—both known to influence central pain processing—future iterations of the study would benefit from incorporating validated measures of central sensitization, such as the Central Sensitization Inventory or quantitative sensory testing. This would not only refine the understanding of mechanisms but could also better stratify responders to lifestyle-based pain interventions.

Second, the implications of reduced pharmacologic dependency merit greater exploration from both efficacy and safety perspectives. In the RA cohort, 44% of patients reduced or discontinued the use of disease-modifying antirheumatic drugs (DMARDs) during follow-up yet maintained or improved disease control. Although these findings are encouraging, the lack of detailed reporting on flare rates or time to disease worsening limits our ability to interpret whether these reductions were clinically appropriate or potentially risky. From a pain clinician's standpoint, the decision to taper medications must be contextualized with flare vigilance and the patient's functional trajectory. We suggest that future research adopt longitudinal disease activity trajectories and flare incidence as coprimary outcomes, to evaluate whether lifestyle interventions can indeed enable safe DMARD de-escalation without compromising long-term joint integrity or function.

Third, the PFJ study provides a valuable platform to consider how nonpharmacologic interventions can be embedded into routine multidisciplinary pain care, but implementation barriers should not be underestimated. Although the intervention demonstrated impressive adherence over two years, this was likely aided by structured group meetings, newsletters, and webinars. However, most rheumatology or pain clinics lack the infrastructure for sustained behavioral support, particularly in underserved populations. Translating PFJ into routine care thus requires simplification, decentralization (eg, community health worker delivery), and digital augmentation (eg, mobile platforms). Understanding which components (dietary change, physical activity, and stress management) are most essential for clinical effect could inform leaner, scalable models of care, especially in resource-limited settings.

In summary, this study makes a compelling case for integrating lifestyle interventions into the long-term management of RA and MSOA, particularly for pain reduction and medication minimization. However, further granularity regarding pain mechanisms, flare dynamics, and pragmatic delivery models will be critical for translating these insights into durable clinical practice.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25646.

Yujie Xu, MM
Hua Xu, MD
pshhuaxu@163.com
Yueyang Hospital of Integrated Traditional Chinese and Western Medicine, Shanghai University of Traditional Chinese Medicine
Shanghai, China

1. Wagenaar CA, Walrabenstein W, van der Leeden M, et al. Two-year follow-up of a multidisciplinary lifestyle intervention for rheumatoid arthritis and osteoarthritis. Arthritis Care Res (Hoboken) 2025;77(9):1141–1148.
2. López-Ruiz M, Losilla JM, Monfort J, et al. Central sensitization in knee osteoarthritis and fibromyalgia: beyond depression and anxiety. PLOS One 2019;14(12):e0225836.
3. Staud R. Evidence for shared pain mechanisms in osteoarthritis, low back pain, and fibromyalgia. Curr Rheumatol Rep 2011;13(6):513–520.
4. Lee YC, Nassikas NJ, Clauw DJ. The role of the central nervous system in the generation and maintenance of chronic pain in rheumatoid arthritis, osteoarthritis and fibromyalgia. Arthritis Res Ther 2011;13(2):211.

DOI 10.1002/acr.25645

Reply

To the Editor:

We thank Xu and Xu for their thoughtful response and for highlighting important considerations not fully addressed in our article, "Two-year follow-up of a multidisciplinary lifestyle intervention for rheumatoid arthritis and osteoarthritis."[1]

First, the authors raise an insightful point regarding the potential role of central sensitization in the observed pain relief, particularly in metabolic syndrome–associated osteoarthritis (MSOA). Our hypothesis was that the intervention would primarily target chronic inflammation, an important driver in both rheumatoid arthritis (RA) and MSOA.[2] The improvements in clinical outcomes support this hypothesis. Interestingly, because our intervention targeted the entire body, including dietary changes, physical activity, and stress reduction, it may have influenced central sensitization as well. Chronic inflammation is known to impact central pain processing by activating glial cells in the central nervous system, which release proinflammatory mediators that enhance neuronal excitability and synaptic transmission.[3] This process contributes to the amplification and persistence of pain and may consequently have been influenced by our intervention. To further investigate these mechanisms, we agree future studies would benefit from incorporating validated tools to assess central sensitization.

Second, the authors address the clinically important topic of disease-modifying antirheumatic drug (DMARD) reduction in people with RA. There is ample evidence that DMARDs can be safely tapered in patients with RA in sustained remission, provided they are monitored for disease flare. In our study, a protocol was in place to reduce DMARDs stepwise during remission and escalate treatment if disease activity, as measured by the Disease Activity Score in 28 joints, subsequently increased. Such increases only occurred a few times. The modest net reduction in DMARD use, along with sustained low disease activity, suggests that lifestyle interventions can support safe medication tapering without compromising long-term joint health.

Third, we fully agree that long-term adherence is critical for the success of any lifestyle intervention. Encouragingly, adherence remained high during the two-year follow-up, supported by relatively modest efforts such as monthly newsletters and bimonthly webinars. We believe this was because of the high motivation of participants, the perceived effectiveness, and the clear emphasis from the outset that the intervention was intended to support lifelong lifestyle change.[4] Nevertheless, translating this intervention into real-world clinical settings requires further adaptation. Drawing from our process evaluation, we identified key working elements and areas for improvement, which informed the development of a more scalable version of the Plants for Joints program.[4] This optimized version includes more individual guidance, digital education and peer support, and a hybrid format combining live and online group sessions, all while retaining the multidisciplinary approach.[5] Importantly, the revised program now offers flexible pathways: some individuals may benefit from comprehensive group support, whereas others may prefer to engage independently through educational resources or one-on-one coaching.[5] This flexibility enhances accessibility and reach. Ultimately, broader integration into health care systems and insurance coverage for such interventions will be essential to ensure equitable access and sustainable impact.

Once again, we thank Xu and Xu for sharing their insightful commentary and expertise. We agree greater granularity on pain mechanisms and practical models for implementation will help to advance the integration of lifestyle medicine in routine care.

Carlijn A. Wagenaar, MD
c.a.wagenaar@amsterdamumc.nl
Dirkjan van Schaardenburg, MD, PhD
Reade Center for Rheumatology and Rehabilitation, Amsterdam University Medical Centers, Universiteit van Amsterdam and Amsterdam Rheumatology and Immunology Center
Amerstand, the Netherlands
Wendy Walrabenstein, PhD
Amsterdam University Medical Centers, Universiteit van Amsterdam
Amsterdam, the Netherlands
Marike van der Leeden, PhD
Amsterdam University Medical Centers, Vrije Universiteit and Amsterdam Movement Sciences Research Institute
Amsterdam, the Netherlands
Franktien Turkstra, MD, PhD
Martijn Gerritsen, MD, PhD
Reade Center for Rheumatology and Rehabilitation
Amsterdam, the Netherlands
Jos W. R. Twisk, PhD
Maarten Boers, MD, PhD
Amsterdam University Medical Centers, Vrije Universiteit
Amsterdam, the Netherlands
Martin van der Esch, PhD
Reade Center for Rheumatology and Rehabilitation and Amsterdam University of Applied Sciences
Amsterdam, the Netherlands
Henriët van Middendorp
Leiden University
Leiden, the Netherlands
Peter J. M. Weijs, PhD
Amsterdam University Medical Centers, Vrije Universiteit and Amsterdam University of Applied Sciences
Amsterdam, the Netherlands

1. Wagenaar CA, Walrabenstein W, van der Leeden M, et al. Two-year follow-up of a multidisciplinary lifestyle intervention for rheumatoid arthritis and osteoarthritis. Arthritis Care Res (Hoboken) 2025;77(9):1141–1148.

2. Furman D, Campisi J, Verdin E, et al. Chronic inflammation in the etiology of disease across the life span. Nat Med 2019;25(12):1822–1832.

3. Ji RR, Nackley A, Huh Y, et al. Neuroinflammation and central sensitization in chronic and widespread pain. Anesthesiology 2018;129(2):343–366.
4. Wagenaar CA, Toonstra A, Walrabenstein W, et al. How the Plants for Joints multidisciplinary lifestyle intervention achieved its effects: a mixed methods process evaluation. BMC Public Health 2024;24(1):1034–1032.
5. Plants for Health. Accessed August 10, 2025. https://www.plants-for-health.com.

Arthritis Care & Research
Vol. 78, No. 3, March 2026, pp 427
DOI 10.1002/acr.25585

American College of Rheumatology
Empowering Rheumatology Professionals

Correction to "Effects of Social Vulnerability and Environmental Burden on Care Fragmentation and Social Needs Among Individuals With Rheumatic Conditions"

Santacroce L, Yang S, Summit R, et al. *Effects of social vulnerability and environmental burden on care fragmentation and social needs among individuals with rheumatic conditions*. Arthritis Care Res (Hoboken) 2025 Jan;77(1):116–126.

In Tables 2 and 3, the footnote "Bolded values signify statistical significance at $P > 0.05$" should instead read "Bolded values signify statistical significance at $P < 0.05$."

We apologize for this error.