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 77 No. 7 July 2025

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 77 • JULY 2025 • NO. 7

Other and Mixed Rheumatic Diseases

Letter

Cover image: The image on the cover (Roberts et al; pages 837–847) shows distribution of the estimated years lived with disability due to low back pain in 2021. Source: Institute for Health Metrics and Evaluation. Used with permission.

Arthritis Care & Research
Vol. 77, No. 7, July 2025, pp 813–826
DOI 10.1002/acr.25512

American College of Rheumatology
Empowering Rheumatology Professionals

Sex-Related Differences in Efficacy and Safety Outcomes in Axial Spondyloarthritis Randomized Clinical Trials: A Systematic Literature Review and Meta-Analysis

Angel Gao,[1] Jordi Pardo Pardo,[2] Steven Dang,[3] Lianne S. Gensler,[4] Philip Mease,[5] and Lihi Eder[3]

Objective. We aimed to assess differences in baseline characteristics, efficacy, and safety of advanced therapies between male and female patients with axial spondyloarthritis (axSpA) in randomized controlled trials (RCTs).

Methods. We conducted a systematic literature search for RCTs assessing the efficacy of advanced therapies in patients with axSpA until March 19, 2023. We extracted the following outcomes by sex: baseline participant characteristics, Assessment in Spondylarthritis International Society (ASAS) 20/40 criteria, and Axial Spondyloarthritis Disease Activity Score low disease activity or inactive disease (ASDAS-LDA/ID). Random-effects models were used to calculate pooled effects for responses in men versus women for different medication classes.

Results. We included 79 RCTs (n = 23,748 patients, 69.7% male). Only 9 trials (11.4%), 22 trials (28%), and 9 trials (11.4%) reported baseline characteristics, efficacy end points, and safety end points by sex, respectively. At baseline, women were significantly older and had higher pain scores, whereas men had higher C-reactive protein levels. Overall, male patients were more likely to achieve an ASAS40 response compared to female patients for all advanced therapies (odds ratio [OR] 1.88, 95% confidence interval [CI] 1.44–2.46) and for interleukin-17A (IL-17A) inhibitors (IL-17Ai) (OR 1.82) and tumor necrosis factor inhibitor (TNFi) (OR 2.42), and male patients had numerically higher values for IL-17A/Fi. Male patients were also more likely to achieve an ASDAS-LDA/ID (OR 2.19, 95% CI 1.47–3.26) across all advanced therapies and for IL-17Ai (OR 2.08) and TNFi (OR 2.42) individually.

Conclusion. Female patients with axSpA are less likely to achieve efficacy outcomes on advanced therapies compared to their male counterparts, with similar differences across medication classes. Future studies should study the biologic (sex-related) and sociocultural (gender-related) mechanisms underlying these differences.

INTRODUCTION

Axial spondylarthritis (axSpA) is characterized by axial inflammation accompanied by peripheral musculoskeletal and extramusculoskeletal manifestations.[1,2] Historically, axSpA was perceived to predominately affect men.[3] Recent research shows a male-to-female ratio ranging from 2:1 to 3:1 for radiographic axSpA (r-axSpA).[2,4,5] However, nonradiographic axSpA (nr-axSpA) shows a more balanced distribution between the sexes.[2]

AxSpA manifests differently in men and women. Male patients tend to develop more severe radiographic spinal damage and have higher C-reactive protein (CRP) levels,[6] whereas female patients consistently report higher pain scores and poorer health-related quality of life.[7–9]

The emergence of advanced therapies directed at various disease-specific inflammatory pathways has significantly improved patient outcomes.[10,11] However, observational studies indicate that female patients with axSpA have a significantly lower treatment

Dr Eder's work was supported by the Canada Research Chair in Inflammatory Rheumatic Diseases (Tier 2).

[1]Angel Gao, BHSc: Queen's University, Kingston, Ontario, Canada; [2]Jordi Pardo Pardo, BCom: University of Ottawa, Ottawa, Ontario, Canada; [3]Steven Dang, BSc, MLT, Lihi Eder, MD, PhD: Women's College Hospital, University of Toronto, Toronto, Ontario, Canada; [4]Lianne S. Gensler, MD: University of California, San Francisco; [5]Philip Mease, MD: Providence Swedish Medical Center and University of Washington, Seattle.

Additional supplementary information cited in this article can be found online in the Supporting Information section (http://onlinelibrary.wiley.com/doi/10.1002/acr.25512).

Author disclosures and graphical abstract are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25512.

Address correspondence via email to Lihi Eder, MD, PhD, at Lihi.Eder@wchospital.ca.

Submitted for publication November 12, 2024; accepted in revised form January 21, 2025.

SIGNIFICANCE & INNOVATIONS

- Our study results revealed that female patients with axial spondyloarthritis (axSpA) participating in clinical trials are less likely to achieve efficacy end points compared to their male patient counterparts across both tumor necrosis factor and interleukin-17 inhibitors.
- Some of these sex-related differences in treatment outcomes may stem from a higher burden of symptoms at baseline among female participants and differences in SpA features (radiographic vs nonradiographic).
- This study improves our understanding of how the patient's sex influences treatment outcomes in axSpA. The differential treatment response may arise from biologic, sex-related mechanisms or sociocultural, gender-related mechanisms.
- We also highlighted the inadequate reporting of sex-disaggregated data, particularly concerning safety end points. Future trials should consider sex and gender in the trial design and prioritize diligent reporting of sex-based analyses to ensure best practices.

response and are less likely to achieve low disease activity (LDA) states than their male counterparts.[7,12–14] Additionally, female patients tend to discontinue therapy prematurely and are more prone to treatment failure.[7,12,13]

Although data from observational studies are important for understanding real-world effectiveness of advanced therapies, randomized controlled trials (RCTs) offer unbiased evidence regarding the efficacy and safety of currently used therapies. However, RCTs typically do not provide sex-disaggregated results.[15] Consequently, it remains unknown whether sex differences exist in the participation and baseline characteristics of patients with axSpA in RCTs investigating advanced therapies. Differences in participation by sex could affect the generalizability and interpretation of the results, and sex-related differences in baseline characteristics could affect the probability of responding to therapy, for example, a higher body mass index (BMI) and lower CRP levels could influence response. More importantly, limited data are available on how patient sex influences treatment efficacy and safety among patients with axSpA and whether such responses vary across different drug classes. Recently, a meta-analysis of RCTs in patients with psoriatic arthritis (PsA) highlighted such differential responses across classes of advanced therapies.[16]

In this systematic literature review (SLR) and meta-analysis of RCTs of advanced therapies in axSpA, we aimed to report the following outcomes by sex: (1) participation, (2) baseline characteristics, (3) efficacy end points, and (4) safety end points. Identifying sex-related differences in treatment outcomes in interventional studies is a crucial for gaining insights into and addressing sex-based disparities in advanced therapies for this patient population.

MATERIALS AND METHODS

We conducted an SLR and meta-analysis in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement (see Supplementary Table 1).[17] This study was registered on PROSPERO (CRD42023412351). The study design followed the Population, Intervention, Comparison, Outcome questions outlined in Supplementary Table 2. Because this study is a systematic review of the literature, it did not require a review by an institutional ethics committee.

Search strategy and eligibility criteria. A search strategy was designed in consultation with a methodology expert (JPP) to find primary references in Medline, Embase, and Cochrane Central Register, from January 1, 2000, to March 19, 2023 (Supplementary Table 3). We also searched clinical trial databases (ClinicalTrials.gov) and databases of regulatory agencies including the European Medicines Agency (EMA) and US Food and Drug Administration (FDA). Furthermore, we screened the references within articles selected for this study to identify potentially relevant studies.

Two reviewers (AG and SD) independently performed abstract/title and full-text screening to identify RCTs meeting the following eligibility criteria: (1) RCTs in adult patients, (2) diagnosis of axSpA including ankylosing spondylitis and radiographic and nonradiographic axSpA, and (3) trials assessing the efficacy of an advanced therapies (biologic or targeted synthetic disease-modifying antirheumatic drug [DMARD]) compared to placebo or another DMARD. Exclusion criteria included the following: (1) cohorts, case controls, reviews, and case reports/series; (2) duration of less than 12 weeks; (3) non-English articles; (4) studies comparing treatment strategies (eg, withdrawal studies); and (5) studies comparing biosimilars to the corresponding originator biologic DMARD. Although we excluded non-English articles, we translated 10 such articles using Google Translate and confirmed that none reported sex-disaggregated results.

Data abstraction. Two reviewers (AG and SD) independently performed data abstraction in duplicate using standardized electronic extraction forms. Any discrepancies were resolved through discussion with the senior investigator (LE).

In cases in which multiple publications described results from the same RCT, we reviewed all publications, including those reporting the primary analysis of the results, as well as any additional publications reporting post hoc analyses and open-label extensions. If subgroup analysis by sex was conducted, secondary publications were included. Otherwise, only the primary publication was considered. As a result, each RCT was counted only

once in the denominator when analyzing sex-disaggregated reporting and participation.

Study outcomes. For each trial, we extracted information on the study design, intervention, and baseline patient characteristics, including the prevalence of male and female participants, as well as whether the study reported sex-disaggregated data on baseline axSpA features, efficacy, and safety outcomes. The terms "sex" and "gender" were used interchangeably in the studies reviewed to denote subgroup analyses of study outcomes in men and women. Therefore, we refer to the term "sex" throughout the article, acknowledging that the variations in study outcomes likely stem from both biologically based, sex-related factors and socioculturally influenced, gender-related factors.

When available, we extracted information on baseline patient characteristics by sex including demographics and measures of disease activity. In addition, we extracted data on the proportion of patients achieving the following efficacy end points by sex at the primary time point of the trial: Assessment of Spondyloarthritis International Society (ASAS) 20/40 response criteria, Bath Axial Spondyloarthritis Disease Activity Index (BASDAI) 50 response, ASDAS inactive disease (ID) (score < 1.3), and ASDAS LDA scores (score < 2.1). ASDAS-LDA and ASDAS-ID were analyzed together due to limited number of trials reporting each one. Lastly, we extracted data on safety end points by sex at the latest time point of the trial.

Statistical analysis. Initially, we computed the weighted pooled proportion of male and female participants in the RCTs. Subsequently, we reported the proportion of trials reporting sex-disaggregated results concerning baseline characteristics, efficacy end points, and safety end points.

Next, we conducted a random-effects meta-analysis to compare the differences in baseline study characteristics between male and female patients. We used random-effects models to combine results from individual trials.[18] The models included variables reported by sex for each study, including means and SDs for continuous variables and frequency and total number for binary variables. Sex differences in baseline patient characteristics were reported as mean differences (MDs) with 95% confidence intervals (CIs) for continuous variables and odds ratios (ORs) with 95% CIs for binary variables in male patients versus female patients.

To evaluate sex differences in efficacy end points, we used a random-effects model to compute ORs for the rates of ASAS40, ASAS20, and BASDAI50 responses and ASDAS LDA or ID (ASDAS-LDA/ID) in men versus women. These models included the frequency of the outcome and total number of participants by sex, stratified by drug class. Pooled ORs were calculated for each drug class individually and for all classes combined. In addition, we conducted a meta-analysis to estimate the pooled effects of achieving ASAS40 and ASAS20 placebo response in men versus women.

We evaluated clinical heterogeneity across studies by examining the variability in participants, interventions, and outcomes. Methodologic heterogeneity was assessed by exploring differences in study design and risk of bias. Additionally, we investigated the presence of substantial heterogeneity in end point results through visual inspection of the forest plot and quantitative analysis using χ^2 statistics (Q statistics) and I^2 statistics.[19] A value of I^2 greater than 40% was considered indicative of heterogeneity across studies.[20]

Furthermore, to further explore sources of heterogeneity, we conducted a series of subgroup analyses based on the following subgroups: drug class, bioexposure, and SpA features (r-axSpA vs nr-axSpA). Data analysis was conducted using the Comprehensive Meta-Analysis software (version 3.3.070).

Given the limited reporting of safety end points by sex, we provided a descriptive summary of the results without conducting a meta-analysis. We visually examined funnel plots to assess publication bias. Additionally, risk of bias was evaluated using the Cochrane collaboration risk of bias tool 1.0 by two independent investigators (AG and LE).[21]

RESULTS

A total of 6,535 records were identified following the initial search and removal of duplicates. Additionally, three more relevant records were found in the FDA and EMA databases (summary basis of approval packages).[22,23] After an initial screening, we selected 1,117 records for full-text review (Figure 1). Of these records, 1,029 were excluded for the following reasons: 586 were post hoc analyses or extensions of the primary study and lacked sex-disaggregated data, 326 were not RCTs, 38 involved the wrong intervention, 33 reported on irrelevant outcomes, 26 had an incorrect duration, 10 were not in English, 4 had the wrong patient population, and 3 could not be retrieved.

A total of 90 records (79 individual trials, 23,748 participants) were included in the SLR (Supplementary Table 4).[7,22–106] Of these trials, 75 trials (23,190 participants) reported the proportion of male and female participants, revealing an overall ratio of male-to-female participation of approximately 2:1 (69.7% male participants). The male-to-female ratio was higher in r-axSpA, at approximately 3:1 (75.3% male participants), whereas it was equal in nr-axSpA (50.1% male participants) (Supplementary Table 5).

Sex-disaggregated data reporting. A minority of the trials provided sex-disaggregated results. Specifically, 9 trials (11.4%, 3,284 participants) reported baseline characteristics by sex,[7,36,48] 22 trials (27.8%, 7,331 participants) reported any efficacy end points by sex,[7,22,23,30,31,35–39,41,44–47,49,61,62,64,66,69,70,82,88–93,105,107–111] and only 9 trials (11.4%, 816 participants) reported safety end

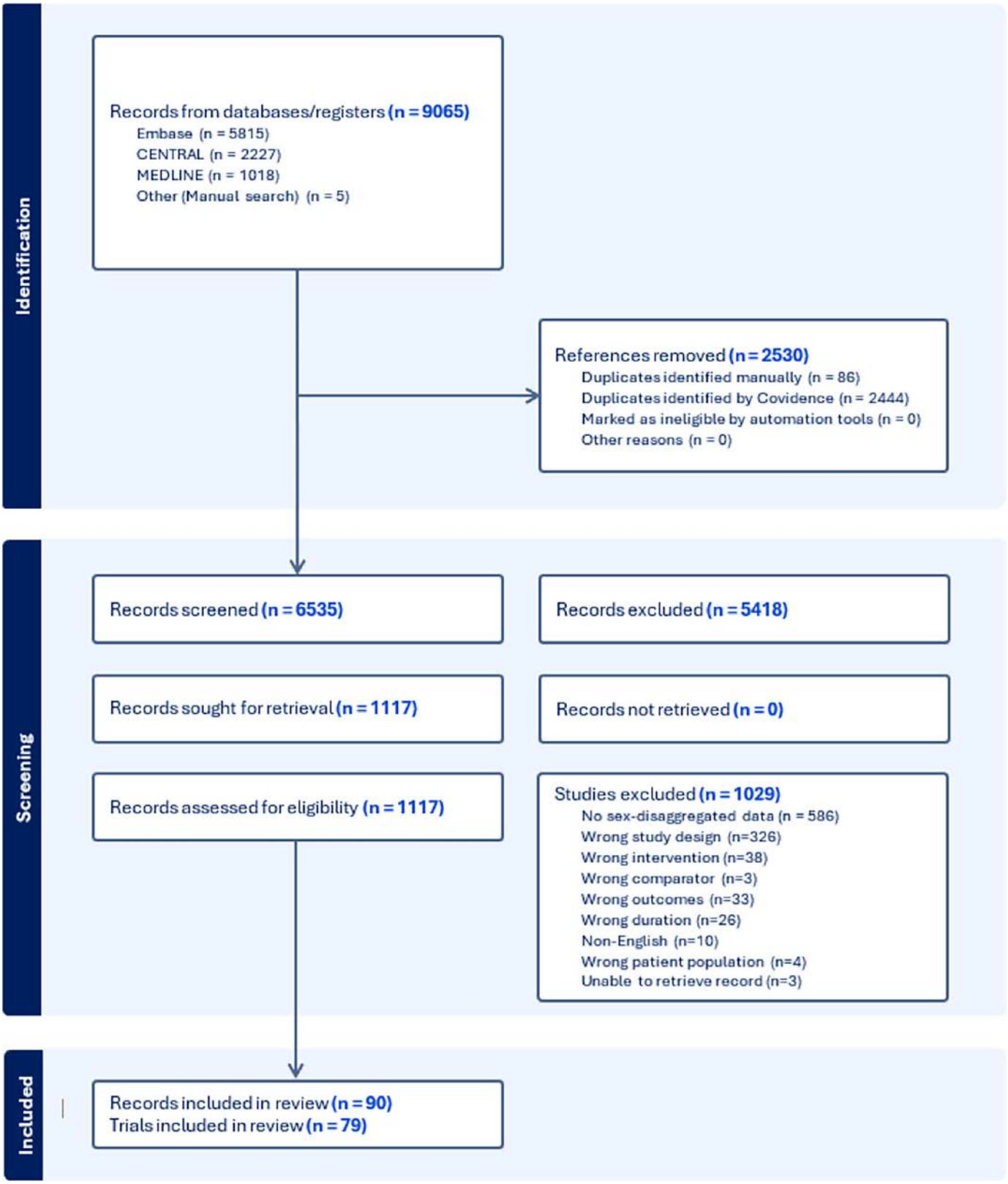

Figure 1. Study flow diagram summarizing the results of the literature search. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25512/abstract.

points by sex.[7,36] For efficacy end points by sex, 13 trials (19.4%, 67 trials)[36,41,48,69,89] reported ASAS40, 9 trials reported ASAS20 (13.2%, 68 trials),[22,23,36,82,92] and 15 trials reported ASDAS (28.8%, 52 trials).[7,36,41,48,70,89]

Baseline characteristics by sex. Nine trials that reported sex-disaggregated data on demographics and disease characteristics were included in the meta-analysis (Table 1, Supplementary Table 6).[7,36,48] Overall, male patients had higher baseline CRP levels (MD 5.9 mg/L, 95% CI 2.7–9.0 mg/L) and were more likely to be HLA-B27 positive (OR 1.58, 95% CI 1.23–2.03). In contrast, male patients had lower BASDAI scores (MD –0.4, 95% CI –0.5 to –0.3). No significant difference was observed in disease duration and ASDAS score between men and women. Sex-related demographic differences included younger age in male patients (MD –4.48 years, 95% CI –6.48 to –2.69 years) and lower BMIs in male patients (MD –0.63, 95% CI –1.17 to –0.09).[7,36,48]

Efficacy end points by sex. ASAS40 response was reported by sex in 13 trials (six publications, 2,407 patients).[36,41,48,69,89] The overall probability of achieving an ASAS40 response across all drug classes was higher in male compared to female patients (OR 1.88, 95% CI 1.44–2.46; Figure 2).[36,41,48,69,89,106] These sex differences were significant for both interleukin-17A inhibitors (IL-17Ai) (OR 1.82, 95% CI 1.32–2.52) and tumor necrosis factor inhibitor (TNFi) (OR 2.42, 95% CI 1.20–4.86). For interleukin 17-A/F inhibitor (IL-17A/Fi) (bimekizumab), a higher ASAS40 response was observed in men with nr-axSpA (OR 2.57, 95% CI 1.24–5.30), whereas the response among patients with r-axSpA was only numerically higher. Substantial heterogeneity in estimated effects was found within each class and across classes (I^2 >40%).

Similarly, sex-related differences were observed in ASAS20 responses, although fewer trials were included in the meta-analysis (nine trials, five publications, 1,672 patients).[22,23,36,82,92] Male patients were more likely to achieve an ASAS20 response compared to female patients across all trials combined (OR 1.79, 95% CI 1.19–2.69). This higher rate of ASAS20 response was significantly higher in TNFi (OR 1.71, 95% CI 1.14–2.15); however, it was only numerically higher in IL-17 inhibitors (IL-17i) (OR 1.22, 95% CI 0.68–2.22; Figure 3). Substantial heterogeneity in effect size was observed across studies and within the IL-17Ai group (I^2 = 58%).

ASDAS-LDA/ID state was reported in 15 trials (six records, 3,423 patients). Male patients were more likely to achieve the ASDAS-LDA/ID state compared to female patients across all trials (OR 2.19, 95% CI 1.47–3.26).[7,36,41,48,70,89] This preferential response in men was significantly higher in both TNFi (OR 2.42, 95% CI 1.21–4.87) and IL-17Ai (OR 2.08, 95% CI 1.28–3.38). Substantial heterogeneity was found across both classes (I^2 = 64%) (Figure 4). No significant sex differences were observed in ASAS40 placebo response (OR 1.35, 95% CI 0.94–1.92; Supplementary Table 7) and ASAS20 placebo response (OR 1.08, 95% CI 0.80–1.47).

A single trial assessing a JAK inhibitors (JAKi) reported ASAS40 response by sex in patients with r-axSpA receiving upadacitinib.[31] However, we were unable to include this study in the meta-analysis because it reported the rate difference in the response to upadacitinib from placebo rather than reporting the rates of response in the study drug alone. Nevertheless, the trial indicated a numerically higher ASAS40 in men versus women (response rate difference in men 30.9, 95% CI 15.0–46.9; response rate difference in women 15.3, 95% CI –9.7 to 20.3).

Subgroup analyses and meta-regression. Subgroup analyses and meta-regression revealed higher rates of ASAS40 and ASAS20 response and ASDAS-LDA/ID in men across various factors including drug class, SpA features, and biologic exposure (Supplementary Tables 8–10). Notably, heterogeneity in effect size was influenced primarily by the underlying SpA feature, namely r-axSpA versus nr-axSpA. Our analysis consistently showed that the odds of achieving efficacy end points among men were higher in patients with nr-axSpA (OR 2.60–3.76) compared to those with r-axSpA (OR 1.21–1.72), as depicted in Figure 5. The meta-regression showed a significantly higher sex-

Table 1. Summary results of random-effects meta-analysis of the differences in baseline patient characteristics in male versus female patients*

Variable	Number of records[a]	Number of male/female patients	Mean difference (male − female) (95% CI) or odds ratio of males vs, females (95% CI)	Heterogeneity, I^2 (%)
Age, y	3	1,428/573	−4.5 (−6.5 to −2.5)	0
BMI, kg/m^2	3	1,428/573	−0.63 (−1.17 to −0.09)	0
Disease/symptom duration, y	4	2,385/899	0.1 (−1.6 to 1.8)	26.9
ASDAS mean score	4	2,385/899	0.1 (−0.1 to 0.2)	27.4
BASDAI mean score	4	2,385/899	−0.4 (−0.5 to −0.3)	0
CRP, mg/L	4	2,385/899	5.9 (2.7 to 9.0)	18.8
HLA-B27 positive	3	1,428/573	1.58 (1.23 to 2.03)	0

* ASDAS, Assessment in Spondyloarthritis International Society; BASDAI, Bath Axial Spondyloarthritis Disease Activity Index; BMI, body mass index; CI, confidence interval; CRP, C-reactive protein.
[a] The term "records" refers to the number of published articles or abstracts.

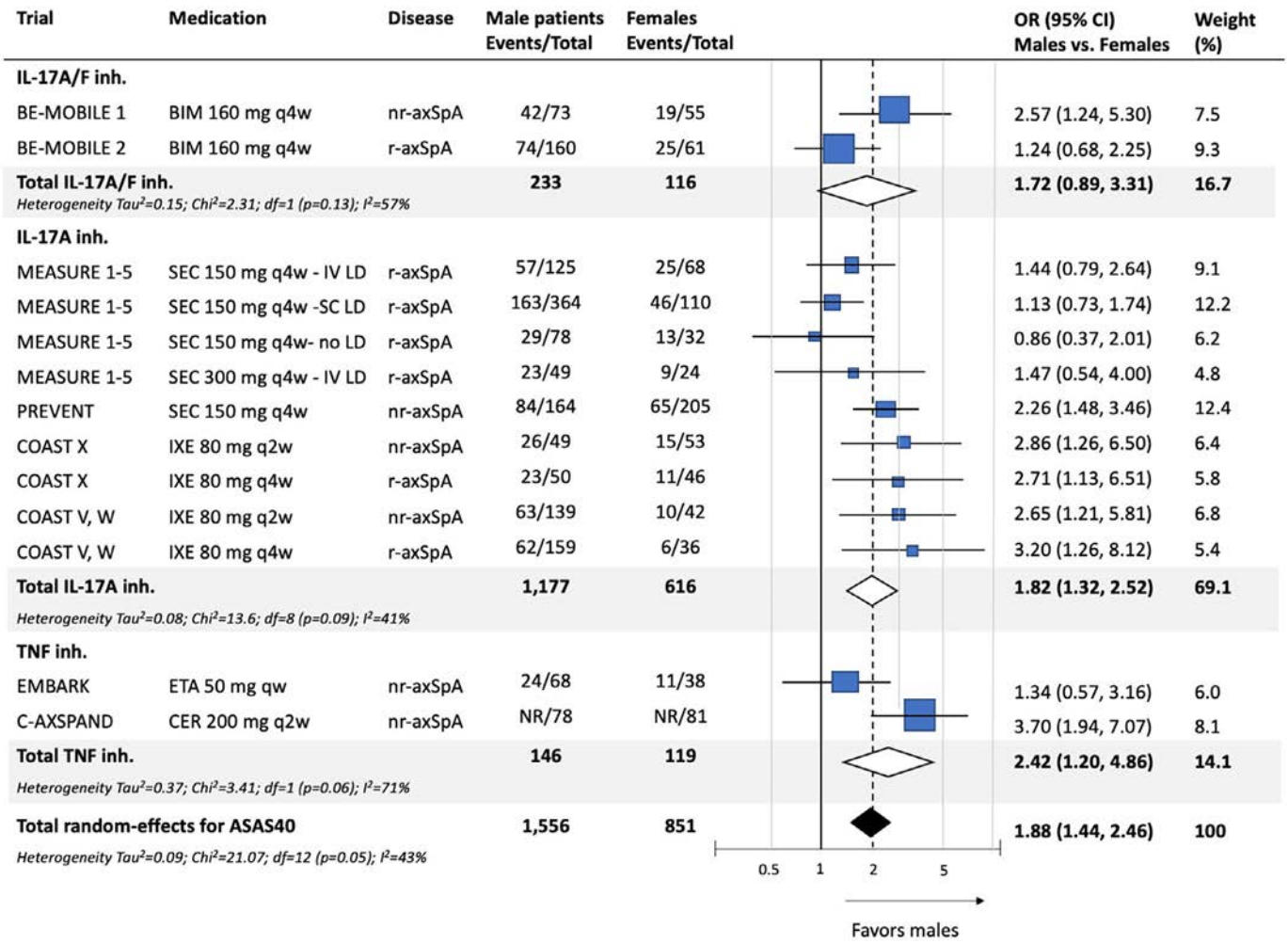

Figure 2. Random-effects meta-analysis of sex differences in ASAS40 response by medication class. ASAS40, Assessment in Spondylarthritis International Society for 40% improvement; C-AXSPAND, Multicenter Study Evaluating Certolizumab Pegol Compared to Placebo in Subjects With axSpA Without X-ray Evidence of AS; CER, certolizumab; CI, confidence interval; ETA, etanercept; IL-17A, interleukin-17A; IL-17A/F, interleukin-17A/F; inh., inhibitor; IV, intravenous; IXE, ixekizumab; LD, loading; NR, non-radiographic; nr-axSpA, nonradiographic axial spondyloarthritis; OR, odds ratio; r-axSpA, radiographic axial spondyloarthritis; SC, subcutanous; SEC, secukinumab; TNF, tumor necrosis factor. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25512/abstract.

related discrepancy in ASDAS-LDA/ID and ASAS40 response among patients with nr-axSpA. No consistent differences in effect size were found between the drug classes and for bionaïve and bioexposed patients.

Imaging. In one trial involving 182 patients treated with TNFi, sex-disaggregated imaging data were analyzed, specifically focusing on magnetic resonance imaging (MRI) vertebral corner inflammation, vertebral corner fat deposition, and radiographic syndesmophyte progression.[109] Overall, male participants were more likely to exhibit progression of spinal imaging findings (adjusted OR ranging from 2.36 to 3.39).[109]

Safety end points. Because of the limited reporting of sex-disaggregated safety data, a meta-analysis could not be conducted. Two publications reported safety end points by sex.[7,36] Pooled data from five secukinumab trials found a numerically higher proportion of female patients treated with secukinumab experiencing any adverse effects (AEs) through 12 months (91% female, 86% male).[36] Furthermore, the rate of serious AEs was numerically higher in men (12%) than women (10%), and the rate of discontinuation due to any AE was numerically higher in men (6%) than women (4%).[36] In another study, pooled data from four etanercept trials revealed a significantly higher rate of discontinuation among men compared to women (hazard ratio 1.49, P = 0.008; male: 7%; female: 10%).[7] Although the overall AE rates were low, men exhibited numerically higher rates of serious infections (men: 0.6%; women: 0%), inflammatory bowel disease (men: 0.3%; women: 0%), and uveitis (men: 2.3%; women: 1.8%).[7]

Risk of bias. Only one study was classified as high risk of bias due to unmasking of the intervention (Supplementary Table 11).[66] The remaining trials were classified as low risk of bias. The funnel plots showed an overall balanced distribution of

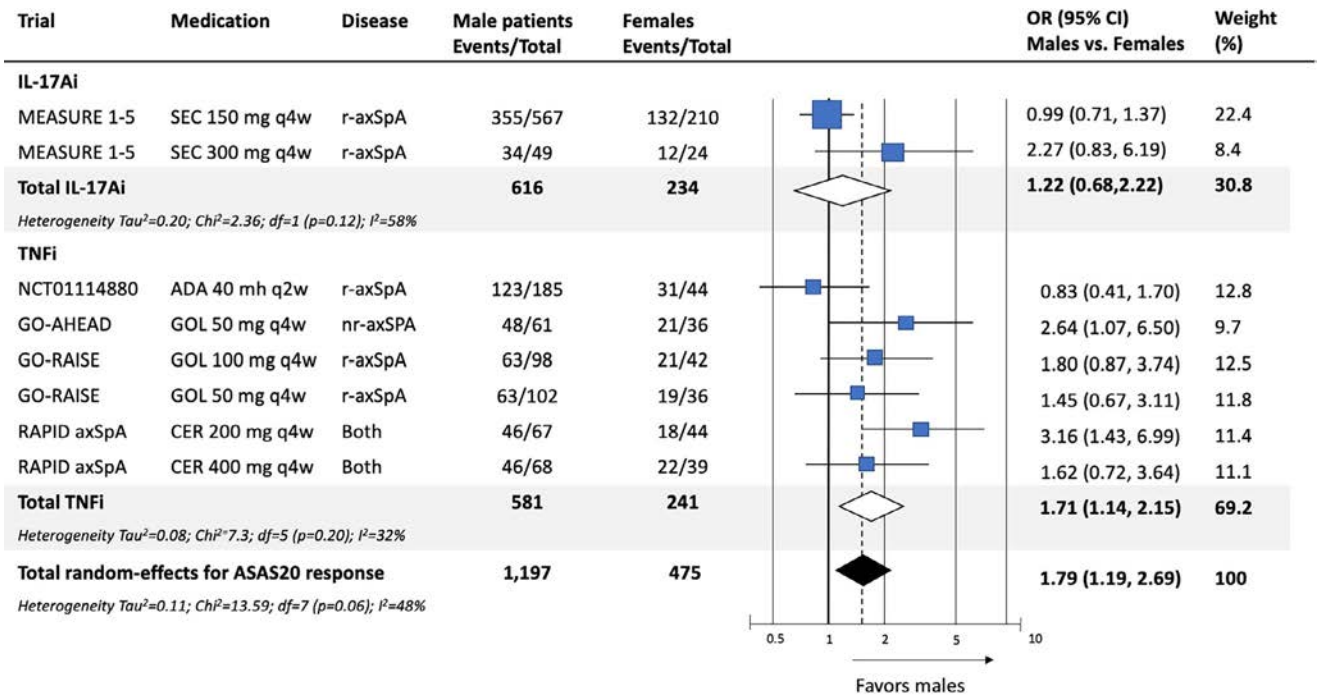

Figure 3. Random-effects meta-analysis of sex differences in ASAS20 response by medication class. ADA, adalimumab; ASAS20, Assessment in Spondylarthritis International Society for 20% improvement; axSpA, axial spondyloarthritis; CER, certolizumab; CI, confidence interval; GOL, golimumab; IL-17Ai, interleukin-17A inhibitor; nr-axSpA, nonradiographic axial spondyloarthritis; OR, odds ratio; r-axSpA, radiographic axial spondyloarthritis; SEC, secukinumab; TNFi, tumor necrosis factor inhibitor. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25512/abstract.

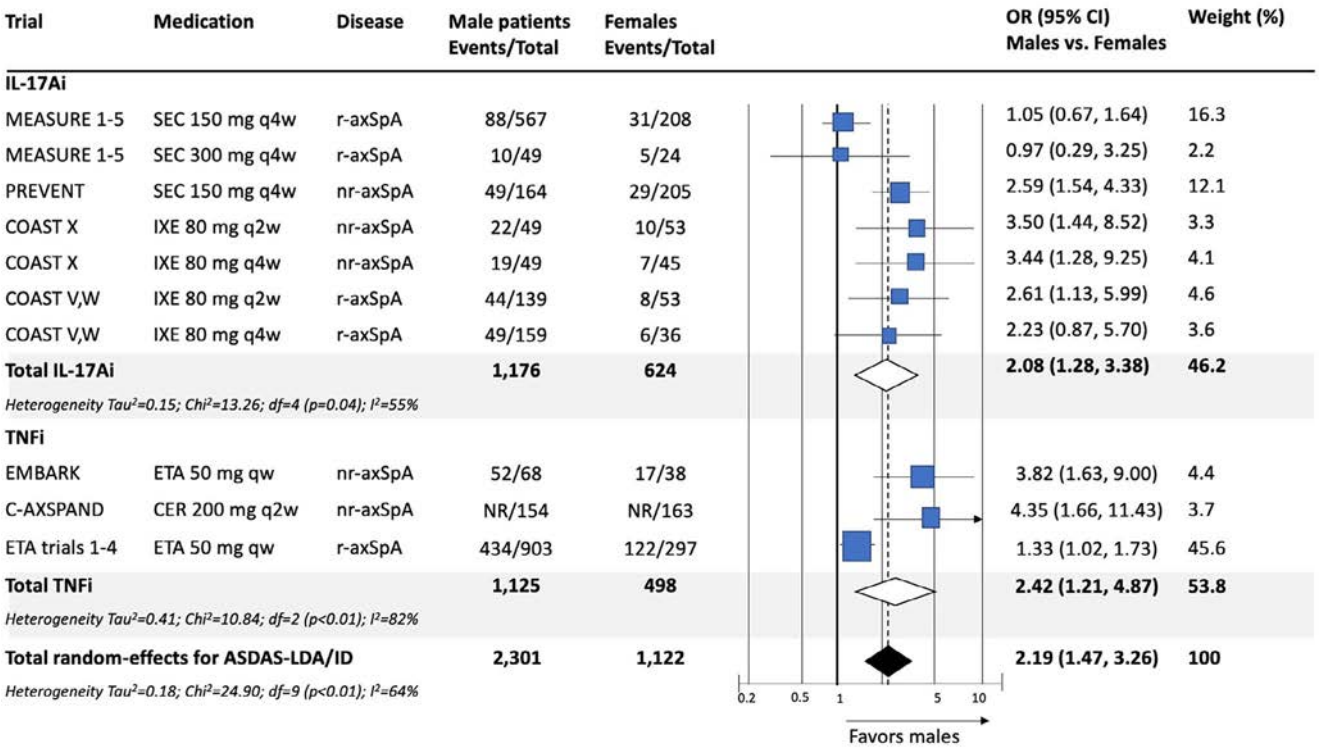

Figure 4. Random-effects meta-analysis of sex differences in ASDAS-LDA/ID medication class. ASDAS-LDA/ID, Axial Spondyloarthritis Disease Activity Score low disease activity or inactive disease; C-AXSPAND, Multicenter Study Evaluating Certolizumab Pegol Compared to Placebo in Subjects With axSpA Without X-ray Evidence of AS; CER, certolizumab; CI, confidence interval; ETA, etanercept; IL-17Ai, interleukin-17A inhibitor; IXE, ixekizumab; NR, non-radiographic; nr-axSpA, nonradiographic axial spondyloarthritis; OR, odds ratio; r-axSpA, radiographic axial spondyloarthritis; SEC, secukinumab; TNFi, tumor necrosis factor inhibitor. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25512/abstract.

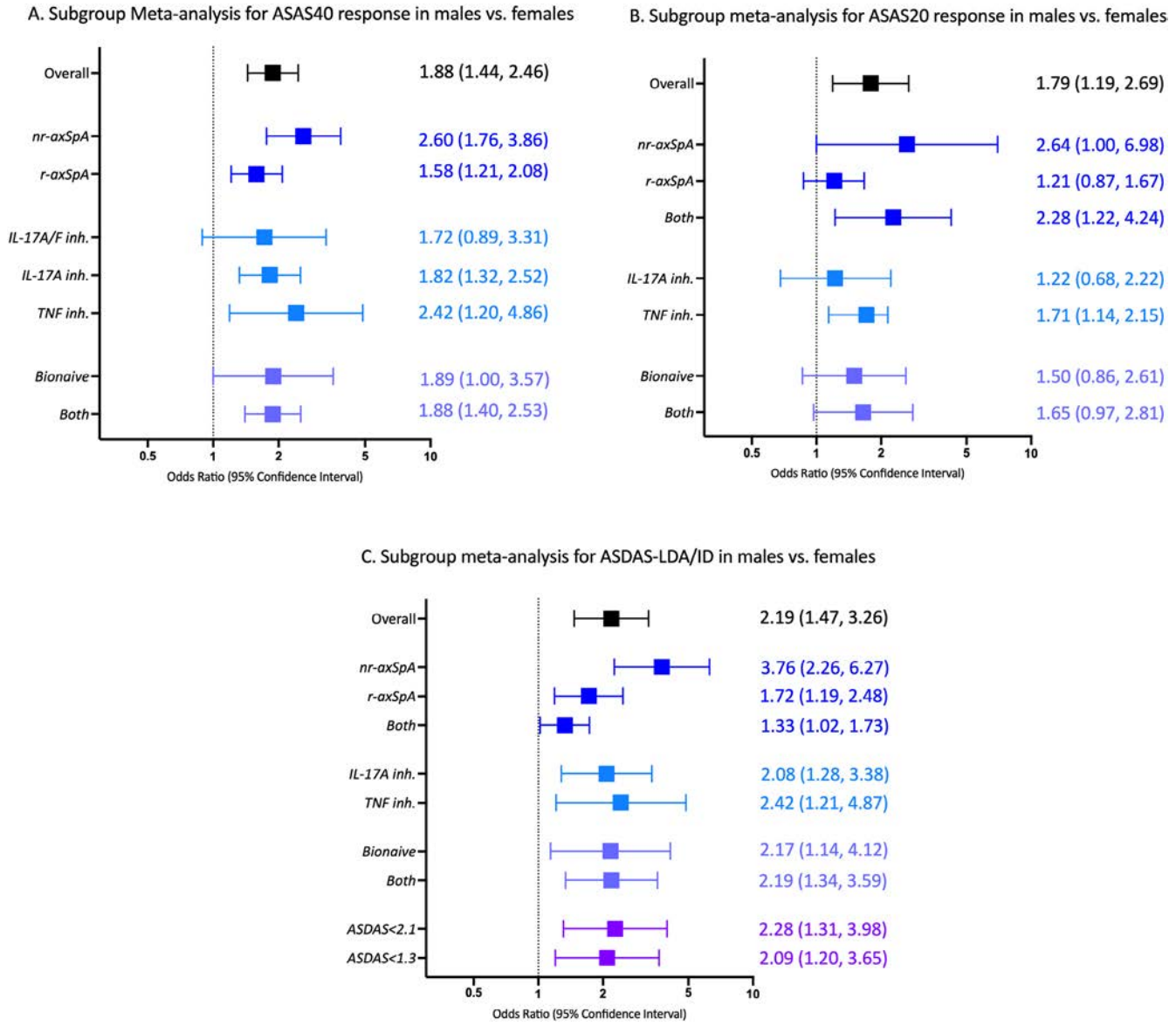

Figure 5. Subgroup meta-analysis of (A) ASAS40 response, (B) ASAS20 response, and (C) ASDAS-LDA/ID in male patients versus female patients (odds ratios and 95% confidence intervals). ASAS20, Assessment in Spondylarthritis International Society for 20% improvement; ASAS40, Assessment in Spondylarthritis International Society for 40% improvement; ASDAS-LDA/ID, Axial Spondyloarthritis Disease Activity Score low disease activity or inactive disease; IL-17A, interleukin-17A; IL-17A/F, uinterleukin 17A/F; inh, inhibitor; nr-axSpA, nonradiographic axial spondyloarthritis; r-axSpA, radiographic axial spondyloarthritis; TNF, tumor necrosis factor.

studies for ASAS40 and ASDAS-LDA/ID. A potential publication bias for ASAS20 response was noted due to a tendency toward the preferential publication of small studies with positive effects and large studies with null effects (Supplementary Table 12).

DISCUSSION

Our meta-analysis found notable sex-related differences in the efficacy of advanced therapies in axSpA RCTs. Across IL-17i and TNFi users, female patients were less likely to achieve efficacy end points compared to their male counterparts. This trend was particularly pronounced among patients with nr-axSpA. Sex differences were also found in patient demographics and disease activity measures at baseline, likely contributing to these differences in treatment response. Finally, similar to recently reported findings in PsA,[16] we identified a significant gap in the reporting of trial end points by sex. Reporting of such data is critical for gaining deeper insights into sex-specific differences in axSpA, ultimately facilitating optimized care for this patient population.

Our meta-analysis findings align with a previous attempt to explore sex-related differences in advanced therapies in

axSpA.[112] This earlier review similarly found lower efficacy in female patients with axSpA. However, the study had some methodologic limitations that may have affected the accuracy of the reported results.[113] One of the limitations is that the authors used fixed effects meta-analysis to pool together study end points despite substantial heterogeneity, which may have led to under estimation of the CIs. The meta-analysis also pooled together studies reporting different end points, which complicates the interpretation of the results, but it failed to report sex differences in baseline characteristics that could have influenced the sex differences in efficacy end points or to report sex differences in safety end points. Finally, the study included fewer trials in the efficacy meta-analysis compared to our study. In contrast, our study offers a more comprehensive and detailed assessment of sex-related difference in axSpA trials. We achieved this by including a larger number of RCTs in the meta-analysis, analyzing multiple end points, and evaluating safety end points by sex. Furthermore, we explored several potential underlying mechanisms and sources of heterogeneity to elucidate our results. For instance, we identified a higher burden of symptoms at baseline among female patients, which may partially account for their lower likelihood of achieving an ID or LDA state. Overall, our meta-analysis findings are consistent with observational studies, indicating reduced effectiveness of biologic medications in axSpA.[12–14,114–116] However, by exclusively including RCTs in our analysis, we mitigated biases inherent to observational studies, thereby enhancing the robustness of our findings.

Whether the patient's sex could guide the selection of advanced therapies remains unknown. Although intriguing, our study did not find any differences in efficacy end points among IL-17Ai, IL-17A/Fi, and TNFi. In a recent meta-analysis of RCTs in PsA, differential sex-related responses were found across drug classes.[16] Specifically, comparable responses were observed between men and women for JAK and Tyk2 inhibitors, whereas biologic therapies showed a preferential response in male patients.[16] Unfortunately, we could not include any JAKi trial in our meta-analysis, which precluded direct comparisons across drug classes. Sex-related dimorphisms have been reported in immune profiles of patients with r-axSpA. Gracey et al[117] reported a lower signature of the IL-17 axis in female patients, whereas both men and women exhibited similar activation of the Th1 axis. Despite these differences in immune profiles, they did not translate into discernible sex-related response disparities across IL-17i and TNFi, both of which showed a similar magnitude of male preferential response.

Reports of lower response among female patients to biologic therapies extend beyond axSpA to other rheumatic conditions, potentially explained by sex- or gender-related mechanisms. Apart from physiologic differences in pain perception or amplification, societal perceptions of feminine or masculine behavior can influence patient-reported pain scores.[118] Additionally, sex-related differences in response to biologic therapies may stem from factors, such as immunogenicity and the development of autoantibodies to drugs in female patients,[119] higher baseline inflammatory burden in male patients, as well as higher BMIs in female patients, which has been linked to reduced response to biologic therapies.[120] In addition, comorbidities that tend to be higher in female patients, such as depression and fibromyalgia, may also affect response to therapy. Unfortunately, we couldn't explore the causal mediating effects of these factors due to the aggregated nature of reported data. However, an axSpA-specific factor that may have contributed to the lower response in female patients is the severity of spinal imaging changes. Our subgroup meta-analysis revealed more pronounced sex differences in patients with nr-axSpA compared to r-axSpA. It is conceivable that misdiagnosis of female patients as nr-axSpA due to nonspecific MRI changes in the sacroiliac joints may have contributed to the lower treatment efficacy in female patients in this group.[121]

A key finding of our study is the strikingly low rate of reporting sex-disaggregated results in RCTs. Merely 28% of trials reported study end points by sex, with an even lower 11% reporting safety data by sex. It is noteworthy that regulatory agencies such as the FDA mandate analysis of drug efficacy by sex as part of the licensing process of new drugs. However, these data remain unpublished in medical journals, possibly due to a combination of lack of awareness regarding their importance and commercial considerations. The Sex and Gender Equity in Research guidelines underscore the importance of considering sex and gender in research publications, specifically reporting sex-disaggregated study outcomes.[122] It is critical that investigators, journals reviewers, and editors adhere to these recommendations, especially when reporting the results of pivotal RCTs of new drugs. In addition, the results of the study can also inform future trial design. For example, trial designers may consider stratifying by sex to reduce the risk of differences in response between placebo and the intervention arms being, at least partly, due to imbalance in sex distribution across study arms. This was recently implemented in the ARGO trial, a phase 2 trial that assessed sonelokimab in patients with PsA.[123]

Our study had several limitations worth noting. Firstly, the scarcity of sex-disaggregated data raises concerns regarding publication bias. The publication of such data appears to be higher in recent years, likely reflecting a growing awareness in sex differences in recent years. Secondly, the lack of patient-level data means that we cannot investigate other factors such as concomitant medications or BMI as potential underlying mechanisms. Lastly, the lack of sex-disaggregated safety end points meant that we were limited to descriptive reporting.

In conclusion, our findings underscore the significant influence of the patient sex on the response to advanced therapies in axSpA. Overall, female patients were less likely to achieve

efficacy outcomes compared to male patients. However, the underlying mechanisms driving this differential response remain unknown. To address this gap in knowledge, it is essential that future trials prioritize the reporting of sex-disaggregated data. A greater understanding of this topic should facilitate the integration of sex and gender considerations in drug prescription ultimately leading to improvement in patient care.

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 Eder 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. Walsh JA, Magrey M. Clinical manifestations and diagnosis of axial spondyloarthritis. J Clin Rheumatol 2021;27(8):e547–e560.
2. Sieper J, van der Heijde D. Review: nonradiographic axial spondyloarthritis: new definition of an old disease? Arthritis Rheum 2013; 65(3):543–551.
3. West HF. Aetiology of ankylosing spondylitis. Ann Rheum Dis 1949; 8(2):143–148.
4. Boonen A, Sieper J, van der Heijde D, et al. The burden of non-radiographic axial spondyloarthritis. Semin Arthritis Rheum 2015; 44(5):556–562.
5. Jovaní V, Blasco-Blasco M, Ruiz-Cantero MT, et al. Understanding how the diagnostic delay of spondyloarthritis differs between women and men: a systematic review and metaanalysis. J Rheumatol 2017; 44(2):174–183.
6. van Tubergen A, Ramiro S, van der Heijde D, et al. Development of new syndesmophytes and bridges in ankylosing spondylitis and their predictors: a longitudinal study. Ann Rheum Dis 2012;71(4): 518–523.
7. van der Horst-Bruinsma IE, Zack DJ, Szumski A, et al. Female patients with ankylosing spondylitis: analysis of the impact of gender across treatment studies. Ann Rheum Dis 2013;72(7):1221–1224.
8. Andreasen RA, Kristensen LE, Egstrup K, et al. The impact of sex and disease classification on patient-reported outcome measures in axial spondyloarthritis: a descriptive prospective cross-sectional study. Arthritis Res Ther 2019;21(1):1–11.
9. Akgul O, Bodur H, Ataman S, et al. Clinical performance of ASAS Health Index in patients with ankylosing spondylitis and non-radiographic axial spondyloarthritis: real-world evidence from Multicenter Nationwide Registry. Rheumatol Int 2020;40(11):1793–1801.
10. Webers C, Ortolan A, Sepriano A, et al. Efficacy and safety of biological DMARDs: a systematic literature review informing the 2022 update of the ASAS-EULAR recommendations for the management of axial spondyloarthritis. Ann Rheum Dis 2023;82(1):130–141.
11. Ortolan A, Webers C, Sepriano A, et al. Efficacy and safety of non-pharmacological and non-biological interventions: a systematic literature review informing the 2022 update of the ASAS/EULAR recommendations for the management of axial spondyloarthritis. Ann Rheum Dis 2023;82(1):142–152.
12. Rusman T, Ten Wolde S, Euser SM, et al. Gender differences in retention rate of tumor necrosis factor alpha inhibitor treatment in ankylosing spondylitis: a retrospective cohort study in daily practice. Int J Rheum Dis 2018;21(4):836–842.
13. Hebeisen M, Neuenschwander R, Scherer A, et al; rheumatologists of the Swiss Clinical Quality Management Program. Response to tumor necrosis factor inhibition in male and female patients with ankylosing spondylitis: data from a Swiss cohort. J Rheumatol 2018;45(4):506–512.
14. Maneiro JR, Souto A, Salgado E, et al. Predictors of response to TNF antagonists in patients with ankylosing spondylitis and psoriatic arthritis: systematic review and meta-analysis. RMD Open 2015; 1(1):e000017.
15. Tannenbaum C, Day D; Matera Alliance. Age and sex in drug development and testing for adults. Pharmacol Res 2017;121:83–93.
16. Eder L, Mylvaganam S, Pardo JP, et al. Sex-related differences in patient characteristics, and efficacy and safety of advanced therapies in randomised clinical trials in psoriatic arthritis: a systematic literature review and meta-analysis. Lancet Rheumatol 2023;5(12): e716–e727.
17. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71.
18. Petitti DB. Meta-Analysis, Decision Analysis, and Cost-Effectiveness Analysis: Methods for Quantitative Synthesis in Medicine. 2nd ed. Oxford University Press; 2000.
19. Haidich AB. Meta-analysis in medical research. Hippokratia 2010; 14(suppl 1):29–37.
20. Cumpston M, Li T, Page MJ, et al. Updated guidance for trusted systematic reviews: a new edition of the Cochrane Handbook for Systematic Reviews of Interventions. Cochrane Database Syst Rev 2019;10(10):ED000142.
21. Schünemann HJ, Vist GE, Higgins JPT, et al. Interpreting results and drawing conclusions. In: Higgins JPT, Thomas HJ, Chandler J, et al. Cochrane Handbook for Systematic Reviews of Interventions. The Cochrane Collaboration; 2019:403–431.
22. Food and Drug Administration. Approval Package for: Application Number: 125160s215. CIMIZA. Center for Drug Evaluation and Research; 2013. Accessed August 15, 2024. https://www.accessdata.fda.gov/drugsatfda_docs/nda/2013/125160Orig1s215.pdf
23. Food and Drug Administration. Application Number 125289. Clinical pharmacology and Biopharmaceutics Review(s). Center for Drug Evaluation and Research 2009. Accessed August 15, 2024. https://www.accessdata.fda.gov/drugsatfda_docs/nda/2009/125289s0000TOC.cfm
24. Deodhar A, Gensler LS, Sieper J, et al. Three multicenter, randomized, double-blind, placebo-controlled studies evaluating the efficacy and safety of ustekinumab in axial spondyloarthritis. Arthritis Rheumatol 2019;71(2):258–270.
25. Baeten D, Østergaard M, Wei JC, et al. Risankizumab, an IL-23 inhibitor, for ankylosing spondylitis: results of a randomised, double-blind, placebo-controlled, proof-of-concept, dose-finding phase 2 study. Ann Rheum Dis 2018;77(9):1295–1302.
26. Peters E, Chou RC, Rozzo SJ, et al. A randomized, double-blind, placebo-controlled phase 2a study of tildrakizumab efficacy and safety in patients with active ankylosing spondylitis. J Clin Rheumatol 2023;29(5):223–229.
27. van der Heijde D, Baraliakos X, Gensler LS, et al. Efficacy and safety of filgotinib, a selective Janus kinase 1 inhibitor, in patients with active ankylosing spondylitis (TORTUGA): results from a randomised,

placebo-controlled, phase 2 trial. Lancet 2018;392(10162):2378–2387.

28. van der Heijde D, Baraliakos X, Sieper J, et al. Efficacy and safety of upadacitinib for active ankylosing spondylitis refractory to biological therapy: a double-blind, randomised, placebo-controlled phase 3 trial. Ann Rheum Dis 2022;81(11):1515–1523.
29. van der Heijde D, Deodhar A, Wei JC, et al. Tofacitinib in patients with ankylosing spondylitis: a phase II, 16-week, randomised, placebo-controlled, dose-ranging study. Ann Rheum Dis 2017;76(8):1340–1347.
30. van der Heijde D, Song IH, Pangan AL, et al. Efficacy and safety of upadacitinib in patients with active ankylosing spondylitis (SELECT-AXIS 1): a multicentre, randomised, double-blind, placebo-controlled, phase 2/3 trial. Lancet 2019;394(10214):2108–2117.
31. Van den Bosch F, Poddubnyy D, Stigler J, et al. POS0923 Influence of baseline demographics on improvements in disease activity measures in patients with ankylosing spondylitis receiving upadacitinib: a post hoc subgroup analysis of SELECT-AXIS 1. Ann Rheum Dis 2021;80(suppl 1):722.2–723.
32. Deodhar A, Sliwinska-Stanczyk P, Xu H, et al. Tofacitinib for the treatment of ankylosing spondylitis: a phase III, randomised, double-blind, placebo-controlled study. Ann Rheum Dis 2021;80(8):1004–1013.
33. Baeten D, Baraliakos X, Braun J, et al. Anti-interleukin-17A monoclonal antibody secukinumab in treatment of ankylosing spondylitis: a randomised, double-blind, placebo-controlled trial. Lancet 2013;382(9906):1705–1713.
34. Baeten D, Sieper J, Braun J, et al. MEASURE 1 Study Group; MEASURE 2 Study Group. Secukinumab, an interleukin-17A inhibitor, in ankylosing spondylitis. N Engl J Med 2015;373(26):2534–2548.
35. Braun J, Baraliakos X, Deodhar A, et al. OP0001 effect of secukinumab, an interleukin-17a inhibitor, on spinal radiographic changes through 2 years in patients with active ankylosing spondylitis: results of the phase 3 study, MEASURE 1. Ann Rheum Dis 2016;75:52.
36. van der Horst-Bruinsma I, Miceli-Richard C, Braun J, et al. A pooled analysis reporting the efficacy and safety of secukinumab in male and female patients with ankylosing spondylitis. Rheumatol Ther 2021;8(4):1775–1787.
37. Kivitz AJ, Pavelka K, Dokoupilova E, et al. Sustained improvements in signs and symptoms of active ankylosing spondylitis and reassuring safety with secukinumab 300mg: 3-year results from a phase 3 study. Rheumatol Ther 2018;70:2084.
38. Pavelka K, Kivitz A, Dokoupilova E, et al. Efficacy, safety, and tolerability of secukinumab in patients with active ankylosing spondylitis: a randomized, double-blind phase 3 study, MEASURE 3. Arthritis Res Ther 2017;19(1):285.
39. Huang F, Sun F, Wan WG, et al. Secukinumab provided significant and sustained improvement in the signs and symptoms of ankylosing spondylitis: results from the 52-week, Phase III China-centric study, MEASURE 5. Chin Med J (Engl) 2020;133(21):2521–2531.
40. *Effect of Secukinumab on Radiographic Progression in Ankylosing Spondylitis as Compared to GP2017 (Adalimumab Biosimilar) (SURPASS)*. National Center for Biotechology Information, US Dept of Health and Human Services; 2023. Accessed August 14, 2024. https://clinicaltrials.gov/study/NCT03259074.
41. Deodhar A, Blanco R, Dokoupilová, et al. Improvement of signs and symptoms of nonradiographic axial spondyloarthritis in patients treated with secukinumab: primary results of a randomized, placebo-controlled phase III study. Arthritis Rheumatol 2021;73(1):110–120.
42. Kiltz U, Baraliakos X, Brandt-Juergens J, et al. POS0910 Rvaluation of the nonsteroidal anti-inflammatory drug-sparing effect of Secukinumab in patients with ankylosing spondylitis: results of the multicenter, randomised, double-blind, phase IV ASTRUM-trial. Ann Rheum Dis 2021;80:714–715.
43. Poddubnyy D, Pournara E, Zielinska A, et al. Rapid improvement in spinal pain in patients with axial spondyloarthritis treated with secukinumab: primary results from a randomized controlled phase-IIIb trial. Ther Adv Musculoskelet Dis 2021;13:1759720X211051471.
44. Deodhar A, Poddubnyy D, Pacheco-Tena C, et al; COAST-W Study Group. Efficacy and safety of ixekizumab in the treatment of radiographic axial spondyloarthritis: sixteen-week results from a phase III randomized, double-blind, placebo-controlled trial in patients with prior inadequate response to or intolerance of tumor necrosis factor inhibitors. Arthritis Rheumatol 2019;71(4):599–611.
45. Deodhar A, van der Heijde D, Gensler LS, et al; COAST-X Study Group. Ixekizumab for patients with non-radiographic axial spondyloarthritis (COAST-X): a randomised, placebo-controlled trial. Lancet 2019;395(10217):53–64.
46. Kiprianos A, van der Horst-Bruinsma IE, Bolce R, et al. P278 Baseline characteristics and treatment response to ixekizumab categorised by sex in radiographic and non-radiographic axial spondylarthritis patients through 52 weeks: data from three phase III, randomized, controlled trials. Rheumatology (Oxford) 2022;61(suppl 1):277–283.
47. Magrey M, de Vlam K, Bolce R, et al. Gender differences in baseline clinical characteristics among patients with ankylosing spondylitis and non-radiographic axial spondyloarthritis: data from 3 randomized ixekizumab controlled trials [abstract]. Arthritis Rheumatol 2020;72(suppl 10). https://acrabstracts.org/abstract/gender-differences-in-baseline-clinical-characteristics-among-patients-with-ankylosing-spondylitis-and-non-radiographic-axial-spondyloarthritis-data-from-3-randomized-ixekizumab-controlled-trials/.
48. van der Horst-Bruinsma IE, de Vlam K, Walsh JA, et al. Baseline characteristics and treatment response to ixekizumab categorised by sex in radiographic and non-radiographic axial spondylarthritis through 52 weeks: data from three phase III randomised controlled trials. Adv Ther 2022;39(6):2806–2819.
49. van der Heijde D, Cheng-Chung Wei J, Dougados M, et al; COAST-V study group. Ixekizumab, an interleukin-17A antagonist in the treatment of ankylosing spondylitis or radiographic axial spondyloarthritis in patients previously untreated with biological disease-modifying anti-rheumatic drugs (COAST-V): 16 week results of a phase 3 randomised, double-blind, active-controlled and placebo-controlled trial. Lancet 2018;392(10163):2441–2451.
50. Erdes S, Nasonov E, Kunder E, et al. Primary efficacy of netakimab, a novel interleukin-17 inhibitor, in the treatment of active ankylosing spondylitis in adults. Clin Exp Rheumatol 2020;38(1):27–34.
51. Gaydukova I, Mazurov V, Erdes S, et al. Netakimab reduces the disease activity of radiographic axial spondyloarthritis. Results of ASTERA study. Ann Rheum Dis 2019;78:193–194.
52. van der Heijde D, Deodhar A, Baraliakos X, et al. Efficacy and safety of bimekizumab in axial spondyloarthritis: results of two parallel phase 3 randomised controlled trials. Ann Rheum Dis 2023;82(4):515–526.
53. van der Heijde D, Gensler LS, Deodhar A, et al. Dual neutralisation of interleukin-17A and interleukin-17F with bimekizumab in patients with active ankylosing spondylitis: results from a 48-week phase IIb, randomised, double-blind, placebo-controlled, dose-ranging study. Ann Rheum Dis 2020;79(5):595–604.
54. Wei JC, Kim TH, Kishimoto M, et al; 4827-006 study group. Efficacy and safety of brodalumab, an anti-IL17RA monoclonal antibody, in patients with axial spondyloarthritis: 16-week results from a randomised, placebo-controlled, phase 3 trial. Ann Rheum Dis 2021;80(8):1014–1021.
55. Sieper J, Braun J, Kay J, et al. Sarilumab for the treatment of ankylosing spondylitis: results of a phase II, randomised, double-blind,

placebo-controlled study (ALIGN). Ann Rheum Dis 2015;74(6):1051–1057.

56. Sieper J, Porter-Brown B, Thompson L,et al. Assessment of short-term symptomatic efficacy of tocilizumab in ankylosing spondylitis: results of randomised, placebo-controlled trials. Ann Rheum Dis 2014;73(1):95–100.
57. Taylor PC, van der Heijde D, Landewé R, et al. A phase III randomized study of apremilast, an oral phosphodiesterase 4 inhibitor, for active ankylosing spondylitis. J Rheumatol 2021;48(8):1259–1267.
58. Pathan E, Abraham S, Van Rossen E, et al. Efficacy and safety of apremilast, an oral phosphodiesterase 4 inhibitor, in ankylosing spondylitis. Ann Rheum Dis 2013;72(9):1475–1480.
59. *Efficacy and Safety of Namilumab for Moderate-to-Severe Axial Spondyloarthritis (NAMASTE)*. National Center for Biotechology Information, US Dept of Health and Human Services; 2022. Accessed August 14, 2024. https://classic.clinicaltrials.gov/ct2/show/NCT03622658.
60. Braun J, van der Horst-Bruinsma IE, Huang F, et al. Clinical efficacy and safety of etanercept versus sulfasalazine in patients with ankylosing spondylitis: a randomized, double-blind trial. Arthritis Rheum 2011;63(6):1543–1551.
61. Calin A, Dijkmans BA, Emery P, et al. Outcomes of a multicentre randomised clinical trial of etanercept to treat ankylosing spondylitis. Ann Rheum Dis 2004;63(12):1594–1600.
62. Davis JC Jr, Van Der Heijde D, Braun J, et al; Enbrel Ankylosing Spondylitis Study Group. Recombinant human tumor necrosis factor receptor (etanercept) for treating ankylosing spondylitis: a randomized, controlled trial. Arthritis Rheum 2003;48(11):3230–3236.
63. Gorman JD, Sack KE, Davis JC Jr. Treatment of ankylosing spondylitis by inhibition of tumor necrosis factor alpha. N Engl J Med 2002;346(18):1349–1356.
64. Rudwaleit M, Listing J, Brandt J, et al. Prediction of a major clinical response (BASDAI 50) to tumour necrosis factor α blockers in ankylosing spondylitis. Ann Rheum Dis 2004;63(6):665–670.
65. Rusman T, van der Weijden MAC, Nurmohamed MT, et al. Is treatment in patients with suspected nonradiographic axial spondyloarthritis effective? Six-month results of a placebo-controlled trial. Arthritis Rheumatol 2021;73(5):806–815.
66. Song IH, Hermann K, Haibel H, et al. Effects of etanercept versus sulfasalazine in early axial spondyloarthritis on active inflammatory lesions as detected by whole-body MRI (ESTHER): a 48-week randomised controlled trial. Ann Rheum Dis 2011;70(4):590–596.
67. Dougados M, Braun J, Szanto S, et al. Efficacy of etanercept on rheumatic signs and pulmonary function tests in advanced ankylosing spondylitis: results of a randomised double-blind placebo-controlled study (SPINE). Ann Rheum Dis 2011;70(5):799–804.
68. Dougados M, Combe B, Braun J, et al. A randomised, multicentre, double-blind, placebo-controlled trial of etanercept in adults with refractory heel enthesitis in spondyloarthritis: the HEEL trial. Ann Rheum Dis 2010;69(8):1430–1435.
69. Dougados M, van der Heijde D, Sieper J, et al. Symptomatic efficacy of etanercept and its effects on objective signs of inflammation in early nonradiographic axial spondyloarthritis: a multicenter, randomized, double-blind, placebo-controlled trial. Arthritis Rheumatol 2014;66(8):2091–2102.
70. Brown MA, Bird PA, Robinson PC, et al. Evaluation of the effect of baseline MRI sacroiliitis and C reactive protein status on etanercept treatment response in non-radiographic axial spondyloarthritis: a post hoc analysis of the EMBARK study. Ann Rheum Dis 2018;77(7):1091–1093.
71. Tu L, Zhao M, Wang X, et al. Etanercept/celecoxib on improving MRI inflammation of active ankylosing spondylitis: a multicenter, open-label, randomized clinical trial. Front Immunol 2022;13:967658.
72. Xu M, Lin Z, Deng X, et al. The Ankylosing Spondylitis Disease Activity Score is a highly discriminatory measure of disease activity and efficacy following tumour necrosis factor-α inhibitor therapies in ankylosing spondylitis and undifferentiated spondyloarthropathies in China. Rheumatology (Oxford) 2011;50(8):1466–1472.
73. Wei JC-C, Tsou H-K, Leong P-Y, et al. Head-to-head comparison of etanercept vs. adalimumab in the treatment of ankylosing spondylitis: an open-label randomized controlled crossover clinical trial. Front Med (Lausanne) 2020;7:566160.
74. You Y, Cai M, Lin J, et al. Efficacy of needle-knife combined with etanercept treatment regarding disease activity and hip joint function in ankylosing spondylitis patients with hip joint involvement: a randomized controlled study. Medicine (Baltimore) 2020;99(19):e20019.
75. Zhao M, Shi G, Tao Y, et al. The effect and safety of yisaipu (yisaipu) in the treatment of patients with nonradiographic axial spondyloarthritis in China. Ann Rheum Dis 2017;76(suppl 2):351–352.
76. Dougados M, Wood E, Combe B, et al. Evaluation of the nonsteroidal anti-inflammatory drug-sparing effect of etanercept in axial spondyloarthritis: results of the multicenter, randomized, double-blind, placebo-controlled sparse trial. Clin Exp Rheumatol 2014;32(5):780–781.
77. *Effects of Etanercept on the Heart, Veins and Thickness of Certain Major Arteries in Ankylosing Spondylitis Patients (CREST)*. National Center for Biotechology Information, US Dept of Health and Human Services; 2016. Accessed August 14, 2024. https://classic.clinicaltrials.gov/ct2/show/NCT00910273?term=NCT00910273&draw=1&rank=1.
78. van der Heijde D, Da Silva JC, Dougados M, et al; Etanercept Study 314 Investigators. Etanercept 50 mg once weekly is as effective as 25 mg twice weekly in patients with ankylosing spondylitis. Ann Rheum Dis 2006;65(12):1572–1577.
79. Lin Z, Liao Z, Huang J, et al. Predictive factors of clinical response of infliximab therapy in active nonradiographic axial spondyloarthritis patients. Biomed Res Int 2015;2015:876040.
80. Haibel H, Rudwaleit M, Listing J, et al. Efficacy of adalimumab in the treatment of axial spondylarthritis without radiographically defined sacroiliitis: results of a twelve-week randomized, double-blind, placebo-controlled trial followed by an open-label extension up to week fifty-two. Arthritis Rheum 2008;58(7):1981–1991.
81. Hu Z, Xu M, Li Q, et al. Adalimumab significantly reduces inflammation and serum DKK-1 level but increases fatty deposition in lumbar spine in active ankylosing spondylitis. Int J Rheum Dis 2012;15(4):358–365.
82. Huang F, Gu J, Zhu P, et al. Efficacy and safety of adalimumab in Chinese adults with active ankylosing spondylitis: results of a randomised, controlled trial. Ann Rheum Dis 2014;73(3):587–594.
83. Krabbe S, Østergaard M, Eshed I, et al. Whole-body magnetic resonance imaging in axial spondyloarthritis: reduction of sacroiliac, spinal, and entheseal inflammation in a placebo-controlled trial of adalimumab. J Rheumatol 2018;45(5):621–629.
84. Lambert RG, Salonen D, Rahman P, et al. Adalimumab significantly reduces both spinal and sacroiliac joint inflammation in patients with ankylosing spondylitis: a multicenter, randomized, double-blind, placebo-controlled study. Arthritis Rheum 2007;56(12):4005–4014.
85. Pedersen SJ, Poddubnyy D, Sørensen IJ, et al. Course of magnetic resonance imaging-detected inflammation and structural lesions in the sacroiliac joints of patients in the randomized, double-blind, placebo-controlled danish multicenter study of adalimumab in spondyloarthritis, as assessed by the Berlin and Spondyloarthritis Research Consortium of Canada Methods. Arthritis Rheumatol 2016;68(2):418–429.
86. Ducourau E, Rispens T, Samain M, et al. Methotrexate effect on immunogenicity and long-term maintenance of adalimumab in axial

spondyloarthritis: a multicentric randomised trial. RMD Open 2020; 6(1): e001047.

87. Sieper J, van der Heijde D, Dougados M, et al. Efficacy and safety of adalimumab in patients with non-radiographic axial spondyloarthritis: results of a randomised placebo-controlled trial (ABILITY-1). Ann Rheum Dis 2013;72(6):815–822.
88. Deodhar A, Gensler LS, Kay J, et al. A fifty-two-week, randomized, placebo-controlled trial of certolizumab pegol in nonradiographic axial spondyloarthritis. Arthritis Rheumatol 2019;71(7):1101–1111.
89. Maksymowych WP, Kumke T, Auteri SE, et al. Predictors of long-term clinical response in patients with non-radiographic axial spondyloarthritis receiving certolizumab pegol. Arthritis Res Ther 2021; 23(1):274.
90. Landewé R, Braun J, Deodhar A, et al. Effect of certolizumab pegol on signs and symptoms of axial spondyloarthritis, including ankylosing spondylitis and non-radiographic axial spondyloarthritis: 24-week results of rapid-axSpA study. Ann Rheum Dis 2013;73(1): 39–47.
91. van der Heijde D, Maksymowych WP, Landewé R, et al. THU0201 factors associated with structural damage in the spine, as measured by x-ray, in patients with axial spondyloarthritis treated with certolizumab pegol over 96 weeks. Ann Rheum Dis 2015;74:268.
92. Dougados M, Bergman G, Maksymowych WP, et al. THU0218 baseline demographic and disease characteristics associated with response to golimumab in patients with active nonradiographic axial spondyloarthritis. Ann Rheum Dis 2015;74:275.
93. Sieper J, van der Heijde D, Dougados M, et al. A randomized, double-blind, placebo-controlled, sixteen-week study of subcutaneous golimumab in patients with active nonradiographic axial spondyloarthritis. Arthritis Rheumatol 2015;67(10):2702–2712.
94. Tam LS, Shang Q, Kun EW, et al. The effects of golimumab on subclinical atherosclerosis and arterial stiffness in ankylosing spondylitis-a randomized, placebo-controlled pilot trial. Rheumatology (Oxford) 2014;53(6):1065–1074.
95. Deodhar A, Reveille JD, Harrison DD, et al. Safety and efficacy of golimumab administered intravenously in adults with ankylosing spondylitis: results through week 28 of the GO-ALIVE study. J Rheumatol 2018;45(3):341–348.
96. Bao C, Huang F, Khan MA, et al. Safety and efficacy of golimumab in Chinese patients with active ankylosing spondylitis: 1-year results of a multicentre, randomized, double-blind, placebo-controlled phase III trial. Rheumatology (Oxford) 2014;53(9):1654–1663.
97. Li EK, Griffith JF, Lee VW, et al. Short-term efficacy of combination methotrexate and infliximab in patients with ankylosing spondylitis: a clinical and magnetic resonance imaging correlation. Rheumatology (Oxford) 2008;47(9):1358–1363.
98. Mok CC, Li A, Chan KL, et al. AB0656 Golimumab versus pamidronate for the treatment of axial spondyloarthropathy (SPA): Aa48-week randomized controlled trial. Ann Rheum Dis 2014;73-(suppl 2):1022.
99. Sieper J, Lenaerts J, Wollenhaupt J, et al. Efficacy and safety of infliximab plus naproxen versus naproxen alone in patients with early, active axial spondyloarthritis: results from the double-blind, placebo-controlled INFAST study, part 1. Ann Rheum Dis 2014;73(1):101–107.
100. Van den Bosch F, Kruithof E, Baeten D, et al. Randomized double-blind comparison of chimeric monoclonal antibody to tumor necrosis factor alpha (infliximab) versus placebo in active spondylarthropathy. Arthritis Rheum 2002;46(3):755–765.
101. van der Heijde D, Dijkmans B, Geusens P, et al; Ankylosing Spondylitis Study for the Evaluation of Recombinant Infliximab Therapy Study Group. Efficacy and safety of infliximab in patients with ankylosing spondylitis: results of a randomized, placebo-controlled trial (ASSERT). Arthritis Rheum 2005;52(2):582–591.
102. Sadeghi A. Effectiveness of adalimumab versus infliximab in patients with ankylosing spondylitis: a randomized double blinded clinical trial. Int J Clin Rheumtol 2021;16(2):73–78.
103. Inman RD, Maksymowych WP, Group CS; CANDLE Study Group. A double-blind, placebo-controlled trial of low dose infliximab in ankylosing spondylitis. J Rheumatol 2010;37(6):1203–1210.
104. Barkham N, Keen HI, Coates LC, et al. Clinical and imaging efficacy of infliximab in HLA–B27–positive patients with magnetic resonance imaging–determined early sacroiliitis. Arthritis Rheum 2009;60(4): 946–954.
105. Braun J, Brandt J, Listing J, et al. Treatment of active ankylosing spondylitis with infliximab: a randomised controlled multicentre trial. Lancet 2002;359(9313):1187–1193.
106. Europen Medicine Agency. Bimzelx: EPAR - Product Information. European Medicines Agency; 2021.
107. Brandt J, Khariouzov A, Listing J, et al. Six-month results of a double-blind, placebo-controlled trial of etanercept treatment in patients with active ankylosing spondylitis. Arthritis Rheum 2003; 48(6):1667–1675.
108. Maksymowych W, Baraliakos X, Lambert R, et al. POS0301 structural outcomes in the sacroiliac joint after ixekizumab treatment for 16 weeks in patients with active non-radiographic axial spondyloarthritis stratified by gender, HLA-B27, and baseline MRI inflammation. Ann Rheum Dis 2022;81:399.
109. Machado PM, Baraliakos X, van der Heijde D, et al. MRI vertebral corner inflammation followed by fat deposition is the strongest contributor to the development of new bone at the same vertebral corner: a multilevel longitudinal analysis in patients with ankylosing spondylitis. Ann Rheum Dis 2016;75(8):1486–1493.
110. van der Heijde D, Kivitz A, Schiff MH, et al; ATLAS Study Group. Efficacy and safety of adalimumab in patients with ankylosing spondylitis: results of a multicenter, randomized, double-blind, placebo-controlled trial. Arthritis Rheum 2006;54(7):2136–2146.
111. Inman RD, Davis JC, van der Heijde D, et al. Efficacy and safety of golimumab in patients with ankylosing spondylitis: results of a randomized, double-blind, placebo-controlled, phase III trial. Arthritis Rheum 2008;58(11):3402–3412.
112. Xie Y, Liu Y, Wu Q. Effect of gender and age on bDMARD efficacy for axial spondyloathritis patients: a meta-analysis of randomized controlled trials. Rheumatology (Oxford) 2024;63(11):2914–2922.
113. Eder L, Pardo JP, Mease P, et al. Comment on: effect of gender and age on bDMARD efficacy for axial spondyloarthritis patients: a meta-analysis of randomized controlled trials. Rheumatology (Oxford) 2024;63(11):e299–e300.
114. Gremese E, Bernardi S, Bonazza S, et al. Body weight, gender and response to TNF-α blockers in axial spondyloarthritis. Rheumatology (Oxford) 2014;53(5):875–881.
115. Fernández-Carballido C, Sanchez-Piedra C, Valls R, et al. Female sex, age, and unfavorable response to tumor necrosis factor inhibitors in patients with axial spondyloarthritis: results of statistical and artificial intelligence–based data analyses of a national multicenter prospective registry. Arthritis Care Res (Hoboken) 2023;75(1):115–124.
116. Lubrano E, Perrotta FM, Manara M, et al. The sex influence on response to tumor necrosis factor-α inhibitors and remission in axial spondyloarthritis. J Rheumatol 2018;45(2):195–201.
117. Gracey E, Yao Y, Green B, et al. Sexual dimorphism in the Th17 signature of ankylosing spondylitis. Arthritis Rheumatol 2016;68(3): 679–689.
118. Myers CD, Riley JL III, Robinson ME. Psychosocial contributions to sex-correlated differences in pain. Clin J Pain 2003;19(4):225–232.
119. Hambardzumyan K, Hermanrud C, Marits P, et al; SWEFOT study group. Association of female sex and positive rheumatoid factor with low serum infliximab and anti-drug antibodies, related to treatment

failure in early rheumatoid arthritis: results from the SWEFOT trial population. Scand J Rheumatol 2019;48(5):362–366.

120. Liew JW, Huang IJ, Louden DN, et al. Association of body mass index on disease activity in axial spondyloarthritis: systematic review and meta-analysis. RMD Open 2020;6(1):e001225.
121. Stovall R, van der Horst-Bruinsma IE, Liu S-H, et al. Sexual dimorphism in the prevalence, manifestation and outcomes of axial spondyloarthritis. Nat Rev Rheumatol 2022;18(11):657–669.
122. Heidari S, Babor TF, De Castro P, et al. Sex and gender equity in research: rationale for the SAGER guidelines and recommended use. Res Integr Peer Rev 2016;1(1):2.
123. Evaluation of sonelokimab for the treatment of patients with active psoriatic arthritis. US Dept of Health and Human Services. National Center for Biotechnology Information. 2024. Accessed August 14, 2024. https://clinicaltrials.gov/study/NCT05640245.

Arthritis Care & Research
Vol. 77, No. 7, July 2025, pp 827–836
DOI 10.1002/acr.25507

American College of Rheumatology
Empowering Rheumatology Professionals

REVIEW ARTICLE

Lupus Flares: More Common in Dialysis Patients Than in Post–Kidney Transplant Recipients: A Systematic Review and Meta-Analysis

Ansaam Daoud,[1] Loai Dweik,[2] Niraj Desai,[1] Sarfaraz A. Hasni,[3] and Omer N. Pamuk[1]

Objective. In this study, we performed a systematic literature review and meta-analysis to assess the frequency of systemic lupus erythematosus (SLE) flares in patients with end-stage renal disease (ESRD) and patients undergoing renal replacement therapy (RRT), hemodialysis (HD), peritoneal dialysis (PD), and kidney transplant (KT).

Methods. Literature from 1973 to 2023 was searched for studies on the frequency of lupus flares after RRT. Data were extracted for ESRD and each RRT modality. Forest plots and random effect models were used to evaluate the odds ratios (95% confidence interval [CI]) of SLE flares after ESRD or RRT, and study heterogeneity was assessed using I^2 statistics.

Results. A total of 57 studies fulfilled the study entry criteria. A total of 29 studies evaluated extrarenal SLE flares after HD/PD, and five studies evaluated extrarenal SLE flares after KT. The frequency of extrarenal SLE flares was compared between HD and PD in seven studies and between HD/PD and KT in four studies. The recurrence of lupus nephritis (LN) was analyzed in 29 studies. Overall, 35.9% of patients with ESRD had at least one extrarenal flare after RRT. The frequency of extrarenal SLE flare was similar in PD and HD (oods ratio [OR] 1.05, 95% CI 0.57–1.94). Extrarenal flare risk was significantly higher in the PD/HD group compared with that of the KT group (OR 4.36, 95% CI 1.66–11.47; $P = 0.0028$). The recurrence of LN after KT was 3.39%.

Conclusion. Extrarenal lupus flares can still occur in more than one-third of patients with ESRD receiving RRT. Dialysis patients have a higher flare risk than those after KT, with comparable flare risk among patients receiving HD and PD.

INTRODUCTION

Systemic lupus erythematosus (SLE) is a multisystem, chronic autoimmune disease characterized by complex immune dysregulation. Although many patients with SLE experience periods of low disease activity or quiescence interspersed with lupus flares, this pattern is not universal.[1] Up to 30% of patients with SLE experience chronically active disease, and 30% have chronically quiescent disease.[1] One of the most serious and common complications of SLE is lupus nephritis (LN), which has a significant impact on quality of life and results in worse outcomes.[2] Renal involvement occurs in up to 60% of patients with SLE, and despite significant advancements in the management of LN, approximately 20% of these patients still progress to end-stage renal disease (ESRD).[3–7] European Renal Association registry data displayed that the incidence of SLE-associated renal replacement therapy (RRT) was 0.8 per million people per year between 1992 and 2016.[8] Among patients with SLE requiring RRT, hemodialysis (HD) was the initial modality in 74% of cases followed by peritoneal dialysis (PD) in 21%. Kidney transplantation (KT) was the initial RRT modality in approximately 5% of patients.[8]

SLE disease activity is generally thought to become quiescent after ESRD allowing for de-escalation or withdrawal of the immunosuppressive regimen. The purported mechanisms of this decline in lupus activity are considered a natural progression of the disease to a "burnout state" and immune modulation owing to uremia and dialysis.[3,5,6,9,10] The patients usually receive high

[1]Ansaam Daoud, MD, Niraj Desai, MD, Omer N. Pamuk, MD: Case Western Reserve University/University Hospitals, Cleveland, Ohio; [2]Loai Dweik, MD: Cleveland Clinic Akron General, Akron, Ohio; [3]Sarfaraz A. Hasni, MD, MSc: NIH, Bethesda, Maryland.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25507.

Address correspondence via email to Ansaam Daoud, MD, at ansaam.daoud@gmail.com.

Submitted for publication March 5, 2024; accepted in revised form January 29, 2025.

SIGNIFICANCE & INNOVATIONS

- This study offers a detailed review and data analysis, providing valuable insights into the risk of systemic lupus erythematosus (SLE) flares in patients with SLE after end-stage renal disease (ESRD) across different renal replacement therapy (RRT) modalities.
- The finding that more than one-third (35.9%) of patients with ESRD experience at least one extrarenal SLE flare after RRT indicates a significant clinical concern for patients with SLE and ESRD.
- The study provides a comparative risk analysis of SLE flares, revealing that patients receiving dialysis (hemodialysis or peritoneal dialysis) are at a higher risk of extrarenal flares compared with those who have undergone kidney transplantation (KT).
- The findings suggest that KT may offer a considerable advantage in the recurrence of lupus nephritis, with a low recurrence rate (3.39%), which is promising.

doses of immune-suppressant medications after KT, which will also suppress any lupus disease activity. However, data from several studies show up to 80% of patients experiencing ongoing SLE disease activity and flares.[3,5–7,11–19] Through this systematic literature review, we aim to evaluate the risk of SLE flares in patients after ESRD and explore how these risks vary according to the type of RRT, including HD, PD, and KT.

PATIENTS AND METHODS

The systematic review and meta-analysis incorporated the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines and checklists.[20] Given the nature of this meta-analysis of existing data from previously published studies, institutional review board or ethics committee approval was not required. This study did not involve direct human or animal subject research.

Data sources. We performed a literature search using the keywords "systemic lupus erythematosus," "end-stage kidney disease," "renal replacement therapy," "hemodialysis," "peritoneal dialysis," "kidney transplantation," "renal transplantation," and "lupus nephritis," combined with the word "OR" in the PubMed, Web of Science, and Cochrane Library databases from 1973 to 2023. Conference abstracts and references of relevant articles published before October 2023 were also reviewed. Three reviewers (AD, LD, and ONP) conducted a literature search in these databases. The titles and abstracts of all potential articles were reviewed, and then the full text of selected articles was evaluated in detail. Another reviewer analyzed a random 20% of all potential abstracts. There was 100% agreement between reviewers for the included and excluded articles.

We included studies that analyzed the frequency of disease flares in patients with SLE who received RRT because of LN. If there were data for SLE disease flares, we included all LN-associated RRT modalities, such as HD, PD, or KT. The data were collected from studies comparing the frequency of disease flares among different RRTs. The criteria for classifying SLE included the American College of Rheumatology (ACR) 1982, modified 1997, Systemic Lupus International Collaborating Clinics 2012, and ACR/EULAR 2019 criteria.[21–24] We included cohort studies, case control studies, population-based studies, and retrospective and prospective studies. However, no clinical trials met our inclusion criteria. We excluded studies written in languages other than English and abstracts lacking the necessary information for data extraction.

Definition of SLE flares. There was variability in defining lupus flares across the studies reviewed. Therefore, we accepted individual the SLE flare definition of each study. In some studies, lupus flares were defined by an increase in the SLE Disease Activity Index score by three or more points only if the clinical symptoms responded to increased corticosteroid therapy. Some studies used serologic markers like complement levels and anti–double-strand DNA to define lupus flares. Other studies relied on a physician's global assessment or the International Flare Consensus Initiative to define flares, focusing on measurable increases in disease activity in one or more organ systems, which were considered clinically significant and often led to changes in treatment. In studies specifically focusing on the recurrence of LN in KT patients, some flares were solely confirmed through kidney biopsies showing histopathologic evidence of disease activity, and some relied on response to clinical symptoms such as proteinuria, hematuria, or renal dysfunction.

Data extraction, synthesis, and quality assessment. Three authors (AD, LD, and ONP) independently reviewed all articles. We extracted data about the publication date, study method (cohort, case control, population-based, prospective, or retrospective), the type of RRT, and SLE flares. We recorded the frequency of SLE flares identified across different RRT modalities, including PD, HD, and KT. When available, we also recorded the data pertaining to the types of SLE flares.

The Newcastle-Ottawa Quality Assessment Scale for cohort and case control studies was used to evaluate the methodologic quality of eligible studies and risk of bias.[25] This scale analyzes cohort selection and comparisons between groups, outcomes, and adequacy of study period. Two authors (A.D. and O.N.P.) rated each study separately, giving a score out of nine possible points. If there were discrepancies in the scores, the issue was resolved by consensus with a third author (LD) (Tables 1 and 2). Two reviewers' inter-reliability was calculated.

Table 1. Overview of included studies on SLE flares*

Study, year	Years	Study design	n	RRT duration	Follow-up duration	NOS/quality	Study outcome
Gaillard et al, 2023[12]	2008–2011	RC	137	NA	5 y	8	SLE flares on dialysis
Kim et al, 2022[6]	1995–2020	RC	121	HD: 19 mo PD: 10 mo	45 mo	8	SLE flares on HD/PD
Tsai et al, 2019[11]	NA	RC	94	HD (6.3 y), PD (6 y), KT (7.1 y)	HD (6.3 y), PD (6 y), KT (7.1 y)	6	SLE flares and survival on HD/PD/KT
Park et al, 2018[30]	2005–2016	RCC	19	43.3 mo of dialysis before KT	70.1 mo	7	Renal outcomes and SLE flares on KT
Barrera-Vargas et al, 2016[5]	1993–2014	RCC	38	NA	≥18 mo after dialysis initiation	8	SLE flares on dialysis
Oliveira et al, 2012[31]	1992–2010	RCS	14	30 mo of dialysis before KT	NA	5	Clinical outcomes on KT
Kang et al, 2011[13]	1990–2007	RMC	59	NA	HD: 5 y PD: 5 y KT: 10y	5	SLE flares and survival on HD/PD/KT
Zhu et al, 2009[32]	1997–2006	NA	29	NA	≥2 y	4	SLE flares and mortality on HD/PD
Ribeiro et al, 2005[33]	2003–2004	PCS MC	57	54 mo	NA	7	LN and nonrenal SLE flares in ESRD
Siu et al, 2005[14]	1995–2003	RCC	18	35.4 mo	NA	5	SLE flares and outcomes on CAPD
Goo et al, 2004[15]	1990–2000	RC	45	NA	53 mo	5	SLE activity and survival in ESRD
Lee et al, 2003[34]	1991–2001	RCC	26	57.5 mo	57.5 mo	5	SLE activity before/after ESRD
Okano et al, 2001[7]	1982–1999	RCS	14	≥6 mo	NA	5	SLE flares and clinical course on HD
Krane et al, 1999[27]	1988–1994	RCS	19	NA	3 y	6	SLE activity before/after ESRD
Bruce et al, 1999[35]	NA	RC	13	NA	12 mo	5	SLE flares in ESRD
Szeto et al, 1998[36]	1987–1996	RC	18	NA	33.8 mo after dialysis	5	SLE activity on dialysis
Kobayashi et al, 1995[26]	1977–1994	RCS	6	5.2 y	6.6 y	6	SLE activity on dialysis
Stock and Krane, 1993[37]	NA	RCS	6	NA	NA	4	SLE activity on HD/PD
Cheigh et al, 1990[16]	1970–1987	RC	59	NA	77.6 mo	5	SLE activity in ESRD
Nossent et al, 1990[10]	1988	RC	55	NA	NA	5	SLE activity before/after ESRD
Sires et al, 1989[38]	1982–1985	CC	9	NA	23 mo	5	SLE flares and clinical course on HD
Rodby et al, 1987[17]	1981–1986	RCC	8	20.8 mo	20.8 mo after dialysis	4	SLE flares and serologic activity on PD
Correia et al, 1984[39]	NA	NA	24	NA	NA	4	SLE activity and survival in ESRD
Pahl et al, 1984[40]	1974–1983	RCS	11	31 mo	31 mo	5	SLE activity before/after dialysis
Coplon et al, 1983[41]	1969–1973	RCS	10	14.3 mo	NA	5	SLE activity in ESRD
Jarrett et al, 1983[42]	1971–1981	RCC	14	3.1 y	NA	6	SLE activity and survival on dialysis
Kimberly et al, 1981[43]	1970–1979	RCS	39	NA	NA	5	SLE activity and progression to ESRD
Fries et al, 1974[44]	NA	NA	13	NA	NA	NA	NA

* CAPD, continuous ambulatory peritoneal dialysis; ESRD, end-stage renal disease; HD, hemodialysis; KT, kidney transplantation; LN, lupus nephritis; MC, multicenter; mo, months; NA, not available; NOS, Newcastle-Ottawa Scale; PCS, prospective cross-sectional; PD, peritoneal dialysis; RC, retrospective cohort; RCC, retrospective case control; RCS, retrospective case series; RMC, retrospective multicenter; RRT, renal replacement therapy; SLE, systemic lupus erythematosus; y, years.

Statistical analysis. Common effect models were used to detect the frequency of disease flares in the LN-ESRD groups. Studies included in our analysis showed crude data, therefore unadjusted analysis was performed. We used the Mantel-Haenszel test to detect the differences in SLE flares between groups. We created forest plots to summarize composite data, generating the frequencies, odds ratio (OR), and corresponding 95% confidence interval (CI) for each subgroup. We used the I^2 statistic to evaluate heterogeneity between studies, and 25%, 50%, and 75% indicated low, moderate, and high heterogeneity, respectively. Funnel plots were employed to detect publication bias. We used the R Statistical Software Meta

Table 2. Overview of included studies on LN flares*

Study, year	Years	Study design	n	RRT duration	Follow-up duration	NOS/quality	Study outcome
Martínez-López et al, 2022[45]	1980–2018	RCC	21	1.63 y	14.2 y	8	Graft survival and SLE flares on KT
Pattanaik et al, 2022[46]	2006–2017	RC	38	4 y	1,230 d	7	LN recurrence and outcomes on KT
Gołębiewska et al, 2016[47]	1999–2014	RC	19	58 mo	1 mo–10.5 y	7	Early/late outcomes on KT
Çeltİk et al, 2016[48]	2000–2013	RC	12	39 mo	63 mo	7	Clinicopathologic findings and SLE flares on KT
Norby et al, 2010[49]	2008	CS	41	13 mo	NA	9	RLN on KT
Contreras et al, 2010[50]	1987–2006	RCC	6850	NA	4.95 y	7	RLN and graft survival on KT
Burgos et al, 2009[4]	1977–2007	RC	202	3.1 y	11.2 y	7	RLN and allograft loss on KT
Ghafari et al, 2008[51]	1989–2006	RCC	23	NA	87 mo	5	LN flares on KT
Yu et al, 2008[2]	1984–2007	CC	23	29.7 mo	107.2 mo	8	SLE flares and outcomes on KT
Lionaki et al, 2008[52]	1985–2005	CC	26	30 mo	79 mo	7	SLE flares and survival on KT
Moroni et al, 2005[53]	1982–2004	RCC	33	42 mo	91 mo	8	RLN and survival on KT
Goral et al, 2003[54]	1976–2000	RC	50	40 mo	6.8 y	6	RLN on KT
Bartosh et al, 2001[55]	1987–1998	RCC	94	514 d	3 y	7	SLE flares and graft outcomes on KT
Stone et al, 1998[56]	1984–1996	RCS	97	33.5 mo	62.6 mo	7	RLN on KT
Azevedo et al, 1998[57]	1975–1994	RCC	45	NA	NA	8	SLE flares and biopsy findings on KT
Grimbert et al, 1998[58]	1971–1993	CC	53	48 mo before KT	NA	8	SLE flares and graft survival on KT
Bitker et al, 1993[59]	NA	NA	10	NA	NA	NA	KT study and NA
Nyberg et al, 1992[60]	NA	NA	16	NA	NA	5	KT study and NA
Contreras-Rodríguez et al, 1992[61]	NA	NA	16	NA	NA	NA	KT study and NA
Sumrani et al, 1992[62]	NA	NA	40	NA	NA	NA	KT study and NA
Goss et al, 1991[63]	1963–1990	RCS	14	36 mo before KT	43.7 mo	5	SLE flares and graft survival on KT
Rivera et al, 1990[64]	1979–1989	RCS	8	22.1 mo before KT	45.1 mo	7	RLN and graft survival on KT
Bumgardner et al, 1988[65]	1969–1987	RCC	32	0–60 mo before KT	7.1 y	7	SLE flares and outcomes on KT
Roth et al, 1987[66]	1979–1985	RCS	15	11.6 mo before KT	3.1 y	7	SLE flares, LN flares, and graft survival on KT
Mejia et al, 1983[67]	1971–1982	RCS	18	1–30 mo receiving dialysis before KT	4.5 y	6	SLE activity, LN flares, and graft survival on KT
Cameron, 1982[68]	NA	NA	6	NA	NA	NA	KT study and NA
Brown et al, 1979[69]	NA	NA	30	NA	NA	NA	KT study and NA
Amend et al, 1977[70]	NA	NA	11	NA	NA	NA	KT study and NA
ASC/NIH, 1975[71]	NA	RMC	56	NA	2 y	6	SLE flares and survival on KT

* CC, case control; CS, cross-sectional; d, days; KT, kidney transplantation; LN, lupus nephritis; mo, months; NA, not available; NOS, Newcastle-Ottawa Scale; RC, retrospective cohort; RCC, retrospective case control; RCS, retrospective case series; RLN, recurrent lupus nephritis; RMC, retrospective multicenter; RRT, renal replacement therapy; SLE, systemic lupus erythematosus; y, years.

Package (version 4.2.1; R Core Team 2022) to perform all analyses.

RESULTS

Study characteristics. The literature review revealed 583 relevant articles; 487 were deemed unsuitable by title or abstract. We reviewed the full text of the remaining 96 articles, and 57 fulfilled the study entry criteria (Figure 1). The Newcastle-Ottawa Quality Assessment Scale score was evaluated in 50 articles and the mean ± SD of articles was 6.1 ± 1.3 with a maximum score of nine points. Interrater reliability for these quality scores was $\kappa = 0.88$ with two independent reviewers. The general characteristics of studies that showed SLE flares after ESRD development are seen in Table 1. Table 2 summarizes studies that showed SLE flares after KT.

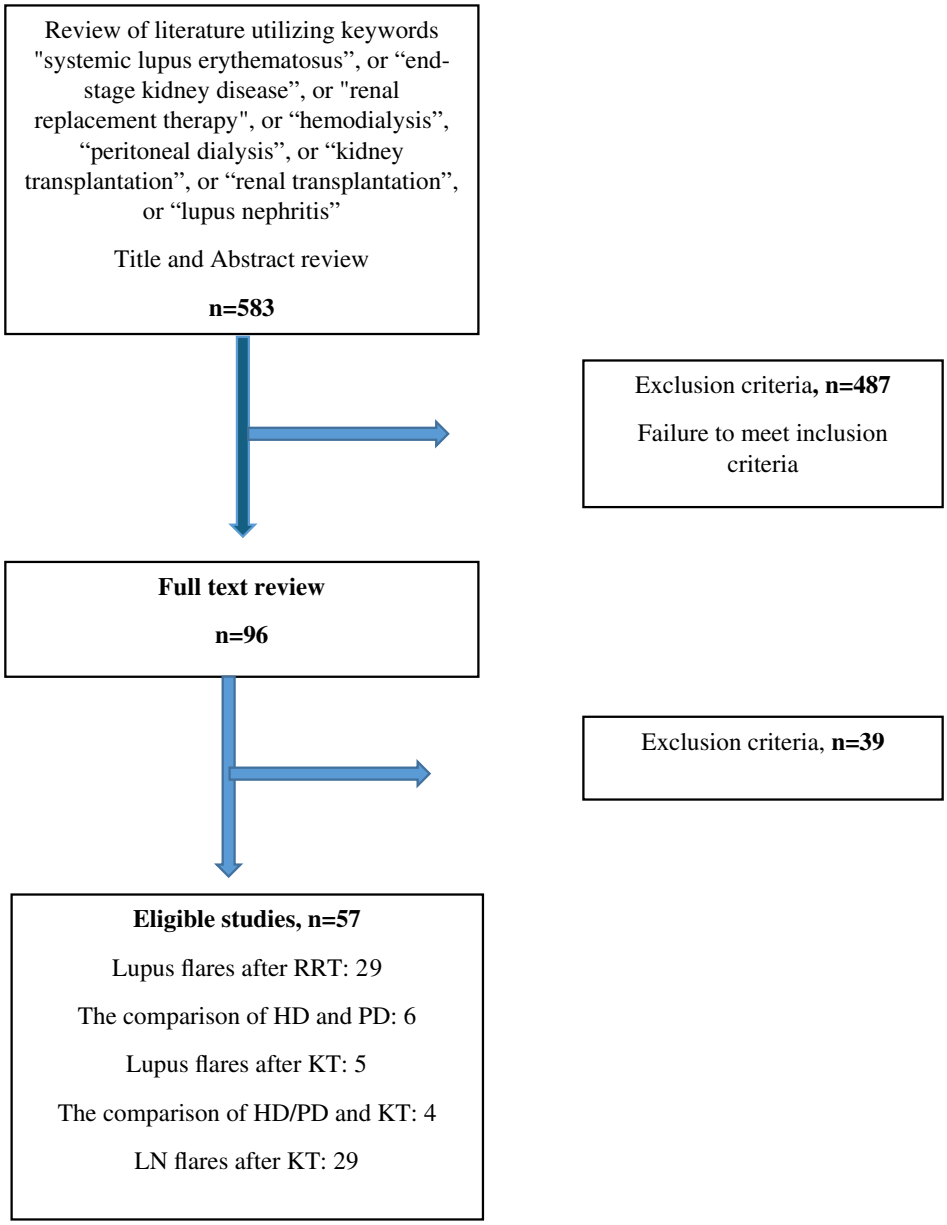

Figure 1. Study algorithm. HD, hemodialysis; KT, kidney transplant; LN, lupus nephritis; PD, peritoneal dialysis; RRT, renal replacement therapy.

Meta-analysis. To identify SLE flares, we used each study's own definitions for SLE flare. Twenty-nine studies (1,016 patients with SLE) evaluated clinical SLE disease flares after dialysis initiation (HD and/or PD). The overall random-effects model pooled a 35.9% (95% CI 33.3–38.9) incidence of at least one lupus flare after dialysis initiation in patients with ESRD with a high level of heterogeneity (I^2 = 89%) (Figure 2). Seven studies (307 patients with SLE) compared SLE disease flares after PD versus HD initiation. The frequency of SLE flares was similar in PD (25.6%) and HD (25.3%) (OR 1.05, 95% CI 0.57–1.94; P = 0.88). Only limited studies evaluated the types of clinical features in SLE flares after RRT. Two studies reported higher than 40% hematologic flares after ESRD development.[4,5] Other common clinical features of flares reported in the studies were arthritis, neurologic disease, mucocutaneous findings, and fever.

Flares after different RRT modalities. Five studies assessed clinical extrarenal lupus flares after KT (13 flares, 206 patients). A total of 6.3% of patients had at least one extrarenal disease flare (95% CI 3.7%–10.6%). Four studies (204 patients with SLE) compared lupus flares in patients with SLE with PD/HD or KT. Flare risk was significantly higher in the PD/HD group compared with the KT group (OR 4.36, 95% CI 1.66–11.47; P = 0.0028) (Figure 3). Twenty-nine studies (7,875 patients with

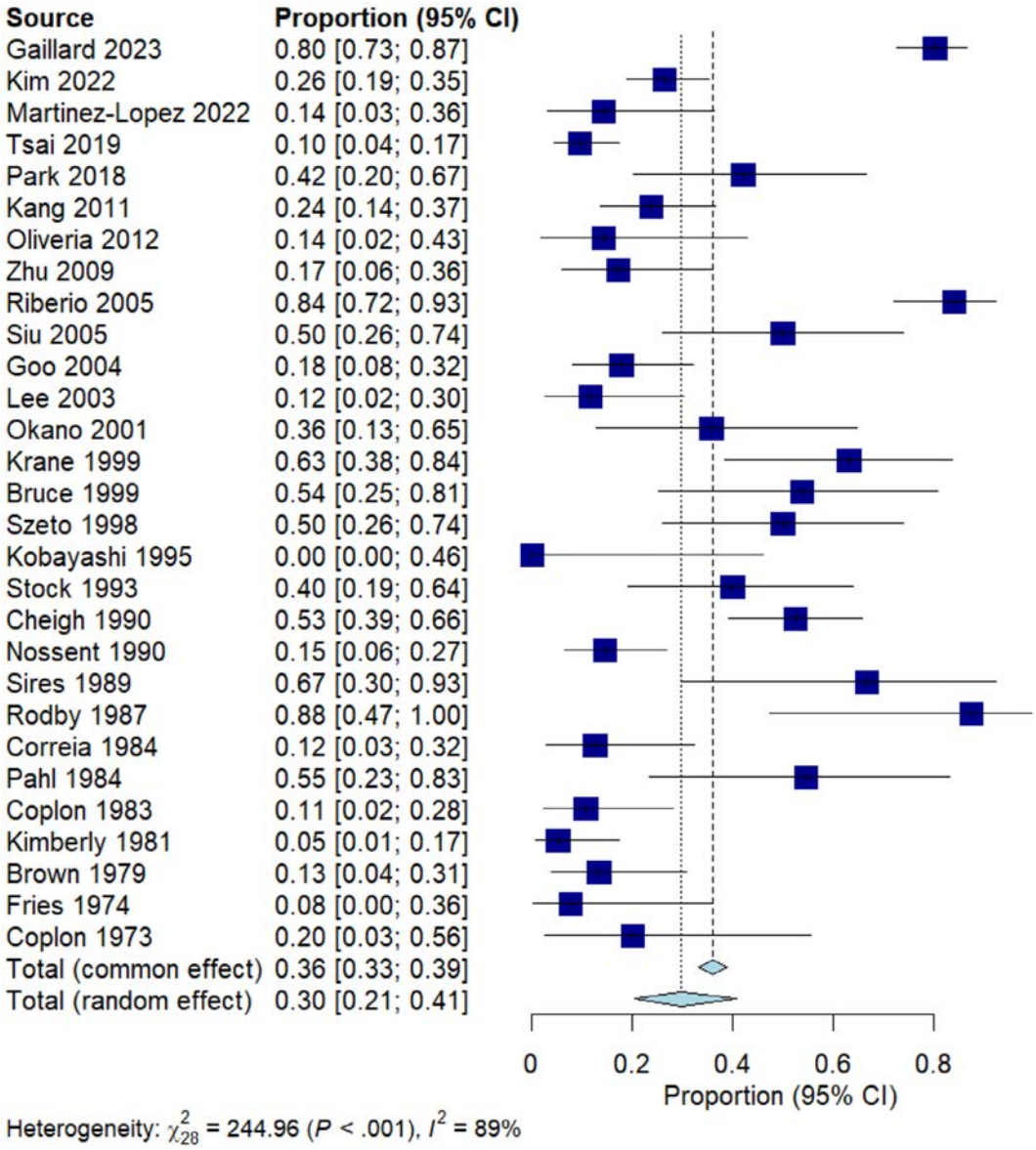

Figure 2. The frequency of systemic lupus erythematosus flares after hemodialysis and/or peritoneal dialysis. CI, confidence interval. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25507/abstract.

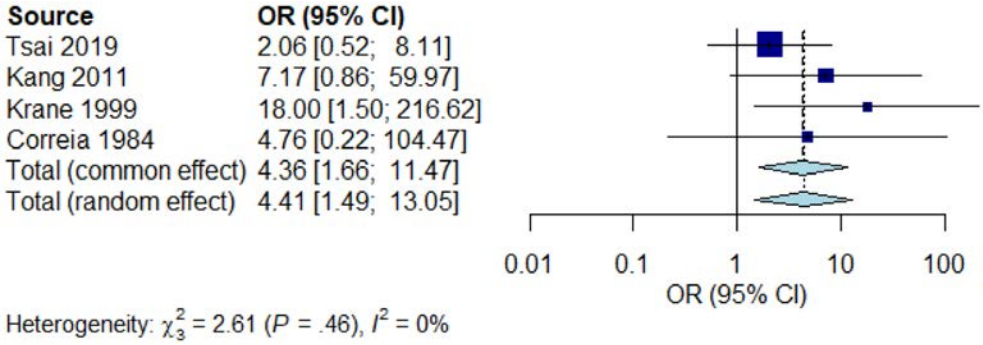

Figure 3. The comparison of systemic lupus erythematosus flares in kidney transplantation (reference group) and dialysis (peritoneal dialysis/hemodialysis). CI, confidence interval; OR, odds ratio. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25507/abstract.

SLE with KT) analyzed the recurrence of LN after KT. A total of 3.39% of patients with lupus with KT had LN relapse (95% CI 3.01%–3.81%; I^2 = 91%) (Figure 4).

DISCUSSION

In the present systematic review, we observed that SLE flares may occur in up to 35.9% of patients with SLE after the development of ESRD. The frequency of SLE flare is similar in PD and HD. On the other hand, SLE flares are relatively uncommon after KT. LN recurrence is possible after KT but is rare.

The frequency of lupus flares after RRT varies among the studies included. These variations could be explained by differences in the definition of lupus flare and geographic location as well as race and ethnicity. The difference in individual study follow-up times could also explain the variations observed. For example, one of the recent studies included in this meta-analysis observed frequent lupus flares (81%) in patients observed for 72 months after starting dialysis.[12] This study found that severe lupus flares were relatively common (19%).[12] Conversely, other studies included found lower rates of lupus flares.[6,11,26] One particular study suggested that lupus flares are common early after dialysis initiation but decrease over time, with clinically active lupus in 55% of patients with SLE in the first year; however, this rate decreased to 6.5% in the fifth year and 0% in the 10th year after dialysis initiation.[16] These longitudinal studies indicate that although clinical disease activity may decline after starting RRT, serologic activity could persist in some patients.

The proposed decline mechanism in SLE activity in the ESRD stage has yet to be fully understood. Studies claim multiple

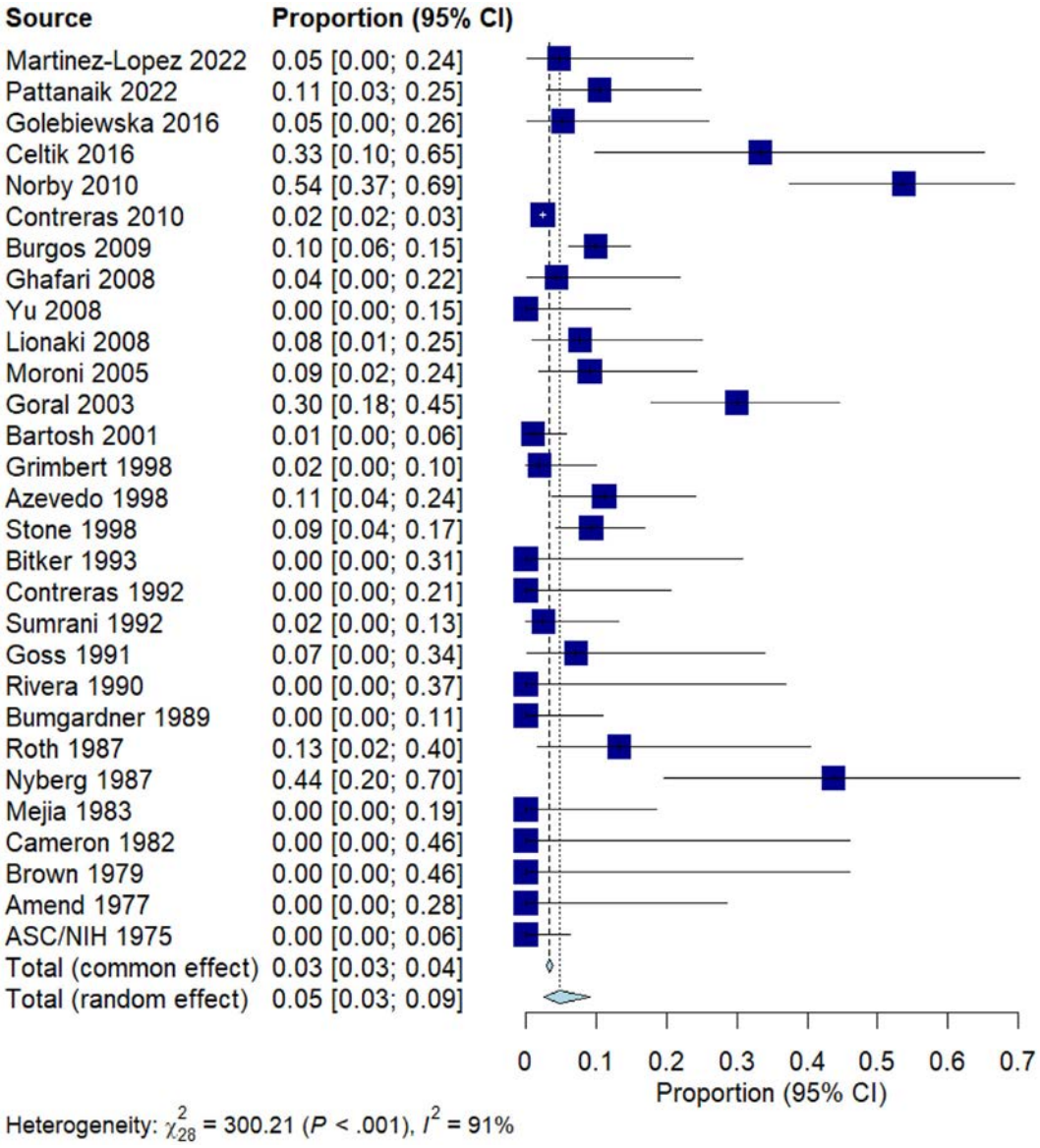

Figure 4. The frequency of lupus nephritis flares after kidney transplantation. CI, confidence interval. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25507/abstract.

immune modulatory factors that contribute to reduced disease activity after ESRD development in SLE. These factors include B cell depletion, impaired B cell maturation, apoptosis of naïve and central memory T cells, shift in T helper ratio, altered maturation of dendritic cells, diminished costimulatory capacity, and decreased interferon α production.[5] In addition, the removal of inflammatory cytokines and autoantibodies by dialysis might be another contributing factor.[6] On the other hand, rapid withdrawal of immunosuppressive treatment, insufficient dose of hydroxychloroquine, and missing rheumatology visits are the main risk factors associated with SLE disease activity after ESRD. It is important to note that our study did not explore these mechanisms directly; rather, they are derived from the literature and proposed as potential explanations.

Our meta-analysis showed similar SLE flares in the HD and PD groups. However, some studies reported higher rates of flares in the HD group than in the PD group.[5,11] The proposed possible mechanisms accounting for differences in immune reactivity between different dialysis modalities include HD-associated cytokine induction through adhesion of mononuclear cells to dialyzer membrane, complement activation of dialyzer membrane, passage of cytokine-inducing bacterial fragments from contaminated dialysate into blood, and HD-associated phagocyte impairment.[15]

Our meta-analysis showed that both renal and extrarenal lupus flares are relatively rare after KT. Patients after KT had better survival rates as compared to those receiving HD or PD.[11,13] For instance, the study from Kang et al found one lupus flare among 17 patients who underwent KT.[13] Another study, which investigated the risk of relapse of LN after KT, did not detect any LN flare in 18 patients who received KT during a 15-year follow-up period.[16] Although the exact mechanism for why lupus flares are less common after KT is unknown, several factors have been proposed. The potent immunosuppressive regimen after KT, bypassing potential pathogenic autoantibody-production from diseased kidneys as a result of KT, and elimination of proinflammatory side effects of HD are the main factors to explain lower renal and extrarenal lupus flares after KT as compared with patients receiving HD or PD.[3] The impact of these proposed mechanisms should be considered when interpreting the lower incidence of lupus flares in KT recipients compared with patients receiving HD or PD.

Most of the studies included in our analysis did not analyze the clinical features of SLE flares after ESRD development. Therefore, we could not evaluate the clinical and serological features of lupus flares in our study. Interestingly, some studies revealed a greater occurrence of hematologic flares compared with other types of lupus flares. Kim et al found that most flare manifestations in patients with SLE receiving dialysis were hematologic (40.6%) followed by neurologic (28%).[6] Another study also reported a similar frequency of hematologic flare (42%) followed by mucocutaneous manifestations (38%) and arthritis (30%).[5] On the other hand, other studies reported various clinical findings of flares, including fever and neurologic, arthritic, and mucocutaneous findings.[7,11,14,27]

Our study has some limitations. First, the classification of lupus and definitions of lupus flares were not uniform across the observational studies included, which could introduce misclassification bias. Second, several studies had no information about the types of SLE flares and concomitant SLE treatment, which can influence outcomes. Therefore, it is important to note that the definition and classification of SLE flares were not uniform across the studies, which could introduce heterogeneity and potential bias in the analysis. Third, variation in follow-up time across studies could lead to heterogeneity in the results, impacting the proportion of patients observed to experience SLE flares. Additionally, the fact that nephrologists, rather than rheumatologists, primarily follow patients with SLE after the development of ESRD may result in the underrecognition of extrarenal lupus flares. Moreover, the extensive timeframe of our study (1973–2023) introduces the potential for bias owing to temporal trends in LN treatment and outcomes. Over the decades, significant advancements in treatment protocols, diagnostic criteria, and patient management have occurred, which may affect the comparability of data from different periods. Future studies should also consider stratifying data by time periods to assess the impact of these temporal changes on study outcomes and further exploring the impact of health care disparities on disease flares in patients with SLE and ESRD.

There are significant health disparities in patients with SLE that may be avoidable. The mechanisms of health disparities in SLE are very complex and involve biologic, sociocultural, physical, and other environmental factors.[28] Health care disparities significantly impact certain racial and ethnic groups, especially African American and Hispanic patients, contributing to worse outcomes in SLE, including a higher risk of LN-ESRD. In addition, self-identified African American patients are 25% less likely to be waitlisted than self-identified non-Hispanic White patients, even after adjusting for medical factors and social determinants of health.[29] Therefore, the differences in the risk of SLE flares in patients receiving dialysis versus in those who received a KT might be due to differences in access to KT by race and ethnicity. African American patients with lupus are more likely to have severe SLE with recurrent flares and are less likely to receive a KT.

In summary, SLE flares after ESRD may be more common than previously thought, and some patients with SLE and ESRD might continue to experience SLE flares. Health care providers should consider continued vigilance for signs of active flares, particularly in the first 5 years after dialysis initiation. Although our study did not specifically examine the role of multidisciplinary care, it is widely recognized that patients with SLE could benefit from ongoing follow-up with both nephrologists and rheumatologists to ensure optimal management of their condition after ESRD. A multidisciplinary approach may help in the timely identification and management of SLE flares even after the onset of ESRD.

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 Daoud 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. Thanou A, Jupe E, Purushothaman M, et al. Clinical disease activity and flare in SLE: current concepts and novel biomarkers. J Autoimmun 2021;119:102615.
2. Yu TM, Chen YH, Lan JL, et al. Renal outcome and evolution of disease activity in Chinese lupus patients after renal transplantation. Lupus 2008;17(7):687–694.
3. Mattos P, Santiago MB. Disease activity in systemic lupus erythematosus patients with end-stage renal disease: systematic review of the literature. Clin Rheumatol 2012;31(6):897–905.
4. Burgos PI, Perkins EL, Pons-Estel GJ, et al. Risk factors and impact of recurrent lupus nephritis in patients with systemic lupus erythematosus undergoing renal transplantation: data from a single US institution. Arthritis Rheum 2009;60(9):2757–2766.
5. Barrera-Vargas A, Quintanar-Martínez M, Merayo-Chalico J, et al. Risk factors for systemic lupus erythematosus flares in patients with end-stage renal disease: a case-control study. Rheumatology (Oxford) 2016;55(3):429–435.
6. Kim YE, Choi SJ, Lim DH, et al. Disease flare of systemic lupus erythematosus in patients with endstage renal disease on dialysis. J Rheumatol 2022;49(10):1131–1137.
7. Okano K, Yumura W, Nitta K, et al. Analysis of lupus activity in end-stage renal disease treated by hemodialysis. Intern Med 2001;40(7):598–602.
8. Derner O, Kramer A, Hruskova Z, et al. Incidence of kidney replacement therapy and subsequent outcomes among patients with systemic lupus erythematosus: findings from the ERA Registry. Am J Kidney Dis 2022;79(5):635–645.
9. Cheigh JS, Stenzel KH. End-stage renal disease in systemic lupus erythematosus. Am J Kidney Dis 1993;21(1):2–8.
10. Nossent HC, Swaak TJ, Berden JH; Dutch Working Party on SLE. Systemic lupus erythematosus: analysis of disease activity in 55 patients with end-stage renal failure treated with hemodialysis or continuous ambulatory peritoneal dialysis. Am J Med 1990;89(2):169–174.
11. Tsai WT, Chang HC, Wang CT, et al. Long-term outcomes in lupus patients receiving different renal replacement therapy. J Microbiol Immunol Infect 2019;52(4):648–653.
12. Gaillard F, Bachelet D, Couchoud C, et al. Lupus activity and outcomes in lupus patients undergoing maintenance dialysis. Rheumatology (Oxford) 2024;63(3):780–786.
13. Kang SH, Chung BH, Choi SR, et al. Comparison of clinical outcomes by different renal replacement therapy in patients with end-stage renal disease secondary to lupus nephritis. Korean J Intern Med 2011;26(1):60–67.
14. Siu YP, Leung KT, Tong MK, et al. Clinical outcomes of systemic lupus erythematosus patients undergoing continuous ambulatory peritoneal dialysis. Nephrol Dial Transplant 2005;20(12):2797–2802.
15. Goo YS, Park HC, Choi HY, et al. The evolution of lupus activity among patients with end-stage renal disease secondary to lupus nephritis. Yonsei Med J 2004;45(2):199–206.
16. Cheigh JS, Kim H, Stenzel KH, et al. Systemic lupus erythematosus in patients with end-stage renal disease: long-term follow-up on the prognosis of patients and the evolution of lupus activity. Am J Kidney Dis 1990;16(3):189–195.
17. Rodby RA, Korbet SM, Lewis EJ. Persistence of clinical and serologic activity in patients with systemic lupus erythematosus undergoing peritoneal dialysis. Am J Med 1987;83(4):613–618.
18. Swai J, Zhao X, Noube JR, et al. Systematic review and meta-analysis of clinical outcomes comparison between different initial dialysis modalities in end-stage renal disease patients due to lupus nephritis prior to renal transplantation. BMC Nephrol 2020;21(1):156.
19. Mojcik CF, Klippel JH. End-stage renal disease and systemic lupus erythematosus. Am J Med 1996;101(1):100–107.
20. Page MJ, Moher D, Bossuyt PM, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ 2021;372(160):n160.
21. Tan EM, Cohen AS, Fries JF, et al. The 1982 revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1982;25(11):1271–1277.
22. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1997;40(9):1725.
23. 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(8):2677–2686.
24. Aringer M, Costenbader K, Daikh D, et al. 2019 European League Against Rheumatism/American College of Rheumatology classification criteria for systemic lupus erythematosus. Arthritis Rheumatol 2019;71(9):1400–1412.
25. Wells GA, Shea B, O'Connell D, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomized studies in meta-analyses. The Ottawa Hospital Research Institute. Accessed August 25, 2024. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp
26. Kobayashi M, Iitsuka T, Ishizu T, et al. Clinicopathological study of dialysis patients with lupus nephritis. Nippon Jinzo Gakkai Shi 1995;37(8):456–461.
27. Krane NK, Burjak K, Archie M, et al. Persistent lupus activity in end-stage renal disease. Am J Kidney Dis 1999;33(5):872–879.
28. Ramos PS. Integrating genetic and social factors to understand health disparities in lupus. Curr Opin Rheumatol 2021;33(6):598–604.
29. Hasan B, Fike A, Hasni S. Health disparities in systemic lupus erythematosus-a narrative review. Clin Rheumatol 2022;41(11):3299–3311.
30. Park ES, Ahn SS, Jung SM, et al. Renal outcome after kidney-transplantation in Korean patients with lupus nephritis. Lupus 2018;27(3):461–467.
31. Oliveira CS, d Oliveira I, Bacchiega AB, et al. Renal transplantation in lupus nephritis: a Brazilian cohort. Lupus 2012;21(5):570–574.
32. Zhu M, Yan Y, Ni Z, et al. Comparison of two-year outcome in maintenance dialysis lupus nephritis patients. Hemodial Int 2009;13(3):385–386.
33. Ribeiro FM, Leite MA, Velarde GC, et al. Activity of systemic lupus erythematosus in end-stage renal disease patients: study in a Brazilian cohort. Am J Nephrol 2005;25(6):596–603.
34. Lee PT, Fang HC, Chen CL, et al. Poor prognosis of end-stage renal disease in systemic lupus erythematosus: a cohort of Chinese patients. Lupus 2003;12(11):827–832.

35. Bruce IN, Hallett DC, Gladman DD, et al. Extrarenal disease activity in systemic lupus erythematosus is not suppressed by chronic renal insufficiency or renal replacement therapy. J Rheumatol 1999;26(7): 1490–1494.
36. Szeto CC, Li PK, Wong TY, et al. Factors associated with active systemic lupus erythematosus after endstage renal disease. J Rheumatol 1998;25(8):1520–1525.
37. Stock GG Jr, Krane NK. Treatment of end-stage renal disease due to lupus nephritis: comparison of six patients treated with both peritoneal and hemodialysis. Adv Perit Dial 1993;9:147–151.
38. Sires RL, Adler SG, Louie JS, et al. Poor prognosis in end-stage lupus nephritis due to nonautologous vascular access site associated septicemia and lupus flares. Am J Nephrol 1989;9(4):279–284.
39. Correia P, Cameron JS, Ogg CS, et al. End-stage renal failure in systemic lupus erythematosus with nephritis. Clin Nephrol 1984;22(6): 293–302.
40. Pahl MV, Vaziri ND, Saiki JK, et al. Chronic hemodialysis in end-stage lupus nephritis: changes of clinical and serological activities. Artif Organs 1984;8(4):423–428.
41. Coplon NS, Diskin CJ, Petersen J, et al. The long-term clinical course of systemic lupus erythematosus in end-stage renal disease. N Engl J Med 1983;308(4):186–190.
42. Jarrett MP, Santhanam S, Del Greco F. The clinical course of end-stage renal disease in systemic lupus erythematosus. Arch Intern Med 1983;143(7):1353–1356.
43. Kimberly RP, Lockshin MD, Sherman RL, et al. "End-stage" lupus nephritis: clinical course to and outcome on dialysis. Experience with 39 patients. Medicine (Baltimore) 1981;60(4):277–287.
44. Fries JF, Powers R, Kempson RL. Late-stage lupus nephropathy. J Rheumatol 1974;1(2):166–175.
45. Martínez-López D, Sánchez-Bilbao L, De Cos-Gómez M, et al. Long-term survival of renal transplantation in patients with lupus nephritis: experience from a single university centre. Clin Exp Rheumatol 2022; 40(3):581–588.
46. Pattanaik D, Green J, Talwar M, et al. Relapse and outcome of lupus nephritis after renal transplantation in the modern immunosuppressive era. Cureus 2022;14(1):e20863.
47. Gołębiewska J, Dębska-Ślizień A, Bułło-Piontecka B, et al. Outcomes in renal transplant recipients with lupus nephritis-a single-center experience and review of the literature. Transplant Proc 2016;48(5):1489–1493.
48. Çeltİk A, Şen S, Tamer AF, et al. Recurrent lupus nephritis after transplantation: clinicopathological evaluation with protocol biopsies. Nephrology (Carlton) 2016;21(7):601–607.
49. Norby GE, Strøm EH, Midtvedt K, et al. Recurrent lupus nephritis after kidney transplantation: a surveillance biopsy study. Ann Rheum Dis 2010;69(8):1484–1487.
50. Contreras G, Mattiazzi A, Guerra G, et al. Recurrence of lupus nephritis after kidney transplantation. J Am Soc Nephrol 2010;21(7):1200–1207.
51. Ghafari A, Etemadi J, Ardalan MR. Renal transplantation in patients with lupus nephritis: a single-center experience. Transplant Proc 2008;40(1):143–144.
52. Lionaki S, Kapitsinou PP, Iniotaki A, et al. Kidney transplantation in lupus patients: a case-control study from a single centre. Lupus 2008;17(7):670–675.
53. Moroni G, Tantardini F, Gallelli B, et al. The long-term prognosis of renal transplantation in patients with lupus nephritis. Am J Kidney Dis 2005;45(5):903–911.
54. Goral S, Ynares C, Shappell SB, et al. Recurrent lupus nephritis in renal transplant recipients revisited: it is not rare. Transplantation 2003;75(5):651–656.
55. Bartosh SM, Fine RN, Sullivan EK. Outcome after transplantation of young patients with systemic lupus erythematosus: a report of the North American pediatric renal transplant cooperative study. Transplantation 2001;72(5):973–978.
56. Stone JH, Millward CL, Olson JL, et al. Frequency of recurrent lupus nephritis among ninety-seven renal transplant patients during the cyclosporine era. Arthritis Rheum 1998;41(4):678–686.
57. Azevedo LS, Romão JE Jr, Malheiros D, et al. Renal transplantation in systemic lupus erythematosus. A case control study of 45 patients. Nephrol Dial Transplant 1998;13(11):2894–2898.
58. Grimbert P, Frappier J, Bedrossian J, et al. Long-term outcome of kidney transplantation in patients with systemic lupus erythematosus: a multicenter study. Groupe Cooperatif de Transplantation d'île de France. Transplantation 1998;66(8):1000–1003.
59. Bitker MO, Barrou B, Ourhama S, et al. Renal transplantation in patients with systemic lupus erythematosus. Transplant Proc 1993; 25(3):2172–2173.
60. Nyberg G, Blohmé I, Persson H, et al. Recurrence of SLE in transplanted kidneys: a follow-up transplant biopsy study. Nephrol Dial Transplant 1992;7(11):1116–1123.
61. Contreras-Rodríguez JL, Bordes-Aznar J, Alberú J, et al. Kidney transplantation in systemic lupus erythematosus: experience from a reference center in Mexico. Transplant Proc 1992;24(5):1798–1799.
62. Sumrani N, Miles AM, Delaney V, et al. Renal transplantation in cyclosporine-treated patients with end-stage lupus nephropathy. Transplant Proc 1992;24(5):1785–1787.
63. Goss JA, Cole BR, Jendrisak MD, et al. Renal transplantation for systemic lupus erythematosus and recurrent lupus nephritis. A single-center experience and a review of the literature. Transplantation 1991;52(5):805–810.
64. Rivera M, Marcen R, Pascual J, et al. Kidney transplantation in systemic lupus erythematosus nephritis: a one-center experience. Nephron J 1990;56(2):148–151.
65. Bumgardner GL, Mauer SM, Payne W, et al. Single-center 1-15-year results of renal transplantation in patients with systemic lupus erythematosus. Transplantation 1988;46(5):703–709.
66. Roth D, Milgrom M, Esquenazi V, et al. Renal transplantation in systemic lupus erythematosus: one center's experience. Am J Nephrol 1987;7(5):367–374.
67. Mejia G, Zimmerman SW, Glass NR, et al. Renal transplantation in patients with systemic lupus erythematosus. Arch Intern Med 1983; 143(11):2089–2092.
68. Cameron JS. Glomerulonephritis in renal transplants. Transplantation 1982;34(5):237–245.
69. Brown CD, Rao TKS, Maxey RW, et al. Regression of clinical and immunological expression of systemic lupus erythematosus (SLE) consequent to development of uremia. Kidney Int 1979;16:884.
70. Amend W, Vincenti F, Covey C, et al. Renal transplantation in systemic lupus erythematosus. Proc Clin Dial Transplant Forum 1977;7:18–22.
71. Renal transplantation in congenital and metabolic diseases. A report from the ASC/NIH renal transplant registry. JAMA 1975;232(2): 148–153.

Arthritis Care & Research
Vol. 77, No. 7, July 2025, pp 837–847
DOI 10.1002/acr.25520

American College of Rheumatology
Empowering Rheumatology Professionals

Global Trends in Risk Factors for Low Back Pain: An Analysis of the Global Burden of Disease Study Data From 1990 to 2021

Katharine E. Roberts,[1] Manuela L. Ferreira,[2] Paula R. Beckenkamp,[1] Sneha Nicholson,[3] Lyn March,[4] and Paulo H. Ferreira[1]

Objective. The increasing burden associated with low back pain (LBP) is a critical issue. This is a novel analysis of trends in risk factors for LBP aiming to identify risk factors that require further attention or consideration in global policies to reduce the burden of LBP.

Methods. The Global Burden of Disease study metadata were used to describe the trends in three modifiable categories of risk factors that contribute to the burden associated with LBP. The trends in occupational/ergonomic, behavioral (smoking), and metabolic (high body mass index [BMI]) risk factors for LBP between 1990 and 2021 have been described with attention to global areas, high sociodemographic index (SDI) areas, and low SDI areas.

Results. The number of years lived with disability (YLDs) caused by LBP increased globally, in high and low SDI areas between 1990 and 2021. The impact of smoking and occupational/ergonomic risk factors have decreased; however, the impact of high BMI has increased markedly in the same time frame, with a particularly concerning impact in high SDI areas and on women.

Conclusion. The burden of LBP is increasing globally, with a significant proportion of the YLDs caused by LBP attributed to three modifiable lifestyle factors: occupation/ergonomics, smoking, and high BMI. Of significant concern is the rapidly increasing impact of high BMI on YLDs caused by LBP, with the greatest impact seen among women in low and high SDI areas. The role of additional risk factors (eg, physical inactivity) still needs to be determined in the context of the global burden of LBP.

INTRODUCTION

The increasing burden associated with low back pain (LBP) is a critical health issue that demands global attention. In 2020, 619 million people experienced LBP globally,[1] with an age-standardized prevalence of 7,460 per 100,000 (range 6,690–8,370), resulting in 69 million years lived with disability (YLDs) worldwide.[1] This represents a significant increase from 1990, when 377.5 million people experienced LBP globally[2] and the associated YLD was 43 million.[1] Driving the demand for attention is the fact that LBP prevalence is forecast to expand to 843 (95% uncertainty interval [UI] 759–933) million by 2050, with the potential for substantial increases in associated disability and burden.[1]

Understanding the mechanisms that explain disease development and analyzing the trajectories of risk factors over time could provide insights into preventive strategies to positively impact disease prevalence and burden as well as identify areas that require further attention or a review in global policy.[3] The Global Burden of Disease (GBD) study generates publicly available datasets and tools that allow the analysis of estimates of global health data. Specifically, GBD study data provide not only estimates of the prevalence of diseases but also the prevalence and impact of risk factors for them, affording the ability

Dr Roberts' work was funded by a University of Sydney Lower Back Pain Scholarship. Dr M. Ferreira holds an Australian National Health and Medical Research Council (NHMRC) Research Fellowship. The Global Burden of Disease project is funded by the Bill and Melinda Gates Foundation.

[1]Katharine E. Roberts, PhD, Paula R. Beckenkamp, PhD, Paulo H. Ferreira, PhD: The University of Sydney, Sydney, New South Wales, Australia; [2]Manuela L. Ferreira, PhD: The George Institute for Global Health, Sydney, New South Wales, Australia; [3]Sneha Nicholson, MS: The University of Washington, Seattle; [4]Lyn March, PhD: The Kolling Institute, 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.25520).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25520.

Address correspondence via email to Katharine Roberts, PhD, at katharine.roberts@sydney.edu.au.

Submitted for publication December 4, 2024; accepted in revised form February 21, 2025.

SIGNIFICANCE & INNOVATIONS

- Understanding the trajectories of risk factors provides insights into preventive strategies that may positively impact low back pain (LBP) burden.
- The impact of occupational/ergonomic risk factors and smoking on the years lived with disability because of LBP is decreasing globally.
- The impact of high body mass index is increasing at an alarming rate and is particularly significant for women in both high sociodemographic index (SDI) areas and low SDI areas.
- We urgently need effective preventive strategies to address obesity, especially among women, if we aim to control the global burden of LBP.

to identify the relative importance of different risk factors both over time and among populations globally.[3]

In 2018, *The Lancet* published an LBP series calling attention to the global dilemma of LBP,[4–6] citing aging and expanding populations, as well as inadequate management, as major contributing factors associated with increasing disability and costs globally.[4] In addition, the study demonstrated that in 2020 more than one-third of the YLDs associated with LBP are associated with three GBD risk factors[1]: occupational/ergonomic, smoking, and high body mass index (BMI). These risk factors have been included in the GBD study LBP estimates, reflecting their relative prevalence in the literature when they meet the evidence of risk-outcome pair criteria.[3] The increased prevalence of these risk factors since 1990 may have been associated with the increasing burden of LBP globally. However, the trends in the risk association between these factors and LBP is still to be ascertained.

The current study aimed, for the first time, to report the global trends in risk factors for LBP from 1990 to 2021. The trends in YLDs have been reported as numbers and percentages providing a comparison between men and women as well as a comparison between the trends in risk factors in high sociodemographic index (SDI) areas and low SDI areas. Understanding the trends of risk factors and their association with the burden of LBP over time could provide insights into preventive strategies to positively impact LBP prevalence and its associated burden as well as important information to guide policy to address the global burden of LBP.

PATIENTS AND METHODS

Data source. The methods used by the GBD study for collecting and calculating LBP and risk factor estimates are described in detail elsewhere.[7] In short, the GBD study uses a large number of data sources to estimate illness, injury, morbidity, and attributable risk for 204 countries.[7,8] Prevalence data and disability weights are then used to calculate YLDs. For LBP, considered to be a nonfatal health outcome, there were 492 input sources from 204 countries and territories identified between 1990 and 2021. Input data sources can be found at https://ghdx.healthdata.org/gbd-2021/sources. The GBD study follows the Guidelines for Accurate and Transparent Health Estimates Reporting Statement.[9] The GBD study systematically reviews and synthesizes data from multiple electronic sources,[7] and a meta-regression approach is used to synthesize the data extracted for each risk-outcome pair as per the Preferred Reporting Items for Systematic Reviews and Meta-analysis framework.[10] The main analytical tools used for GBD 2021 are disease model meta-regression 2.1, spatiotemporal Gaussian process regression, and meta-regression—Bayesian, regularised, trimmed.[7]

Outcome. The GBD study defines LBP as pain experienced in the posterior body between the lower margin of the 12th ribs and the lower gluteal folds, which may or may not include referred pain into one or both lower limbs, and lasts for at least 1 day.[1,11] YLDs were used to describe the burden associated with LBP. Although disability-adjusted life years (DALYs) are often used to describe the burden associated with disease, YLDs are commonly used to describe the burden of nonfatal diseases such as LBP.[1] Estimates are presented as numbers (count), age-standardized percentages, and age-standardized rates (per 100,000 population). The GBD study categorizes risk factors into hierarchical order. Level 1 represents three broad categories (occupational, behavioral, and metabolic), which are progressively broken down into level 2, 3, and 4 risk factors.[3] For example, behavioral risk factors (level 1) may be disaggregated into tobacco (level 2) and smoking, chewing tobacco, and second-hand smoke (level 3). This hierarchy allows analysis of individual risk factors or groups of risk factors.[7]

Risk factors. In the context of LBP, the risk factors reported in the GBD study are occupational/ergonomic (level 3), smoking (levels 2 and 3), and BMI (level 2).[3] Occupational/ergonomic exposures include lifting, forceful movements, vibrations, and awkward postures. High BMI for adults aged >20 years is defined as >20 to 25 kg/m^2.[12] SDI is an indicator of the social and economic conditions (ie, development status) that influence health outcomes. SDI is calculated by the GBD combining total fertility rate, mean education, and lag-distributed income per capita,[8] generating a score between 0 and 1. A score of 0 represents minimum socioeconomic development and, therefore, poorer associated health outcomes.[7] SDI is divided into quintiles with low SDI incorporating countries with an SDI of 0.00 to 0.45 and high SDI incorporating countries with an SDI of 0.81 to 1.00.[13]

The GBD study assesses the impact of each risk factor using a comparative risk assessment framework[7] estimating risk with a six-step meta-analytical method: (1) risk-outcome pairs that meet specific criteria; (2) relative risks as a function of exposure; (3) levels of exposure in each age, sex, location and year; (4) the

theoretical minimum risk exposure; (5) computed attributable deaths, years of life lost, YLDs, and DALYs; and (6) population-attributable fractions and attributable burden.[3] The GBD study currently provides data on only the three risk factors discussed in this study, for which there is credible evidence of risk-outcome relationships. This includes findings supported by more than one study type, data from at least two cohorts, minimal and explained heterogeneity, low risks of confounding and selection bias, and biologically plausible dose-response gradients.[1,7]

Data presentation, UIs, and data access. All data were downloaded from the results[14] and compare tools[15] for presentation. The GBD metadata are publicly available through the Institute for Health Metrics and Evaluation (IHME) Global Health Data Exchange (GHDx) at https://www.healthdata.org/data-tools-practices/interactive-data-visuals and https://vizhub.healthdata.org/gbd-compare/.

Trends in risk factors were viewed using the IHME data visualization tools. Specifically, the GBD "results tool"[14] was used to visualize the trends in risk factors for LBP using 5-year brackets (1990 to 2020 and 2021). The GBD "compare tool"[15] was used to visualize the risk factors for LBP when comparing high SDI countries with low SDI countries. Percentage change was generated by the data visualization tools. The visualization tools were accessed in July 2024.

No patients were involved in this study, and ethical approval was not required for this study. All data relevant to the study are included in the article or uploaded as online supplemental information. The GBD metadata are publicly available through the IHME GHDx.

RESULTS

In 1990, 42.4% of the YLDs caused by LBP globally were attributable to exposure to the three risk factors assessed in this study: occupational/ergonomic, behavioral (smoking), and metabolic (high BMI). By 2020, the proportion had decreased slightly to 38.8%.[1] A global map displaying YLDs caused by LBP in 2021 is presented in Figure 1, and global maps displaying LBP attributable to occupational/ergonomic, behavioral (smoking), and metabolic (high BMI) risk factors for both sexes and all ages in 2021 are available in Supplementary Figures 1, 2, and 3, respectively.

Global trends in YLDs attributable to occupational/ergonomic, smoking, and high BMI risk factors. Globally, the YLDs caused by LBP as a percentage of total YLDs attributable to all causes decreased from 1990 to 2021 for occupational/ergonomic and smoking risk factors, whereas the percentage of YLDs caused by LBP attributable to high BMI increased in the same period (Figure 2). The rank of risk factors by contribution to YLDs caused by LBP has occupational/ergonomic as the most contributing factor, followed by smoking and high BMI, with occupational/ergonomic factors contributing the most YLDs caused by LBP and high BMI contributing the least (Figure 3). That rank order has not changed since 1990.

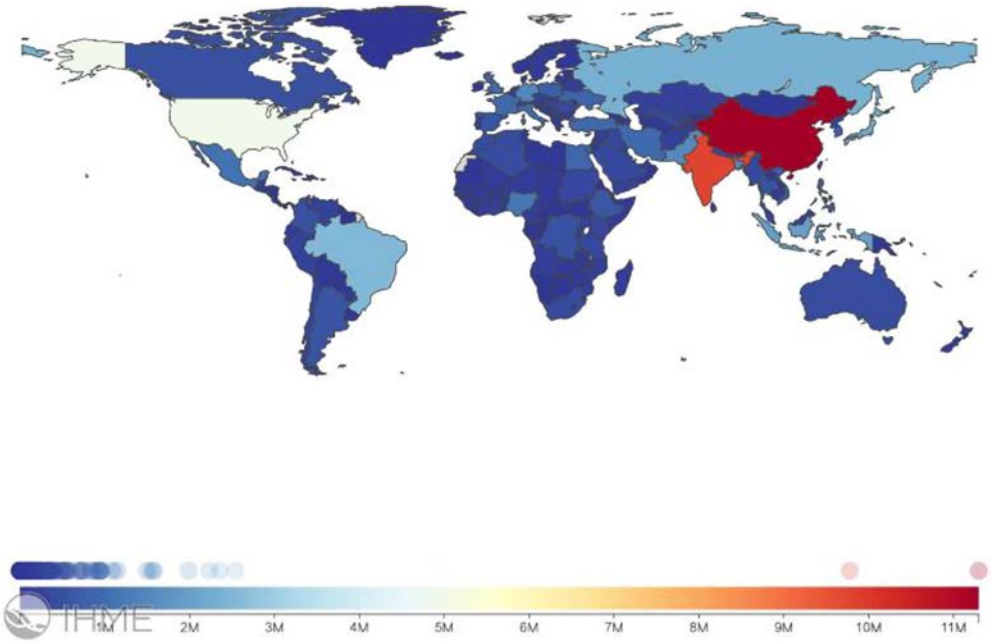

Figure 1. Global distribution of the estimated number of years lived with disability caused by low back pain in 2021 presented as all ages and both sexes.

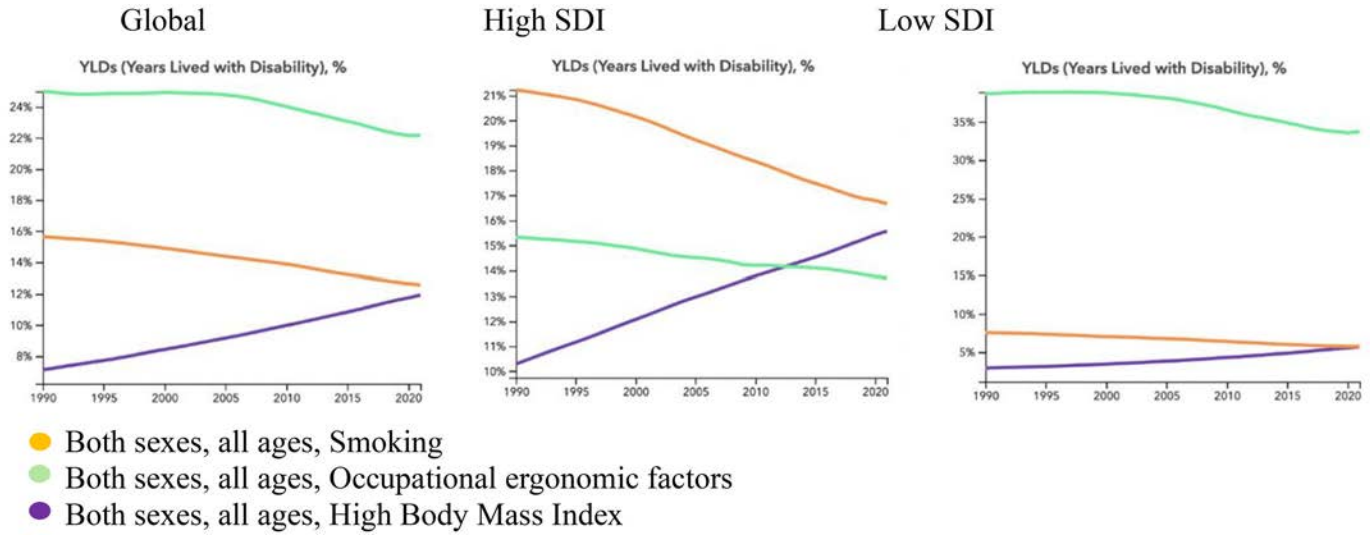

Figure 2. Estimated trends in YLDs caused by occupational/ergonomic, behavioral (smoking), and metabolic (high body mass index) risk factors presented as percentages from 1990 to 2021. SDI, Sociodemographic Index; YLD, year lived with disability. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25520/abstract.

The global rate of YLDs caused by LBP attributed to occupational/ergonomic risk factors has decreased for both sexes combined and when men and women are considered separately. The percentage of YLDs caused by LBP attributed to occupational/ergonomic risk factors and smoking as a proportion of total YLDs has also decreased (by 12% and 21%, respectively) from 1990 to 2021. In comparison, the percentage of YLDs caused by LBP attributed to high BMI as a proportion of total YLDs increased by 65% from 1990 to 2021. Importantly, the number of YLDs caused by exposure to high BMI also increased markedly from 3,100,000 in 1990 to 8,400,000 in 2021 (Table 1; Figure 4; and Supplementary Figures 4 and 5).

The impact of smoking on LBP is more pronounced for men than it is for women globally. The total number of YLDs caused by LBP for men that were attributable to smoking increased from approximately 4,400,000 in 1990 to 5,900,000 in 2021 (Supplementary Figure 4). In comparison, the YLDs caused by LBP that were attributable to smoking for women increased from approximately 2,400,000 in 1990 to 2,900,000 in 2021 (Supplementary Figure 4). Most notably, the impact of high BMI on YLDs caused by LBP has increased globally since 1990. The rate of YLDs per 100,000 caused by LBP attributed to high BMI for men increased by 45% from 1990 to 2021. This trend is also significant for women in the same period, for whom the rate of

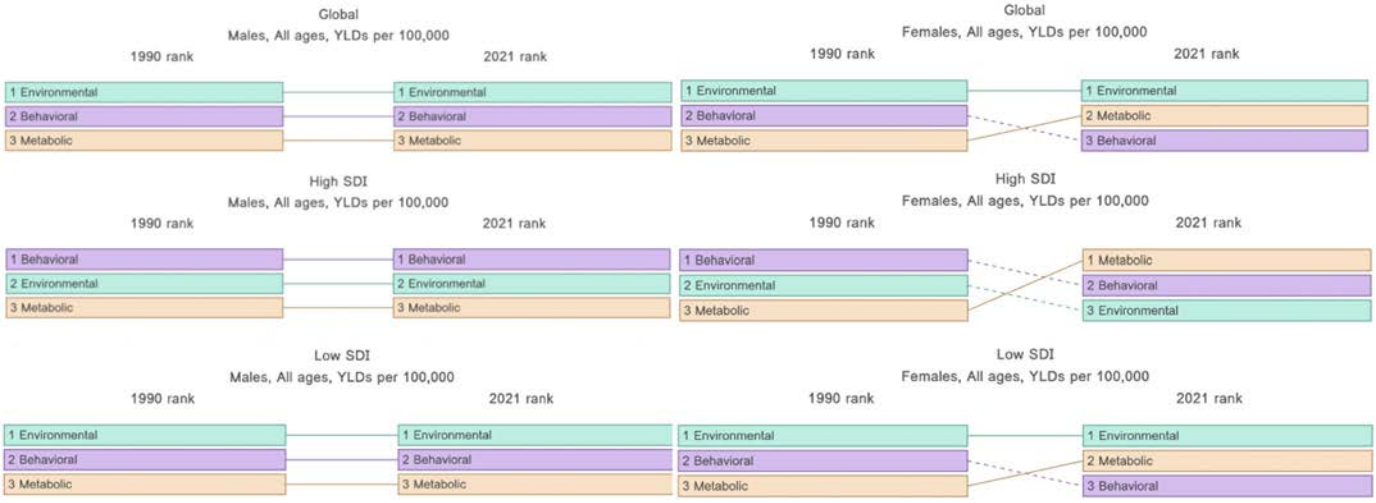

Figure 3. Rank of (occupational/ergonomic), behavioral (smoking), and metabolic (high body mass index) risk factors as an estimated rate of YLDs per 100,000 caused by low back pain in 1990 compared with 2021 for all Global Burden of Disease age groups. SDI, Sociodemographic Index; YLD, years lived with disability. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25520/abstract.

YLDs per 100,000 caused by LBP because of high BMI increased by 37% (Table 1 and Supplementary Figure 4).

Trends in YLDs attributable to occupational/ergonomic, smoking, and high BMI risk factors in high SDI areas. The trends in high SDI areas mimic those of global trends, with YLDs caused by LBP as a proportion of total YLDs decreasing for occupational/ergonomic and smoking risk factors and increasing for high BMI between 1990 and 2021 (Figure 2). In high SDI areas, when considering both sexes combined, the impact of high BMI increased by 36% from 118.84 YLDs per 100,000 in 1990 to 161.80 YLDs per 100,000 in 2021 (Table 2; Supplementary Figure 4). A slight 11% decrease occurred in rate of YLDs per 100,00 caused by occupational/ergonomic risk factors with a more significant decrease of 29% in rate of YLDs per 100,000 caused by smoking for both sexes combined in that timeframe. Notably, in high SDI areas, the rank order differs, and smoking is more impactful than occupational/ergonomic risk factors or high BMI for both sexes combined. This indicates that smoking has a bigger impact on the risk of YLDs caused by LBP in high SDI areas than low SDI areas or globally when considering both sexes combined. In 2012/2013, the impact of high BMI surpassed that of occupational/ergonomic factors and the rank order changed, with smoking as the most impactful risk factor followed by high BMI and then occupational/ergonomic risk factors. This trend in the ranking of risk factors remained stable from 2013 to 2021 (Table 1 and Supplementary Figure 4).

When considering men alone, a decrease is observed of 8% and 25% of total YLDs caused by exposure to occupational/ergonomic risk factors and smoking, respectively, and an increase of 53% of total YLDs caused by exposure to high BMI between 1990 and 2021. In comparison, the YLDs caused by LBP attributed to smoking in women decreased by 24% and only 0.1% for occupational/ergonomic risk factors between 1990 and 2021. The rate of YLDs per 100,000 caused by LBP attributed to high BMI has increased from 95.48 in 1990 to 134.88 in 2021, representing an increase of 35% in total YLDs. Of note, the rank order of risk factors for women has changed significantly, with high BMI becoming more impactful throughout the period from 1990 to 2021 (Figure 3). In 1990, smoking was the most impactful risk factor for women followed by occupational/ergonomic risk factors then high BMI. In 1996, high BMI outranked occupational risk factors, and by 2012 high BMI became

Table 1. Percentage change in YLDs caused by LBP/total YLDs, number of YLDs, and rate of YLDs per 100,000 attributed to occupational, smoking, and high BMI risk factors for men, women, and both sexes combined between 1990 and 2021*

Location	Risk factor	Men, %[a]	Women, %[a]	Both sexes, %[a]
Percentage change in the number of YLDs (all ages)				
Global	Behavioral (smoking)	35 (29 to 40)	22 (16 to 27)	30 (25 to 35)
Global	Metabolic (high BMI)	178 (160 to 194)	167 (154 to 179)	171 (157 to 183)
Global	Occupational/ergonomic	37 (32 to 43)	50 (40 to 61)	43 (37 to 50)
High SDI	Behavioral (smoking)	8 (2 to 14)	3 (−3 to 8)	6 (0 to 11)
High SDI	Metabolic (high BMI)	116 (98 to 134)	96 (83 to 111)	103 (88 to 119)
High SDI	Occupational/ergonomic	18 (12 to 24)	23 (14 to 33)	20 (14 to 27)
Low SDI	Behavioral (smoking)	71 (63 to 77)	66 (50 to 85)	69 (62 to 77)
Low SDI	Metabolic (high BMI)	401 (320 to 452)	308 (273 to 334)	333 (289 to 360)
Low SDI	Occupational/ergonomic	86 (79 to 94)	102 (91 to 114)	95 (88 to 102)
Percentage change in the YLDs owing to LBP/total YLDs (age-standardized)				
Global	Behavioral (smoking)	−21 (−24 to −20)	−31 (−34 to −29)	−25 (−27 to −23)
Global	Metabolic (high BMI)	65 (54 to 73)	53 (46 to 60)	57 (48 to 63)
Global	Occupational/ergonomic	−12 (−15 to −9)	−6 (−11 to 0.5)	−9 (−12 to −5)
High SDI	Behavioral (smoking)	−25 (−28 to −22)	−24 (−27 to −21)	−24 (−26 to −21)
High SDI	Metabolic (high BMI)	53 (40 to 63)	46 (37 to 56)	48 (37 to 58)
High SDI	Occupational/ergonomic	−8 (−11 to −5)	0.1 (−6 to 7)	−4 (−7 to 0)
Low SDI	Behavioral (smoking)	−21 (−24 to −18)	−24 (−32 to −16)	−23 (−26 to −20)
Low SDI	Metabolic (high BMI)	129 (92 to 154)	77 (63 to 88)	92 (73 to 102)
Low SDI	Occupational/ergonomic	−15 (−18 to −12)	−12 (−17 to −7)	−14 (−17 to −10)
Percentage change in the rate of YLDs per 100,000 (age-standardized)				
Global	Behavioral (smoking)	−31 (−33 to −29)	−39 (−41 to −36)	−33 (−35 to −32)
Global	Metabolic (high BMI)	45 (36 to 53)	37 (30 to 43)	39 (32 to 45)
Global	Occupational/ergonomic	−23 (−25 to −20)	−16 (−21 to −10)	−19 (−22 to −16)
High SDI	Behavioral (smoking)	−30 (−36 to −27)	−30 (−33 to −27)	−29 (−32 to −26)
High SDI	Metabolic (high BMI)	41 (29 to 52)	35 (26 to 45)	36 (26 to 46)
High SDI	Occupational/ergonomic	−15 (−18 to −11)	−8 (−13 to −1)	−11 (−15 to −8)
Low SDI	Behavioral (smoking)	−26 (−29 to −23)	−29 (−36 to −21)	−27 (−31 to −24)
Low SDI	Metabolic (high BMI)	114 (80 to 137)	66 (53 to 78)	81 (63 to 92)
Low SDI	Occupational/ergonomic	−21 (−24 to −18)	−17 (−22 to −12)	−19 (−22 to −16)

* BMI, body mass index; LBP, low back pain; SDI, sociodemographic index; YLD, year lived with disability.
[a] The data in parentheses are uncertainty intervals.

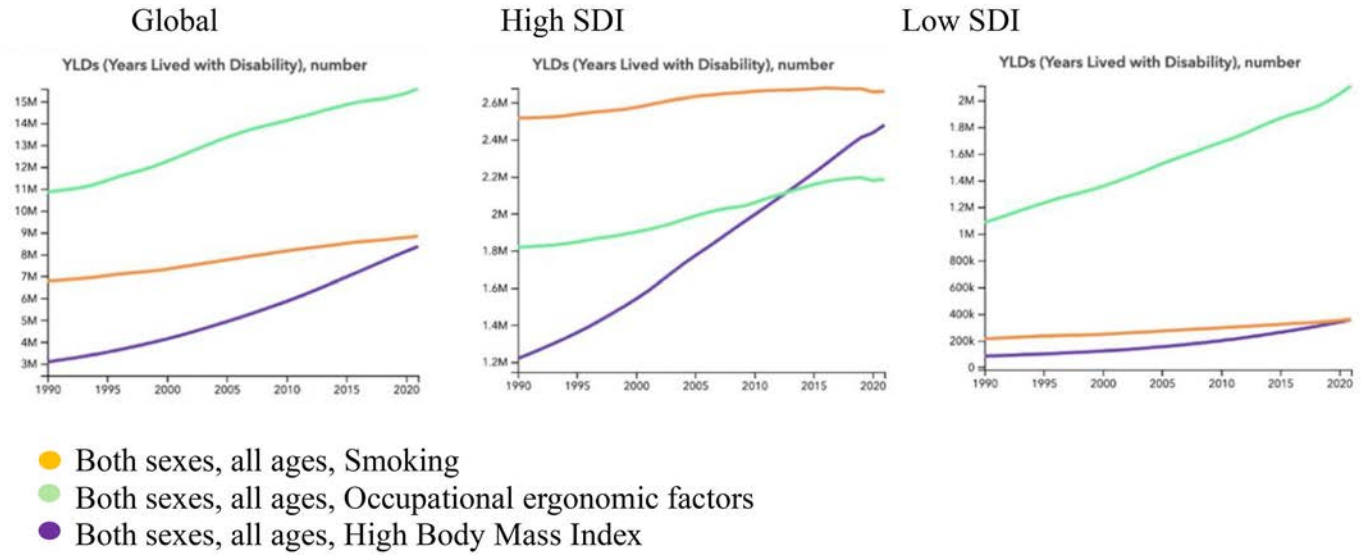

Figure 4. Estimated trends in YLDs caused by occupational/ergonomic, behavioral (smoking), and metabolic (high body mass index) risk factors presented as counts from 1990 to 2021. SDI, Sociodemographic Index; YLD, year lived with disability. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25520/abstract.

the most impactful risk factor for YLDs caused by LBP in women in high SDI areas.

Trends in YLDs attributable to occupational/ergonomic, smoking, and high BMI risk factors in low SDI areas. In low SDI areas, the ranks mirror those of the global trends for both sexes combined (Figure 3). However, when women are considered separately, high BMI becomes more impactful than behavioral risk factors from 1991 to 2021 (Supplementary Figure 6). Similar to the global and high SDI area trends, the number of YLDs caused by LBP attributable to occupational/ergonomic risk factors, smoking, and high BMI increased from 1990 to 2021. However, in contrast to global and high SDI trends, the percentage of YLDs attributable to occupational/ergonomic risk factors did not start to decline until 2005, and the percentage of YLDs attributable to high BMI demonstrated a slow increase from 1990 to 2021 (Figure 2).

The impact of risk factors in low SDI areas for both sexes combined have the same rank order as the global trends: occupational risk factors are ranked first followed by smoking and high BMI (Figure 3). In low SDI areas, occupational/ergonomic risk factors have been the most impactful risk factors associated with YLDs caused by LBP since 1990, declining by a minimal 14% of total YLDs by 2021 (Table 1). The rate of YLDs per 100,000 caused by LBP attributable to occupational risk factors also decreased slightly for both sexes combined. Of note, YLDs caused by LBP attributed to exposure to smoking and high BMI increased throughout the period, resulting in 3,600,000 total YLDs caused by LBP in 2021 for each risk factor. Despite this increase in total YLDs, the impact of smoking as a proportion of total YLDs decreased by 23% when considering both sexes combined.

Of particular significance, the impact of high BMI as a proportion of total YLDs increased by 92% by 2021 in both sexes combined. Considering women alone, a decrease is seen in the impact of occupational/ergonomic risk factors and smoking of 12% and 24% of total YLDs caused by LBP as a proportion of total YLDs, respectively. The impact of occupational/ergonomic risk factors and smoking on the rate of YLDs per 100,000 similarly decreased by 17% and 29%, respectively. The impact of high BMI on all measures of YLDs caused by LBP increased markedly. Most notably, the number of YLDs caused by LBP that are attributed to high BMI increased by a staggering 308% by 2021 (Table 1 and Supplementary Figure 4).

DISCUSSION

This descriptive study aimed to explore the trends in trajectories of three modifiable risk factors for LBP (occupational/ergonomic, behavioral [smoking], and metabolic [high BMI]) from 1990 to 2021, using data from the IHME and GBD study. Significant changes were observed over time of the impact of these risk factors on the total YLDs caused by LBP, the rate of YLDs per 100,000, the YLDs caused by LBP as a proportion of total YLDs owing to all causes, and, importantly, the rank order of these risk factors.

Although occupational/ergonomic risk factors are the most impactful risk factors for LBP globally and in low SDI areas, a decline in the percentage of YLDs owing to occupational/ergonomic risk factors has been seen worldwide. Of note, this decline was not observed in low SDI areas until the early 2000s. In addition, a decrease in average annual percent change in DALY rates and age-standardized disability rates since 1990 that has been seen in high SDI areas has not been seen in low SDI

Table 2. Years lived with disability caused by low back pain presented as counts (millions), percentage (age-standardized), and rate (age-standardized) for global, high SDI, and low SDI areas attributed to occupational, behavioral (smoking), and metabolic (high BMI) risk factors in 1990 and 2021 (both sexes combined)*

Risk	Metric[a]	1990[b]	2021[b]
Global			
Metabolic (high BMI)	Number	3.09 (0.31–6.48)	8.36 (0.84–17.42)
Metabolic (high BMI)	Percent	7.49 (0.79–14.72)	11.74 (01.22–22.84)
Metabolic (high BMI)	Rate	70.22 (7.14–146.48)	97.66 (9.78–204.00)
Occupational/ergonomic	Number	10.85 (7.60–14.54)	1.57 (11.03–20.91)
Occupational/ergonomic	Percent	24.18 (22.35–25.88)	22.07 (20.34–23.69)
Occupational/ergonomic	Rate	226.82 (158.99–304.34)	183.82 (129.49–247.24)
Behavioral (smoking)	Number	6.78 (4.07–10.07)	8.82 (5.18–13.13)
Behavioral (smoking)	Percent	16.34 (11.13–21.32)	12.26 (8.16–16.33)
Behavioral (smoking)	Rate	153.22 (91.37–226.59)	102.04 (60.03–152.10)
High SDI			
Metabolic (high BMI)	Number	1.22 (0.12–2.55)	2.48 (0.25–5.09)
Metabolic (high BMI)	Percent	9.99 (1.03–19.78)	14.80 (1.55–28.27)
Metabolic (high BMI)	Rate	118.84 (11.60–248.80)	161.80 (15.99–332.59)
Occupational/ergonomic	Number	1.82 (1.27–2.45)	2.19 (1.56–2.95)
Occupational/ergonomic	Percent	15.68 (14.47–16.82)	15.12799 (14.10–16.23)
Occupational/ergonomic	Rate	186.74 (129.62–252.22)	165.69 (118.62–222.72)
Behavioral (smoking)	Number	2.52 (1.50–3.74)	2.66 (1.56–3.99)
Behavioral (smoking)	Percent	20.77 (14.02–27.24)	15.88 (10.35–21.31)
Behavioral (smoking)	Rate	247.25 (147.56–367.62)	173.83 (101.74–261.13)
Low SDI			
Metabolic (high BMI)	Number	0.08 (0.01–0.17)	0.36 (0.04–0.73)
Metabolic (high BMI)	Percent	3.22 (0.36–6.10)	6.16 (0.62–11.91)
Metabolic (high BMI)	Rate	27.81 (3.09–55.76)	50.30 (5.07–103.18)
Occupational/ergonomic	Number	1.08 (0.77–1.45)	2.11 (1.47–2.84)
Occupational/ergonomic	Percent	38.38 (35.72–40.91)	33.04 (30.52–35.33)
Occupational/ergonomic	Rate	331.67 (238.02–442.90)	269.53 (191.97–364.56)
Behavioral (smoking)	Number	0.21 (0.12–0.32)	0.36 (0.20–0.55)
Behavioral (smoking)	Percent	8.69 (5.70–11.74)	6.69 (4.29–9.08)
Behavioral (smoking)	Rate	75.09 (44.42–112.87)	54.56 (31.26–83.63)

* BMI, body mass index; SDI, sociodemographic index.
[a] Number is presented as counts in millions for all ages. Percentages and rates are age-standardized.
[b] The data in parentheses are uncertainty intervals.

areas,[16] suggesting that low SDI areas have been slower to respond to the burden of musculoskeletal disorders,[16] potentially explaining this delay.

In high SDI areas, the number of YLDs that can be attributed to smoking is a decreasing trend that has been occurring since 2011, and the percentage of YLDs caused by LBP as a proportion of total YLDs and the rate of YLDs per 100,000 that can be attributed to smoking has decreased globally since 1990, demonstrating the success that global policy can have in modifying behavior. The World Health Organization Framework Convention on Tobacco Control was enforced in 2005,[17] with a large number of countries globally experiencing their most significant reductions in age-standardized prevalence of smoking between 2005 and 2009[17] despite disparities still existing in global taxation and legislation.[3] A positive correlation between smoking and SDI has been reported,[12,18] which may reflect the significant use of smokeless tobacco in low SDI areas[18] and the ongoing challenge of changing lifestyle behaviors globally.[18] Furthermore, increasing global populations and population ageing[18] are resulting in increased numbers of people smoking and increased YLDs owing to LBP attributed to smoking, despite a decrease in the global prevalence of smoking.[17]

The most concerning trend identified is the rapidly increasing impact of high BMI globally on YLDs caused by LBP with an increase of 171% globally, 103% in high SDI areas and 333% in low SDI areas. Globally, the DALYs attributed to high BMI increased from 33.1 million in 1990 to 70.7 million in 2017,[19] which is likely because of increasing and aging populations[19] as well as increasingly sedentary lifestyles, increased caloric intake,[3] and increased life-stress.[19] The risk-weighted exposure for high BMI is noted to increase with increasing SDI,[20] with summary exposure values increasing by >40% between 1990 and 2016.[18] The rapidly increasing impact of high BMI on YLDs caused by LBP in high SDI areas is particularly worrisome as all-cause DALYs attributable to high BMI have remained stable in high SDI areas over this same period of time.[21]

In 2021, the impact of metabolic risk factors on YLDs caused by LBP was markedly more significant for women than for men, with the discrepancy likely to be multifactorial. Obesity is more prevalent in women than men globally[22,23] and is more common in older age groups.[22] Women may be more susceptible to psychopathology associated with obesity and are up to two times more likely to report anxiety or affective disorders associated with

obesity,[24] as well as decreased health-related quality of life,[25] compared with men. Increased abdominal and visceral fat associated with hormonal changes that occur during perimenopause and menopause may also be important in this relationship,[26] as chronic low-grade inflammation is specifically associated with abdominal and visceral adiposity.[27] Furthermore, low levels of physical activity have been found to be associated with an increased incidence of radiating LBP in people with obesity (odds ratio [OR] 3.3, 95% confidence interval [CI]1.01–10.4).[28] In addition, the bidirectional relationship among high BMI, low physical activity, and LBP is likely to be important.

Increased awareness and intervention to mitigate risks, as well as technology and automation and a shift away from manufacturing and agriculture, has likely driven the decreasing impact of occupational/ergonomic risk factors on YLDs caused by LBP.[20] Globalization has led to increased manufacturing in lower SDI areas, which may have decreased the impact of occupational risk factors on YLDs caused by LBP in high SDI areas while increasing the impact in low SDI areas. In high SDI areas, fewer YLDs caused by LBP can be attributed to occupational/ergonomic risk factors than to smoking.

Low physical activity is almost certainly an important risk factor that is contributing to the increase in YLDs caused by LBP globally, and it would be appropriate for the IHME to consider inclusion of low physical activity as a risk factor for musculoskeletal conditions, including LBP, in future GBD releases. There is global recognition of an increasing epidemic of sedentary behavior partly associated with changing occupational and leisure time activities.[29,30] Alzrahani et al performed two systematic reviews reporting an inverse relationship between physical activity and LBP, with moderate physical activity associated with lower LBP prevalence,[31] and >3 hours per day of sedentary behavior associated with increased LBP disability.[32] Psychological state is also recognized as a potential risk factor for LBP,[33–35] and a history of LBP may be associated with recurrences of LBP,[36] both of which warrant further attention on a global scale.

Although the impact of occupational/ergonomic risk factors on LBP is a decreasing trend, it warrants continued attention, especially in low SDI areas. Lifting, forceful movements, vibrations, and awkward postures are most commonly associated with manufacturing, farming, and manual labor,[37,38] however it is still unclear whether LBP is more prevalent in rural or urban settings.[39] Recent evidence also suggests psychosocial aspects, such as high stress levels,[40–42] may also be important contributors to LBP associated with occupational/ergonomic risk factors.[39]

The mechanisms through which high BMI impacts LBP are multifactorial with systemic inflammation,[43,44] increased mechanical stress,[45,46] metabolic syndrome,[45] and deconditioning associated with low physical activity[47,48] thought to contribute to the development of LBP. People with a high BMI commonly report higher disability levels associated with their LBP,[49] with the odds of experiencing high levels of LBP disability associated with high BMI being 1.39 (95% CI 1.15–1.68) in men and 1.45 (95% CI 1.29–1.62) in women.[49] Moreover, men and women who are obese are more likely to seek medical care for their LBP (OR 1.56, 95% CI 1.46–1.67)[50] compared with people with a healthier BMI. The impact of high BMI on the chronicity of LBP, disability associated with LBP, and care-seeking for LBP[49,50–52] suggests that lifestyle programs focusing on maintaining a healthy BMI could potentially be important factors in managing the global burden associated with LBP.[53]

There appears to be a bidirectional relationship between smoking and LBP,[54] with smoking impacting LBP via central processing,[54] altered pain processing, and impaired oxygen delivery[55] with further possible links via low mood, socioeconomic status, and opioid use.[55] A recent meta-analysis reported that current smokers are 30% more likely to suffer from chronic disabling LBP than nonsmokers.[56] Smokers with chronic pain also report increased pain intensity,[54,57] poorer function, worse mental health and mood,[54] worse pain interference,[54] and greater opioid use[57] compared with nonsmokers.[54,58] Smoking cessation is likely to improve musculoskeletal health,[59] day-to-day functional capacity,[60] and the incidence of LBP.[61]

The LBP series published in *The Lancet*[4–6] made a call for LBP management to focus on healthy lifestyles, suggesting that lifestyle programs that focus on these modifiable lifestyle behaviors may be important for managing the global burden associated with LBP. Understanding the trends in the trajectories of these modifiable lifestyle risk factors for LBP may also provide important information for guiding global conversation and policy to address the burden associated with LBP. Strategies may need to vary between areas of different SDIs, addressing differences in population and access to resources as well as society and culture.[6] Understanding the impact and trends of these lifestyle risk factors in different SDI areas may, therefore, aid in directing focus for more effective policy changes and priorities globally.

To our knowledge, this is the first study describing the trends in three modifiable risk factors for LBP from 1990 to 2021, which is a significant strength of this study. The large volume of data generated by the GBD study and the 29 years represented in the data are further strengths of this study. However, the data generated by the GBD study are modeled estimates, which may be seen as a limitation of this study.[62] The LBP prevalence data are primarily obtained from high-income countries, which may impact the generalizability of the results.[1] However, we did not discuss subnational data or specific countries in an effort to maximize the completeness of data used in the trend visualizations. Furthermore, although temporal changes in GBD data are thought to be potentially unreliable,[62] we focused on a general trend analysis rather than specific changes by year. Data from low SDI countries are scarce in the GBD analyses, impacting the strength of conclusions from low SDI areas,[63] and conclusions should therefore be considered cautiously. Additionally, a large

component of risk of LBP is unaccounted for in GBD data, and additional factors such as low physical activity, if measured, may influence the rank order of modifiable factors.

The burden of LBP continues to increase globally, with a significant portion of the YLDs caused by LBP attributed to three modifiable lifestyle factors: occupational/ergonomic, smoking, and high BMI. The number of YLDs caused by LBP continues to increase globally despite the percentage of total YLDs decreasing, which is reflective of an expanding and aging global population. Exposure to occupational risk factors remains the biggest contributor to YLDs caused by LBP globally, but of greater concern is the rapidly increasing impact of high BMI, especially among women. The global increase in BMI is likely related to increases in sedentary lifestyles, which may further impact YLDs caused by LBP. Local and global strategy and policy urgently need to consider the management of the modifiable lifestyle risk factors of occupational/ergonomic, behavioral (smoking), and especially metabolic (high BMI) risk factors to address the growing global burden associated with LBP.

ACKNOWLEDGMENTS

The authors would like to acknowledge the IHME and the GBD study for providing the data visualization tools used for this descriptive analysis. The authors would like to acknowledge the funding support provided by the University of Sydney, the Australian NHMRC, and the Bill and Melinda Gates Foundation. 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 Roberts 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. Low Back Pain Collaborators GBD; GBD 2021 Low Back Pain Collaborators. Global, regional, and national burden of low back pain, 1990-2020, its attributable risk factors, and projections to 2050: a systematic analysis of the Global Burden of Disease study 2021. Lancet Rheumatol 2023;5(6):e316–e329.
2. Wu A, March L, Zheng X, et al. Global low back pain prevalence and years lived with disability from 1990 to 2017: estimates from the Global Burden of Disease study 2017. Ann Transl Med 2020;8(6):299.
3. Risk Factors Collaborators GBD; GBD 2019 Risk Factors Collaborators. Global burden of 87 risk factors in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease study 2019. Lancet 2020;396(10258):1223–1249.
4. Hartvigsen J, Hancock MJ, Kongsted A, et al; Lancet Low Back Pain Series Working Group. What low back pain is and why we need to pay attention. Lancet 2018;391(10137):2356–2367.
5. Foster NE, Anema JR, Cherkin D, et al; Lancet Low Back Pain Series Working Group. Prevention and treatment of low back pain: evidence, challenges, and promising directions. Lancet 2018;391(10137):2368–2383.
6. Buchbinder R, van Tulder M, Öberg B, et al; Lancet Low Back Pain Series Working Group. Low back pain: a call for action. Lancet 2018;391(10137):2384–2388.
7. Risk Factors Collaborators GBD; GBD 2021 Risk Factors Collaborators. Global burden and strength of evidence for 88 risk factors in 204 countries and 811 subnational locations, 1990-2021: a systematic analysis for the Global Burden of Disease study 2021. Lancet 2024;403(10440):2162–2203.
8. Diseases Injuries Collaborators GBD; GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease study 2019. Lancet 2020;396(10258):1204–1222.
9. Stevens GA, Alkema L, Black RE, et al; The GATHER Working Group. Guidelines for accurate and transparent health estimates reporting: the GATHER statement. Lancet 2016;388(10062):e19–e23.
10. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372(71):n71.
11. Hoy D, March L, Brooks P, et al. The global burden of low back pain: estimates from the Global Burden of Disease 2010 study. Ann Rheum Dis 2014;73(6):968–974.
12. Risk Factor Collaborators GBD; GBD 2017 Risk Factor Collaborators. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease study 2017. Lancet 2018;392(10159):1923–1994.
13. Global Burden of Disease Collaborative Network. Global Burden of Disease study 2019 (GBD 2019) Sociodemographic Index (SDI). Institute for Health Metrics and Evaluation. Published 2020. Accessed January 10, 2025. https://doi.org/10.6069/D8QB-JK35
14. Global Health Data Exchange. Global Burden of Disease study 2021 (GBD 2021) results. Institute for Health Metrics and Evaluation. Accessed July 1, 2024. https://vizhub.healthdata.org/gbd-results/.
15. Global Health Data Exchange. GBD compare data visualization. Institute for Health Metrics and Evaluation. Accessed July 30, 2024. https://vizhub.healthdata.org/gbd-compare/
16. 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.
17. Tobacco Collaborators GBD; GBD 2019 Tobacco Collaborators. Spatial, temporal, and demographic patterns in prevalence of smoking tobacco use and attributable disease burden in 204 countries and territories, 1990-2019: a systematic analysis from the Global Burden of Disease study 2019. Lancet 2021;397(10292):2337–2360.
18. Risk Factors Collaborators GBD; GBD 2016 Risk Factors Collaborators. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2016: a systematic analysis for the Global Burden of Disease study 2016. Lancet 2017;390(10100):1345–1422.
19. Dai H, Alsalhe TA, Chalghaf N, et al. The global burden of disease attributable to high body mass index in 195 countries and territories, 1990-2017: an analysis of the Global Burden of Disease study. PLoS Med 2020;17(7):e1003198.
20. Occupational Risk Factors Collaborators GBD; GBD 2016 Occupational Risk Factors Collaborators. Global and regional burden of

disease and injury in 2016 arising from occupational exposures: a systematic analysis for the Global Burden of Disease study 2016. Occup Environ Med 2020;77(3):133–141.

21. Zhou XD, Chen QF, Yang W, et al. Burden of disease attributable to high body mass index: an analysis of data from the Global Burden of Disease study 2021. EClinicalMedicine 2024;76:102848.
22. Li M, Gong W, Wang S, et al. Trends in body mass index, overweight and obesity among adults in the USA, the NHANES from 2003 to 2018: a repeat cross-sectional survey. BMJ Open 2022;12(12): e065425.
23. Li Y, Zhao D, Wang M, et al. Association between body mass index, waist circumference, and age at natural menopause: a population-based cohort study in Chinese women. Women Health 2021;61(9): 902–913.
24. Desai RA, Manley M, Desai MM, et al. Gender differences in the association between body mass index and psychopathology. CNS Spectr 2009;14(7):372–383.
25. Zhang J, Xu L, Li J, et al. Gender differences in the association between body mass index and health-related quality of life among adults:a cross-sectional study in Shandong, China. BMC Public Health 2019;19(1):1021.
26. Palacios S, Chedraui P, Sánchez-Borrego R, et al. Obesity and menopause. Gynecol Endocrinol 2024;40(1):2312885.
27. Khanna D, Khanna S, Khanna P, et al. Obesity: A chronic low-grade inflammation and its markers. Cureus 2022;14(2):e22711.
28. Shiri R, Solovieva S, Husgafvel-Pursiainen K, et al. The role of obesity and physical activity in non-specific and radiating low back pain: the Young Finns study. Semin Arthritis Rheum 2013;42(6): 640–650.
29. Dempsey PC, Biddle SJH, Buman MP, et al. New global guidelines on sedentary behaviour and health for adults: broadening the behavioural targets. Int J Behav Nutr Phys Act 2020;17(1):151.
30. World Health Organization. Global action plan on physical activity 2018–2030: more active people for a healthier world. World Health Organization; 2018. Accessed August 1, 2024. https://www.who.int/publications/i/item/9789241514187
31. Alzahrani H, Mackey M, Stamatakis E, et al. The association between physical activity and low back pain: a systematic review and meta-analysis of observational studies. Sci Rep 2019;9(1):8244.
32. Alzahrani H, Alshehri MA, Alzhrani M, et al. The association between sedentary behavior and low back pain in adults: a systematic review and meta-analysis of longitudinal studies. PeerJ 2022; 10:e13127.
33. Parreira P, Maher CG, Steffens D, et al. Risk factors for low back pain and sciatica: an umbrella review. Spine J 2018;18(9):1715–1721.
34. Felício DC, Filho JE, de Oliveira TMD, et al. Risk factors for non-specific low back pain in older people: a systematic review with meta-analysis. Arch Orthop Trauma Surg 2022;142(12):3633–3642.
35. Pinheiro MB, Ferreira ML, Refshauge K, et al. Symptoms of depression as a prognostic factor for low back pain: a systematic review. Spine J 2016;16(1):105–116.
36. Krishnamurthy I, Othman R, Baxter GD, et al. Risk factors for the development of low back pain: an overview of systematic reviews of longitudinal studies. Phys Ther Rev 2018;23(3):162–177.
37. Yang F, Di N, Guo WW, et al. The prevalence and risk factors of work related musculoskeletal disorders among electronics manufacturing workers: a cross-sectional analytical study in China. BMC Public Health 2023;23(1):10.
38. Kirkhorn SR, Earle-Richardson G, Banks RJ. Ergonomic risks and musculoskeletal disorders in production agriculture: recommendations for effective research to practice. J Agromedicine 2010;15(3): 281–299.
39. Hurwitz EL, Randhawa K, Torres P, et al. The Global Spine Care Initiative: a systematic review of individual and community-based burden of spinal disorders in rural populations in low- and middle-income communities. Eur Spine J 2018;27(S6)(suppl 6):802–815.
40. Jia J, Zhang M, Cao Z, et al. Prevalence of and risk factors for low back pain among professional drivers: a systematic review and meta-analysis. J Orthop Surg Res 2024;19(1):551.
41. Joseph L, Vasanthan L, Standen M, et al. Causal relationship between the risk factors and work-related musculoskeletal disorders among professional drivers: a systematic review. Hum Factors 2023;65(1): 62–85.
42. Hubner FCL, Telles RW, Giatti L, et al. Job stress and chronic low back pain: incidence, number of episodes, and severity in a 4-year follow-up of the ELSA-Brasil Musculoskeletal cohort. Pain 2024; 165(11):2554–2562.
43. Issa RI, Griffin TM. Pathobiology of obesity and osteoarthritis: integrating biomechanics and inflammation. Pathobiol Aging Age Relat Dis 2012;2(2012).
44. Stienstra R, Duval C, Müller M, et al. PPARs, obesity, and inflammation. PPAR Res 2007;2007:95974.
45. McVinnie DS. Obesity and pain. Br J Pain 2013;7(4):163–170.
46. Janke EA, Collins A, Kozak AT. Overview of the relationship between pain and obesity: what do we know? Where do we go next? J Rehabil Res Dev 2007;44(2):245–262.
47. Arranz LI, Rafecas M, Alegre C. Effects of obesity on function and quality of life in chronic pain conditions. Curr Rheumatol Rep 2014; 16(1):390.
48. Tukker A, Visscher TL, Picavet HS. Overweight and health problems of the lower extremities: osteoarthritis, pain and disability. Public Health Nutr 2009;12(3):359–368.
49. Hussain SM, Urquhart DM, Wang Y, et al. Fat mass and fat distribution are associated with low back pain intensity and disability: results from a cohort study. Arthritis Res Ther 2017;19(1):26.
50. Shiri R, Karppinen J, Leino-Arjas P, et al. The association between obesity and low back pain: a meta-analysis. Am J Epidemiol 2010; 171(2):135–154.
51. Lucha-López MO, Hidalgo-García C, Monti-Ballano S, et al. Body mass index and its influence on chronic low back pain in the Spanish population: a secondary analysis from the European Health Survey (2020). Biomedicines 2023;11(8):2175.
52. Dario AB, Ferreira ML, Refshauge KM, et al. The relationship between obesity, low back pain, and lumbar disc degeneration when genetics and the environment are considered: a systematic review of twin studies. Spine J 2015;15(5):1106–1117.
53. Brady SRE, Hussain SM, Brown WJ, et al. Sat0527 predictors of back pain in middle aged women: data from the Australian Longitudinal Study on Women's Health. Ann Rheum Dis 2016;75(Suppl 2):860. https://doi.org/doi:10.1136/annrheumdis-2016-eular.1499
54. Khan JS, Hah JM, Mackey SC. Effects of smoking on patients with chronic pain: a propensity-weighted analysis on the Collaborative Health Outcomes Information Registry. Pain 2019;160(10):2374–2379.
55. Shi Y, Weingarten TN, Mantilla CB, et al. Smoking and pain: pathophysiology and clinical implications. Anesthesiology 2010;113(4): 977–992.
56. Shiri R, Karppinen J, Leino-Arjas P, et al. The association between smoking and low back pain: a meta-analysis. Am J Med 2010; 123(1):87.e7–e35.
57. Hooten WM, Townsend CO, Bruce BK, et al. Effects of smoking status on immediate treatment outcomes of multidisciplinary pain rehabilitation. Pain Med 2009;10(2):347–355.
58. Weingarten TN, Moeschler SM, Ptaszynski AE, et al. An assessment of the association between smoking status, pain intensity, and

functional interference in patients with chronic pain. Pain Physician 2008;11(5):643–653.

59. US Centers for Disease Control and Prevention. Smoking and tobacco use: benefits of quitting smoking. Published May 15, 2024. Accessed September 15, 2024. https://www.cdc.gov/tobacco/about/benefits-of-quitting.html
60. Lindström D, Sadr Azodi O, Wladis A, et al. Effects of a perioperative smoking cessation intervention on postoperative complications: a randomized trial. Ann Surg 2008;248(5):739–745.
61. Ikeda T, Cooray U, Murakami M, et al. Assessing the impacts of smoking cessation and resumption on back pain risk in later life. Eur J Pain 2023;27(8):973–980.
62. Maher C, Ferreira G. Time to reconsider what Global Burden of Disease studies really tell us about low back pain. Ann Rheum Dis 2022;81(3):306–308.
63. Tamrakar M, Kharel P, Traeger A, et al. Completeness and quality of low back pain prevalence data in the Global Burden of Disease study 2017. BMJ Glob Health 2021;6(5):e005847.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 77, No. 7, July 2025, pp 848–854
DOI 10.1002/acr.25510

Long-Term Outcomes of Children Born to Anti-Ro Antibody–Positive Mothers With and Without Rheumatic Disease

Talia Diaz,[1] Ashely Danguecan,[2] Daniela Dominguez,[3] Andrea Knight,[3] Carl A. Laskin,[4] Deborah M. Levy,[3] Edgar Jaeggi,[3] Melissa Misztal,[3] Piushkumar Mandhane,[5] Theo Moraes,[6] Lawrence Ng,[3] Franklin Silverio,[3] Earl D. Silverman,[3] Elinor Simons,[7] Stuart E. Turvey,[8] Padmaja Subbarao,[9] and Linda T. Hiraki[3]

Objective. The objective of this study was to estimate the prevalence of allergy, and/or neurodevelopmental and autoimmune diagnoses in children born to anti-Ro antibody–positive mothers.

Methods. We conducted a cohort study of children born to anti-Ro antibody–positive mothers observed in the neonatal lupus erythematosus (NLE) clinic at The Hospital for Sick Children. Participants one year of age or older were invited to complete a health status questionnaire. Prevalence of allergic, neurodevelopmental, and autoimmune disease diagnoses was compared between the NLE cohort and the non-NLE population-based CHILD Cohort Study cohort. Descriptive statistics were used for demographics, NLE manifestations, and outcomes. Fisher's exact test compared the prevalence of diagnoses between subgroups. We tested the association between allergies and neurodevelopmental conditions and NLE with logistic regression models. A P-value < 0.006 was considered significant.

Results. We included 321 participants born to anti-Ro antibody–positive mothers. The median age at survey completion was six years, 51% of participants were female, and 50% (n = 162) had NLE. We found no significant difference in any disease prevalence between children with and without NLE manifestations (P = 0.57) or between children born to mothers with and without a rheumatic disease (P = 0.11). Disease prevalence was similar between the NLE and CHILD cohorts (allergic disease 30% vs 22% [P= 0.25], neurodevelopmental conditions 5% vs 2% [P = 0.45], autoimmune disease 4% vs 2% [P = 0.68]).

Conclusion. In a large multiethnic cohort of infants born to anti-Ro antibody–positive mothers, there was no significant difference in the prevalence of allergic, neurodevelopmental, or autoimmune diseases between children with and without NLE or between those born to anti-Ro antibody–positive mothers and a population-based non-NLE cohort.

INTRODUCTION

Neonatal lupus erythematosus (NLE) is an acquired autoimmune syndrome in infants, secondary to the transplacental passage of maternal anti-Ro antibodies.[1] NLE includes a wide range of clinical manifestations: cardiac, cutaneous, hepatic, hematologic, and neurologic. Complete congenital heart block is the most severe NLE manifestation because it is irreversible and may require pacemaker insertion at birth. Complete congenital heart block is present in 2% of children born to women with positive anti-Ro antibodies.[2] The risk of complete congenital heart block increases to 19% for children born to women with a previously affected child.[3] Fetal exposure

Dr Hiraki holds a Tier 2, Canada Research Chair. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

[1]Talia Diaz, MD: The Hospital for Sick Children, Toronto, Ontario, Canada, and Hospital San José, Tecnológico de Monterrey, Monterrey, Nuevo Leon, Mexico; [2]Ashely Danguecan, PhD, Daniela Dominguez, MD, MSc, Deborah M. Levy, MD, MS, Edgar Jaeggi, MD, Melissa Misztal, MHSc, Lawrence Ng, BSc, Franklin Silverio, BSc: The Hospital for Sick Children, Toronto, Ontario, Canada; [3]Andrea Knight, MD, MSCE, Earl D. Silverman, MD, Linda T. Hiraki, MD, ScD: The Hospital for Sick Children and SickKids Research Institute, Toronto, Ontario, Canada; [4]Carl A. Laskin, MD: Mount Sinai Hospital and University of Toronto, Toronto, Ontario, Canada; [5]Piushkumar Mandhane, MD, PhD: University of Alberta, Edmonton, Alberta, Canada, and Canadian Healthy Infant Longitudinal Development Study; [6]Theo Moraes, MD, PhD: The Hospital for Sick Children, Toronto, Ontario, Canada, and Canadian Healthy Infant Longitudinal Development Study; [7]Elinor Simons, MD, PhD: University of Manitoba, Winnipeg, Manitoba, Canada, and Canadian Healthy Infant Longitudinal Development Study; [8]Stuart E. Turvey, MBBS, DPhil: BC Children's Hospital and University of British Columbia, Vancouver, British Columbia, Canada, and Canadian Healthy Infant Longitudinal Development Study; [9]Padmaja Subbarao, MD, MSc: The Hospital for Sick Children and University of Toronto, Toronto, Ontario, Canada, and Canadian Healthy Infant Longitudinal Development Study.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25510.

Address correspondence via email to Linda T. Hiraki, MD, ScD, at linda.hiraki@sickkids.ca.

Submitted for publication April 26, 2024; accepted in revised form December 23, 2024.

SIGNIFICANCE & INNOVATIONS

- Little is known about long-term health outcomes of children born to anti-Ro antibody–positive mothers with and without a rheumatic disease.
- In this large multiethnic cohort of children born to anti-Ro antibody–positive mothers, we did not observe a higher prevalence of allergic, neurodevelopmental, or autoimmune diseases in children compared to population controls.
- This study on longterm outcomes in NLE may inform counseling to affected families and aid physicians who care for these children.

to antimalarials throughout pregnancy may decrease the risk of developing cardiac NLE.[4,5]

Anti-Ro antibodies are present in 0.86% of healthy women, in 40% of patients with systemic lupus erythematosus (SLE), and in 60% to 100% of patients with Sjögren disease.[6,7] Despite the high prevalence of these specific autoantibodies in women with rheumatic diseases, 50% of the children with NLE are born to asymptomatic woman or women with an undifferentiated autoimmune syndrome.[8] Maternal rheumatic disease status has been suggested to be a risk factor for NLE, yet findings are inconsistent.[9]

Previous studies have suggested that children with NLE-associated complete congenital heart block may have an increased risk of allergic, autoimmune, neuropsychiatric, and neurodevelopmental conditions compared to the general population.[10,11] To date, little is known about the long-term outcomes of children without features of NLE born to anti-Ro antibody–positive mothers with or without a rheumatic disease diagnosis.

The aim of this study was to examine the prevalence of allergic, neurodevelopmental, and autoimmune (rheumatic and nonrheumatic) diagnoses in children born to anti-Ro antibody–positive mothers. Secondary aims were to compare the prevalence of allergic, neurodevelopmental, and autoimmune diseases in (1) children born to anti-Ro antibody–positive mothers, with and without NLE manifestations, (2) children born to mothers with and without a rheumatic disease diagnosis, and (3) children born to anti-Ro antibody–positive mothers compared to the population based CHILD Cohort Study participants.

PATIENTS AND METHODS

Study population. *SickKids NLE clinic cohort.* We conducted a cohort study of all patients observed in our NLE clinic at The Hospital for Sick Children (SickKids) born to women with anti-Ro antibodies during pregnancy with and without a rheumatic disease diagnosis. The clinic serves a multiethnic greater Toronto area population. Children included in the study were discharged from the NLE clinic between September 1987 and August 2021. The anti-Ro antibody status of mothers was confirmed in the first trimester of pregnancy or at NLE diagnosis. Anti-Ro antibodies were tested by enzyme-linked immunosorbent assay for most participants. Beginning in June 2016, the majority of anti-Ro antibodies were tested by chemiluminescence immunoassay, enabling quantification of anti-Ro52 and anti-Ro60 titers. Postnatally, children born to these mothers were seen in the NLE clinic at 2, 4, and 12 months of age, irrespective of the presence or absence of NLE manifestations. If they were persistently autoantibody positive and/or symptomatic, they were followed with subsequent rheumatology clinic visits or by their primary care provider. Clinical and laboratory manifestations of NLE were prospectively collected and stored in a dedicated NLE database.

Inclusion criteria. Participants were included in the study if all the following criteria were met: (1) documentation of maternal positive anti-Ro antibodies during pregnancy (2) age ≥one year old, (3) discharged from the NLE clinic, and (4) consented to participate in research at discharge. Unconsented patients were sent an invitation letter by mail or email to participate in the study. In addition, we posted on social media platforms (SickKids' Twitter account) informing NLE graduates of the study. Participants were excluded if there was no documentation of an NLE diagnosis and specific NLE manifestations. This study was approved by the Research Ethics Board at SickKids (REB no. 1000034004). We extracted demographic data on mothers and children, which included the following: sex, self-reported ethnicity, maternal anti-Ro antibody status during pregnancy, maternal health status (diagnoses of rheumatic diseases were provided by their rheumatologists), NLE diagnosis, and specific manifestations from the NLE database and medical charts.

NLE cohort health status questionnaire. A self-administered health status questionnaire for graduated patients with NLE or their parents was designed by three pediatric rheumatologists with experience in NLE (LTH, AK, and TD) and a clinical pediatric psychologist (AD) with experience in neonatal neurodevelopmental follow-up. The questionnaire was designed in REDCap and approved by the Research Ethics Board at SickKids. There was a parent version of the questionnaire for parents of children <18 years of age and a participant version for graduated patients with NLE ≥18 years of age. The questionnaire included the following: (1) sociodemographic information; (2) presence or absence of allergic, neurodevelopmental, and autoimmune (rheumatic and nonrheumatic) conditions; (3) family history; and (4) disclosure of personal health information. The questionnaire was sent by email to participants, followed by up to three biweekly reminders if necessary. If there was no response after three reminders, participants were considered as lost to follow-up and they were excluded from the study.

CHILD Cohort Study. The CHILD Cohort Study is a multicenter, prospective, longitudinal, population-based birth cohort study following children over time from maternal pregnancy to adolescence (https://childstudy.ca/). The CHILD Cohort Study includes participants from Vancouver, Edmonton, Manitoba, and Toronto, Canada. The CHILD Cohort Study collects children's and parents' health data via self-reported questionnaire. We focused on the self-administered health status questionnaire for five-year-old children regarding allergy, neurodevelopmental, and autoimmune (rheumatic and nonrheumatic) diagnoses that were similar to those included in our NLE cohort health status questionnaire. We included children with a completed five-year child health questionnaire. Children born to mothers with a rheumatic disease diagnosis were excluded from this comparator group.

Exposure of interest. Our exposures of interest were (1) NLE diagnosis (present or absent) and (2) maternal rheumatic disease during pregnancy (present or absent). NLE was defined as one or more of the following manifestations: (1) cardiac (conduction [congenital atrioventricular block] and/or myocardial disease), (2) cutaneous, (3) hepatic, (4) hematologic, and (5) neurologic. Maternal rheumatic disease status during pregnancy (present or absent) was extracted from the NLE database.

Outcomes. The outcomes of interest were the presence of diagnoses categorized in four groups of disorders as obtained from the NLE clinic health status and CHILD questionnaires: (1) allergic diseases (asthma, eczema, urticaria, and seasonal allergies), (2) neurodevelopmental conditions (autism and attention deficit hyperactivity disorder [ADHD]), (3) autoimmune rheumatic diseases (arthritis, SLE, vasculitis, and other), and (4) autoimmune nonrheumatic diseases (type 1 diabetes, inflammatory bowel disease, psoriasis, thyroid disease, and other).

Statistical analysis. Summary statistics were used for demographic and clinical data. We compared the prevalence of disease diagnoses between (1) children with and without an NLE diagnosis and (2) children born to women with and without a rheumatic disease diagnosis during pregnancy using Fisher's exact test. Secondary analyses compared allergic, neurodevelopmental, and autoimmune disease prevalence between children born to anti-Ro antibody–positive mothers and children from the CHILD Cohort Study across Canada (Edmonton, Toronto, Vancouver, and Winnipeg) and from the Toronto subgroup. The ethnic groups of the NLE SickKids cohort and the CHILD Cohort Study cohort were compared by using the chi-square test. We tested the association of NLE manifestations in the NLE cohort and maternal anti-Ro antibody positivity compared with no anti-Ro antibody exposure in the CHILD cohort with allergy and neurodevelopmental diagnoses using logistic models. The association between these exposures and autoimmune disease was not tested because of limited numbers of patients with autoimmune disease. Multivariable models included covariates for child's sex, child's ethnicity, maternal rheumatic disease status during pregnancy (present or absent), and maternal use of antimalarials during pregnancy (taken or not taken). A corrected P value <0.006 (adjusted for nine independent tests) was used for significance. Data analysis was performed using R version 1.2.5042 (R Foundation for Statistical Computing). This study was approved by the Research Ethics Board at SickKids (REB no. 1000034004).

RESULTS

A total of 711 surveys were sent to the SickKids NLE clinic cohort, of those 321 (45%) surveys were completed (parents 91% [n = 292], patients 9% [n = 29]). Participants were 51% female (n = 163). The median age at survey time was 6.1 years (interquartile range [IQR] 3–11.3 years). Participants' ages at survey time were as follows: <3 years in 22% (n = 70), 3 to ≤5 years in 14% (n = 45), 5 to 10 years in 34% (n = 110), ≥10 to 15 years in 15% (n = 50), and >15 years in 14% (n = 46). The ethnic distribution of the SickKids NLE clinic cohort was as follows: European in 44% (n = 140), East Asian in 17% (n = 53), admixed in 9% (n = 30), South Asian in 8% (n = 27), African in 7% (n = 23), American in 3% (n = 11), other ethnicities (eg, Middle Eastern, First Nations, Pacific Islander) in 2% (n = 6), and missing ethnicity (if the questionnaire was not completed or if the child had one parent with an unknown ethnicity) in 10% (n = 31). Of the SickKids NLE clinic cohort, 50% (n = 162) of the participants had NLE manifestations. The most frequent NLE manifestations were hepatic (49%, n = 80), cutaneous (23%, n = 38), and cardiac involvement (19%, n = 31) (Table 1). Sixty-eight percent (n = 203) of participants were born to a mother with a rheumatic disease diagnosis during pregnancy, with SLE being the most frequently reported diagnosis in 66% (n = 101) of the mothers. Information on antimalarial use during pregnancy was available in 60% (n = 121) of the mothers. Of those mothers, 52% (n = 63) were on antimalarial medication during pregnancy.

Among the NLE clinic cohort, the prevalence of allergic disease at any point over the follow-up period was 30% (n = 84). Within this group of participants with allergy diagnoses, eczema was the most reported (58%, n = 49), followed by seasonal allergies (29%, n = 24) and asthma (23%, n = 19). Neurodevelopmental conditions were reported in 5% (n = 15), comprising ADHD in 66% (n = 10) and autism in 33% (n = 5). Other reported neurodevelopmental conditions included learning disorders in 3% (n = 9), language disorders in 1% (n = 4), global developmental delay in 1% (n = 4), and intellectual disability in 0.7% (n = 2). Rheumatic autoimmune diagnoses were reported in 1% (SLE n = 2 and reactive arthritis n = 1). Nonrheumatic autoimmune diagnoses were reported in 2% (n = 7), with type 1 diabetes in 71% (n = 5) of the patients, vitiligo in 14% (n = 1), and alopecia in 14% (n= 1). Within

Table 1. Child demographic and clinical characteristics of the SickKids NLE clinic (n = 321) and CHILD (n = 3,261) cohorts*

	NLE clinic cohort	CHILD cohort
Female, n (%)	163 (51)	1,638 (47)
Age, median (IQR), y	6 (3–11.3)	5.3 (5.1–5.4)
Ethnicity, n (%)		
European	140 (44)	2,010 (62)
East Asian	53 (17)	231 (7)
Admixed	30 (9)	674 (21)
South Asian	27 (8)	67 (2)
African	23 (7)	44 (1)
American	11 (3)	34 (1)
Other/missing[a]	37 (12)	46 (1)
NLE, n (%)	162 (50)	–
Hepatic	80 (49)	–
Hematologic	51 (31)	–
Cutaneous	38 (23)	–
Cardiac	31 (19)	–
Neurologic	13 (8)	–
Maternal rheumatic disease, n (%)	154 (64)	–
SLE	101 (66)	–
Sjögren disease	28 (18)	–
Rheumatoid arthritis	10 (6)	–
Other	16 (10)	–

* IQR, interquartile range; NLE, neonatal lupus erythematosus; SLE, systemic lupus erythematosus.
[a] "Other" includes Middle Eastern, First Nations, and Pacific Islander; "missing" represents if the questionnaire was not completed or if the child has one parent with unknown ethnicity.

the cohort with anti-Ro antibody–positive mothers, there was no statistically significant difference in disease prevalence of allergic diseases, neurodevelopmental conditions, and/or autoimmune (rheumatic and nonrheumatic) diagnoses between children with and without NLE manifestations (Figure 1) or between children born to a mother with and without a rheumatic disease (Figure 2). No association was found between NLE and allergy (odds ratio [OR] 0.93, 95% confidence interval [CI] 0.56–1.57, $P = 0.81$) and neurodevelopmental diagnoses (OR 0.80, 95% CI 0.25–2.51, $P = 0.71$) in univariate and multivariate adjusted logistic models (Table 2).

The comparator CHILD Cohort Study included 3,261 participants, with 47% (n = 1,638) female. The CHILD Cohort Study Toronto subgroup included n = 761 participants. The median age at survey time was 5.3 years (IQR 5.1–5.4). The ethnic distribution of the CHILD Cohort Study cohort was as follows: European in 62% (n = 2,010), admixed in 21% (n = 674), East Asian in 7% (n = 231), South Asian in 2% (n = 67), African in 1% (n = 44), American in 1% (n = 34), other ethnicities (eg, Middle Eastern, Pacific Islander) in 1% (n = 46), and missing ethnicity (if the questionnaire was not completed or if the child had one parent with an unknown ethnicity) in 5% (n = 155). The ethnic groups of the SickKids cohort and the CHILD Cohort Study cohort were compared using the chi-square test (4.11×10^{-38}) and Fisher's exact test (5.91×10^{-30}), with significant difference in the distribution of overall ethnicity.

In the CHILD Cohort Study, the most frequently reported allergic disease was eczema (26%, n = 632). Asthma was reported in 7% (n = 154), urticaria was reported in 7% (n = 164), and seasonal allergies were reported in 7% (n = 159). Of the neurodevelopmental conditions, autism was reported in 1% (n = 20) and ADHD was reported in 1% (n = 19). Rheumatic autoimmune diagnoses were reported in 1% (n = 6; IgA vasculitis n = 2; juvenile idiopathic arthritis n = 2; and periodic fever, apthous stomatitis, pharyngitis, adenitis n = 2), and nonrheumatic autoimmune diagnoses were reported in 1% (n = 9; celiac disease n = 5, anti–*N*-methyl-D-aspartate encephalitis n = 1, immune thrombocytopenia n = 2, and hypothyroidism n = 1).

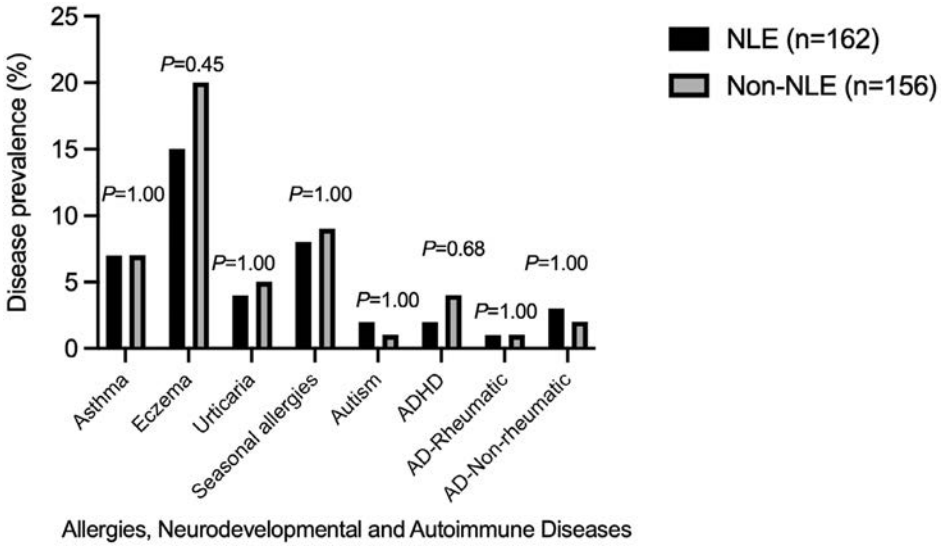

Figure 1. Comparison of disease prevalence in children with and without NLE features within the cohort of children born to anti-Ro antibody positive mothers (n=318). AD, Autoimmune Disease; ADHD, attention deficit hyperactivity disorder; NLE, neonatal lupus erythematosus.

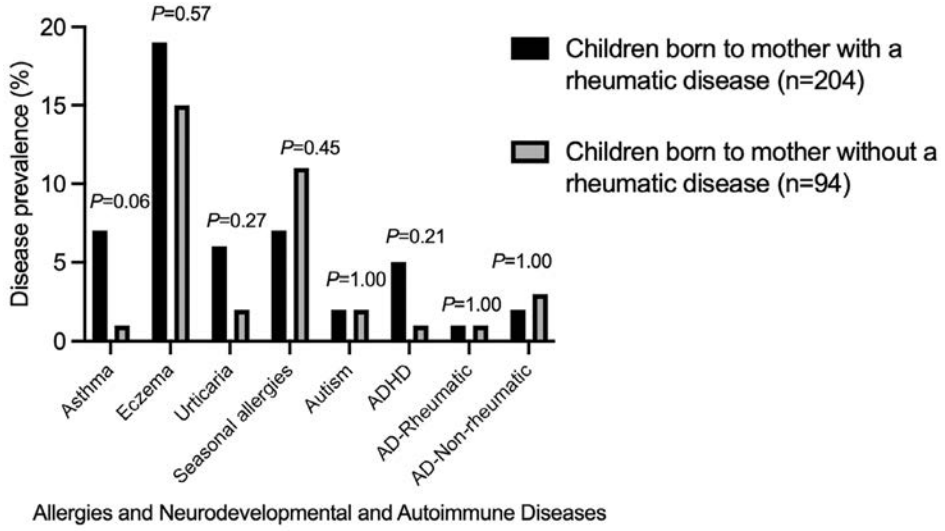

Figure 2. Diseases prevalence in children born to anti-Ro positive antibody mothers with and without a rheumatic disease (n=298). AD, Autoimmune Disease; ADHD, attention deficit hyperactivity disorder.

When comparing children born to anti-Ro antibody–positive mothers and children from the CHILD Cohort Study, we found no significant difference in disease prevalence of allergic diagnoses (asthma 7%, eczema 18%–26%, urticaria 5%–7%, seasonal allergies 7%–9%) or neurodevelopmental diagnoses (ADHD 1%–4% and autism 1%–2%) (Figure 3). Repeat analyses restricted to the Toronto subset of the CHILD cohort did not show a significant difference in prevalence of allergic diseases and neurodevelopmental conditions (Figure 3).

DISCUSSION

In a large multiethnic population of children born to anti-Ro antibody–positive mothers, we did not observe a higher prevalence of allergic, neurodevelopmental, or autoimmune diagnoses among children with NLE compared to those without NLE. We also found no difference in the prevalence of these diagnoses when comparing children born to anti-Ro antibody–positive mothers to children from the Canadian population-based CHILD Cohort Study. Our findings suggest that children born to anti-Ro antibody–positive mothers do not have a higher prevalence of allergic, neurodevelopmental, and/or autoimmune diagnoses compared to the general population. These results provide reassurance to the families of children born to anti-Ro antibody–positive mothers.

Prior studies have investigated the prevalence of neurodevelopmental conditions in children born to mothers with rheumatic disease. In a prospective study, 49 children born to women with SLE were matched with children born to women without SLE by age, sex, race, and socioeconomic status.[12] These children

Table 2. Allergic and neurodevelopmental conditions risk on children born to anti-Ro antibody–positive mothers (n = 287)*

	Allergies, OR (95% CI)	*P* value	Neurodevelopmental conditions, OR (95% CI)	*P* value
NLE	0.93 (0.56–1.57)	0.81	0.80 (0.25–2.51)	0.71
Child ethnicity (referent: European)				
Admixed	1.20 (0.49–2.97)	0.68	–[a]	
American	0.53 (0.10–2.68)	0.44	–[a]	
African	1.35 (0.51–3.60)	0.54	2.95 (0.60–14.36)	0.17
East Asian	1.22 (0.60–2.47)	0.57	1.59 (0.41–6.13)	0.49
South Asian	0.70 (0.25–1.91)	0.48	0.79 (0.08–7.21)	0.83
Other/not available	0.80 (0.29–2.21)	0.67	–[a]	
Female (referent)				
Male	1.38 (0.82–2.31)	0.21	3.76 (1.00–14.1)	0.04
Antimalarials in pregnancy	0.88 (0.49–1.82)	0.73	1.74 (0.36–8.38)	0.48
Antimalarials not available	0.85 (0.46–1.56)	0.61	1.15 (0.28–4.67)	0.84
Maternal rheumatic disease	1.41 (0.76–2.59)	0.26	1.08 (0.24–4.81)	0.91

* CI, confidence interval; NLE, neonatal lupus erythematosus; OR, odds ratio.
[a] Insufficient numbers to generate a point.

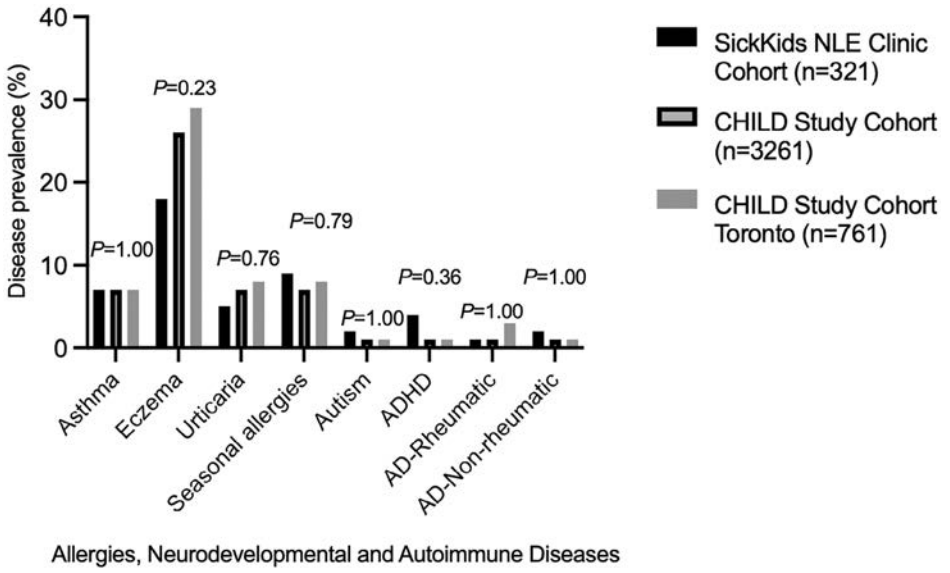

Figure 3. Diseases prevalence of NLE (n = 321) vs. CHILD (n = 3261) vs. CHILD Toronto (n = 761) Cohorts. AD Autoimmune Disease; ADHD, attention deficit hyperactivity disorder; NLE, neonatal lupus erythematosus.

underwent psychological evaluation to determine the prevalence of neuropsychological dysfunction. The authors found greater odds of impairment in learning and memory (OR 3.45 [95% CI 1.25–9.09], P = 0.02) and in behavior (OR 4.00 [95% CI 1.12–14.3], P = 0.03) among children born to women with SLE compared to those born to women without SLE.[12] Although the study observed these statistical differences in neuropsychological testing, the clinical and functional implications are unclear. Of note, the prevalence of neurodevelopmental conditions such as global developmental delay, language disorder, and learning and intellectual disabilities ranged between 1% and 3%, comparable with the general population.[13,14]

Previous studies have suggested that children born to women with SLE may have a higher risk of allergic conditions. A large cohort study by Couture et al[10] showed that there was an increased risk of allergic conditions (asthma, allergic rhinitis, eczema, urticaria, angioedema, anaphylaxis) in 719 children born to mothers with SLE compared to 8,493 matched controls (OR 1.35, 95% CI 1.13–1.61). In another study,[15] using a large Swedish population-based study sample of 775 children born to mothers with SLE and 11,225 children born to mothers without SLE, the risk ratio for childhood asthma was 1.46 (95% CI 1.16–1.84). In contrast, 45% (n = 144) of the children included in our study were born to mothers with SLE; however, we did not find a higher prevalence of allergic conditions in our NLE population compared to population based CHILD cohort.

Our finding of a similar prevalence of autoimmune diagnoses in children born to anti-Ro antibody–positive mothers and the general population also differs from prior studies of long-term outcomes of infants born to mothers with SLE. A large cohort study of 719 children born to mothers with SLE, using *International Classification of Diseases, Ninth and 10th Revision* diagnostic codes to determine the prevalence of autoimmune diseases in children, found that children born to women with SLE as compared to women without SLE (mean age of 9.1 years) had an increased risk of nonrheumatic autoimmune diseases (OR 2.30, 95% CI 1.06–5.03).[16] A limitation of using administrative claims data is the inability to validate the accuracy of the diagnostic codes used to identify cases of disease. Increased vigilance and medical investigation of children born to mothers with SLE, from both parents and care providers, can lead to potential misclassification of these children as having a disease diagnosis. This may result in a higher frequency of billing for diseases that represent investigations rather than confirmed diagnoses. Although 45% of the children included in the SickKids NLE clinic cohort were born to a mother with SLE, we did not find an association between maternal rheumatic disease diagnosis and diagnosis of an autoimmune disease in the child. This may be due to the relatively young age of our study participants (median age at survey was 6.1 years [IQR 3–11.3 years]), since autoimmune diseases more frequently develop during adolescence or young adulthood.

Other studies have found an association between NLE and future autoimmune diseases. In a large cohort study[11] 119 children with complete congenital heart block secondary to NLE, their siblings (n = 128) all born to anti-Ro antibody positive mothers, and were matched by age, sex, month of birth, and region of birth, to healthy controls (n = 1,190) and their siblings (n = 1,071). An increased risk of systemic connective tissue diseases was reported in children with NLE-associated complete congenital heart block (hazard ratio 11.8, 95% CI 4.0–11.8). In the SickKids NLE clinic cohort, 19% (n = 31) of the participants had cardiac NLE manifestations; however, we did not observe a higher

proportion of autoimmune diseases when comparing children with cardiac (0%) versus noncardiac (4%) NLE manifestations.

We acknowledge some limitations of our study. Our power to detect differences in autoimmune and neurodevelopmental conditions was limited given the sample size, the rare nature of these diagnoses, and self-reported diagnoses. Due to the median age of six years among participants, the prevalence of autoimmune diseases was low, as these diseases are typically diagnosed in adolescence or young adulthood. This resulted in limited power to detect a difference in the prevalence of autoimmune diseases. However, we had >99% power to detect a difference in allergic disease prevalence between the NLE and CHILD cohorts if we presumed an allergic disease prevalence of 38% in the CHILD cohort and a previously reported increased relative risk for allergic disease in the NLE cohort of 1.35.

This study has several strengths. Little is known about long-term health outcomes of children born to anti-Ro antibody–positive mothers with and without a rheumatic disease diagnosis. The majority of prior studies examining the prevalence of disease diagnoses focus on children born to mothers with SLE. In our study, we were able to compare the prevalence of diagnoses among children observed in the NLE SickKids clinic cohort, irrespective of maternal disease status, and we also compared our clinic cohort to the CHILD Cohort Study cohort, a Canadian population-based study. Our study provides essential information for parental counseling and education, which may guide disease screening for physicians who care of these children.

In conclusion, in this large multiethnic cohort of children born to anti-Ro antibody–positive mothers, we did not observe a higher prevalence of allergic, neurodevelopmental, or autoimmune diseases in children compared to population controls. Continued follow-up beyond six years of age will provide more information regarding the risk of developing autoimmune diseases in children born to anti-Ro antibody–positive mothers.

AKNOWLEDGMENTS

We thank the CHILD Cohort Study participant families for their dedication and commitment to advancing health research. CHILD was initially funded by The Canadian Institutes of Health Research and AllerGen NCE. Visit CHILD at childstudy.ca.

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. Hiraki 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. Cimaz R, Spence DL, Hornberger L, et al. Incidence and spectrum of neonatal lupus erythematosus: a prospective study of infants born to mothers with anti-Ro autoantibodies. J Pediatr 2003;142(6):678–683.
2. Brucato A, Frassi M, Franceschini F, et al. Risk of congenital complete heart block in newborns of mothers with anti-Ro/SSA antibodies detected by counterimmunoelectrophoresis: a prospective study of 100 women. Arthritis Rheum 2001;44(8):1832–1835.
3. Kan N, Silverman ED, Kingdom J, et al. Serial echocardiography for immune-mediated heart disease in the fetus: results of a risk-based prospective surveillance strategy. Prenat Diagn 2017;37(4):375–382.
4. Barsalou J, Jaeggi E, Laskin CA, et al. Prenatal exposure to antimalarials decreases the risk of cardiac but not non-cardiac neonatal lupus: a single-centre cohort study. Rheumatology (Oxford) 2017;56(9) 1552–1559.
5. Izmirly P, Kim M, Friedman DM, et al. Hydroxychloroquine to prevent recurrent congenital heart block in fetuses of anti-SSA/Ro-positive mothers. J Am Coll Cardiol 2020;76(3):292–302.
6. Keogan M, Kearns G, Jefferies CA. Extractable nuclear antigens and SLE: specificity and role in disease pathogenesis. In: Lahita RG, ed. Systemic Lupus Erythematosus. 5th ed. Academic Press; 2011: 259–274.
7. Satoh M, Chan EKL, Ho LA, et al. Prevalence and sociodemographic correlates of antinuclear antibodies in the United States. Arthritis Rheum 2012;64(7):2319–2327.
8. Vanoni F, Lava SAG, Fossali EF, et al. Neonatal systemic lupus erythematosus syndrome: a comprehensive review. Clin Rev Allergy Immunol 2017;53(3):469–476.
9. Diaz T, Dominguez D, Jaeggi E, et al. Ethnicity and neonatal lupus erythematosus manifestations risk in a large multiethnic cohort. J Rheumatol 2021;48(9):1417–1421.
10. Couture J, Ben-Shoshan M, Pineau CA, et al. Risk of allergic conditions in children born to women with systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2018;70(2):315–319.
11. Mofors J, Eliasson H, Ambrosi A, et al. Comorbidity and long-term outcome in patients with congenital heart block and their siblings exposed to Ro/SSA autoantibodies in utero. Ann Rheum Dis 2019; 78(5):696–703.
12. Urowitz MB, Gladman DD, MacKinnon A, et al. Neurocognitive abnormalities in offspring of mothers with systemic lupus erythematosus. Lupus 2008;17(6):555–560.
13. Shevell M, Ashwal S, Donley D, et al; Practice Committee of the Child Neurology Society. Practice parameter: evaluation of the child with global developmental delay: report of the Quality Standards Subcommittee of the American Academy of Neurology and the Practice Committee of the Child Neurology Society. Neurology 2003;60(3): 367–380.
14. Regier DA, Kuhl EA, Kupfer DJ. The DSM-5: classification and criteria changes. World Psychiatry 2013;12(2):92–98.
15. Rossides M, Nguyen C, Arkema EV, et al. Asthma in children of mothers with systemic lupus erythematosus and the role of preterm birth. Arthritis Care Res (Hoboken) 2018;70(8):1269–1274.
16. Couture J, Bernatsky S, Scott S, et al. Risk of childhood rheumatic and nonrheumatic autoimmune diseases in children born to women with systemic lupus erythematosus. Arthritis Rheumatol 2018; 70(11):1796–1800.

Arthritis Care & Research
Vol. 77, No. 7, July 2025, pp 855–866
DOI 10.1002/acr.25504

American College of Rheumatology
Empowering Rheumatology Professionals

Geographic Clustering of Systemic Sclerosis in Areas of Environmental Pollution

Noelle N. Kosarek,[1] Megan E. Romano,[1] Erika L. Moen,[1] Robert W. Simms,[2] Ashleigh Erickson,[1] Dinesh Khanna,[3] Patricia A. Pioli,[1] and Michael L. Whitfield[2]

Objective. Systemic sclerosis (SSc) is a rare autoimmune disease characterized by fibrosis of the skin and other organs. SSc is thought to arise in genetically predisposed individuals with occupational triggers, although further environmental etiologies still need to be identified. Limited research exists detailing which environmental factors lead to the downstream inflammatory and fibrotic symptoms experienced by patients with SSc across the United States. This study describes a retrospective cohort of 179,188 individuals with an SSc or SSc-related diagnosis code enrolled in the Medicare beneficiary program between the years 2014 and 2018.

Methods. The incidence of SSc and SSc-related diagnosis codes in all US zip codes with beneficiary counts greater than 11 was calculated. We conducted global and local Moran's Index (Moran's I) as well as a hot spot analysis with the Getis Ord Gi statistic to determine whether SSc and SSc-related diagnosis codes exhibited clustered or dispersed patterns across the United States. We identified clusters of SSc and SSc-related diagnosis code with high incidences in or around Superfund sites, which are federally identified areas of environmental contamination.

Results. SSc exhibited clustered patterns in two analyzed cohorts based on global Moran's I statistics of 0.588 and 0.521. Results of local Moran's I indicated clusters of disease in Mississippi, New York, Wisconsin, and Michigan, among others. Some zip codes with high disease incidences were home to at least one Superfund site.

Conclusion. SSc exhibits nonrandom, clustered distributions in a US Medicare beneficiary cohort composed of 179,188 individuals from 2014 to 2018.

INTRODUCTION

Systemic sclerosis (SSc) is a rare autoimmune disease characterized by fibrosis of the skin and internal organs, vascular abnormalities, and autoantibody formation.[1] SSc has the highest case fatality rate of any autoimmune disease: 30% to 40% of individuals with SSc die within 10 years of diagnosis.[2]

Data on the geographic distribution of SSc and related disorders are sparse. SSc prevalence and incidence have been estimated globally. A meta-analysis of 100 studies found a pooled prevalence of 17.6 per 100,000 people and an incidence[3–5] of 1.4 per 100,000 people. The study noted wide regional variation, with North America reporting significantly higher estimates than other areas. SSc geographic distribution has been studied on the nationwide level in other countries such as Denmark, Italy, and the United Kingdom.[6–9] Although the findings of these studies are significant, they are not generalizable to a US population. Similarly, many US studies are limited to specific states or cities and are not generalizable to the greater US population.[10,11] Although, one of these examined the distribution of patients with SSc from two large academic centers in Massachusetts and found a significant enrichment of patients with SSc around hazardous waste facilities and oil release disposal sites. Of the nationwide US studies, researchers have found geographic clusters of high SSc deaths in South Dakota, New Mexico, and Montana, among others.[12] A second nationwide study of SSc provided updated incidence and prevalence for SSc and SSc-associated

Supported by the Scleroderma Research Foundation (to Dr Whitfield). Dr Kosarek's work was supported by the National Institutes of Health Big Data to Knowledge (grant 5T32-LM0-12204) and the Arnold Postlethwaite Scleroderma Foundation Pre-Doctoral Summer Fellowship Award.

[1]Noelle N. Kosarek, PhD, Megan E. Romano, PhD, Erika L. Moen, PhD, Ashleigh Erickson, MPH, Patricia A. Pioli, PhD: Geisel School of Medicine at Dartmouth, Lebanon, New Hampshire; [2]Robert W. Simms, MD, Michael L. Whitfield, PhD: Geisel School of Medicine at Dartmouth and Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire; [3]Dinesh Khanna, MD, MSc, University of Michigan, Ann Arbor.

Additional supplementary information cited in this article can be found online in the Supporting Information section (http://onlinelibrary.wiley.com/doi/10.1002/acr.25504).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25504.

Address correspondence via email to Michael L. Whitfield, PhD, at michael.whitfield@dartmouth.edu.

Submitted for publication May 30, 2024; accepted in revised form January 17, 2025.

SIGNIFICANCE & INNOVATIONS

- SSc exhibits nonrandom, clustered distributions in a US Medicare beneficiary cohort.
- Clusters of SSc and SSc-related diagnosis codes were found across multiple geographic locations. A significant fraction of the zip codes with high rates and risk ratios in clusters of SSc and SSc-related diagnosis codes contained Superfund sites.

interstitial lung disease but did not focus on geographic distributions.[13]

SSc arises in a genetically predisposed individual and has been reported to coincide with occupational triggers such as crystalline silica[14,15] or solvent exposure, although further environmental etiologies still need to be identified. Occupational, bacterial, and viral factors have been cited as potential environmental triggers of SSc. Higher levels of serum silica micro- and nanoparticles were found in individuals diagnosed with SSc compared to healthy controls after occupational exposure to silica dust,[16] and several case reports have documented SSc development after silica dust exposure.[17,18] Trichloroethylene (TCE), which is used as a metal degreasing and dry cleaning solvent, has been reported to induce SSc[19,20] and has been associated with the development of pulmonary fibrosis in mice.[21] Many other organic compounds, such as chlorinated solvents, welding fumes, and vinyl chloride, have been associated with SSc presentation and progression.[22–24]

The potential sources of environmental triggers are extensive and vary by geographic region. Given this, investigating SSc incidence at federally identified sites of hazardous contamination may aid in narrowing the list of exposures associated with SSc presentation. Superfund sites, where severe contamination has occurred because of the improper management of hazardous waste, are designated by the US Environmental Protection Agency (EPA) under the Comprehensive Environmental Response, Compensation, and Liability Act. Many of these sites are associated with former manufacturing facilities, landfills, and mining operations. EPA funding is available to support environmental remediation efforts, as exposure to toxic substances at Superfund sites is linked to cancer incidences in 48 states.[25] Notable examples include the high incidence of bladder cancer observed in residents near the Drake Superfund site in Pennsylvania[26] and the increased incidences of melanoma, lymphoma, kidney, brain, and breast cancers[27] in individuals working at an IBM location near the Endicott Village Well Field (Endicott, NY), which was identified as a source of contaminated drinking water.

Given the contribution of environmental inducers to SSc development and progression, investigation of associations between SSc incidence in the greater United States and near Superfund sites may aid in the identification of specific occupational or environmental regulators of disease. To assess these potential associations, Medicare beneficiary data were used to build two cohorts of Medicare beneficiaries carrying SSc or SSc-related diagnosis codes between the years 2014 and 2018. In this work, we report the spatial distribution of SSc and SSc-related diagnosis codes across the United States and within proximity to Superfund sites.

PATIENTS AND METHODS

SSc Medicare cohort. The Committee for the Protection of Human Subjects (CPHS) at Dartmouth College approved this study (CPHS study number STUDY00032176; institutional review board number IRB00006768). Medicare beneficiaries enrolled in the program between 2014 and 2018 and carrying at least one of the SSc or SSc-related *International Classification of Diseases* (ICD)-9 and/or ICD-10 diagnosis codes listed in Table 1 were used to assemble the two cohorts analyzed in this study. Beneficiaries were included in the cohort if they were the minimum age of 65 years to qualify for Medicare or turned 65 during the 2014 to 2018 capture period. Cohort 1 was composed of individuals with primary SSc ICD-9 and ICD-10 codes, whereas cohort 2 included individuals with a diagnosis code of "history of scleroderma" (ICD-10 Z87.39), which additionally included individuals with a "personal history of other diseases of the musculoskeletal system and connective tissue."

Nationwide descriptive statistics and zip code incidence rates. Descriptive statistics including incidence by age, sex, race, and diagnosis code were generated in SAS Studio v3.8 (SAS Institute, Inc). Incidence rates for all US zip codes were calculated by extracting the number of beneficiaries enrolled in Medicare and the number of beneficiaries in our cohorts in each zip code. Zip codes with cohort 1 or cohort 2 beneficiary counts less than or equal to 11 were suppressed. It is a Centers for Medicare and Medicaid Services rule that counts less than or equal to 11 in individual zip codes must be suppressed to ensure that beneficiaries are not identifiable. The number of cohort beneficiaries was divided by the total number of beneficiaries to obtain incidence rates for SSc and SSc-related diagnosis codes in individual zip codes.

Incidence rates in zip codes containing at least one Superfund site. A list of current, proposed, and retired Superfund sites was compiled from the US EPA Socioeconomic Data and Applications Center. The link to these data can be found at https://sedac.ciesin.columbia.edu/data/set/Superfund-epa-national-priorities-list-ciesin-mod-v2. Risk ratios for cohort 1 and cohort 2 in zip codes with Superfund sites were calculated by dividing the incidence rate for each zip code by the nationwide incidence rate for the diagnosis codes used to assemble cohorts 1 and 2 in the United States. Maps outlining zip code and state boundaries displaying disease incidence were generated in

Table 1. Characteristics of cohort 1 and cohort 2 in Medicare beneficiary data by sex, age, race, diagnosis code, and US region*

Characteristic	Cohort 1 (n = 58,379)	Cohort 2 (n = 179,188)
Sex, n (%)		
Male	18,854 (32.30)	6,4214 (35.84)
Female	39,525 (67.70)	114,974 (64.16)
Age, n (%), yr		
Under 65	2,250 (3.85)	6,496 (3.63)
65–74	26,356 (45.15)	87,971 (49.09)
75–84	20,529 (35.17)	59,666 (33.30)
Over 85	9,244 (15.83)	25,055 (13.98)
Race, n (%)		
Unknown	562 (0.96)	153,317 (85.56)
White	49,065 (84.05)	150,00 (8.37)
Black	5,462 (9.36)	2,563 (1.43)
Other	977 (1.67)	2,769 (1.55)
Asian	1,103 (1.89)	2,517 (1.40)
Hispanic	976 (1.67)	749 (0.42)
North American Native	234 (0.40)	153,317 (85.56)
Diagnosis and ICD codes, n (%)		
Systemic sclerosis (7101)	9,428 (16.15)	9,428 (5.26)
Progressive systemic sclerosis (M340)	460 (0.79)	460 (0.26)
CREST syndrome (M341)	1,829 (3.13)	1,829 (1.02)
Other systemic sclerosis (M3489)	375 (0.64)	375 (0.21)
Systemic sclerosis unspecified (M349)	4,561 (7.81)	4,561 (2.55)
Systemic sclerosis with lung involvement (M3481)	214 (0.37)	214 (0.12)
Lung involvement in systemic sclerosis (5172)	91 (0.16)	91 (0.05)
Systemic sclerosis with myopathy (M3482)	30 (0.05)	30 (0.02)
Systemic sclerosis polyneuropathy (M3483)	385 (0.66)	385 (0.21)
Pulmonary hypertension secondary to scleroderma (I2729)	29,923 (51.26)	29,923 (16.70)
Renal involvement in scleroderma (N08)	11,083 (18.98)	11,083 (6.19)
Personal history of other disease of the musculoskeletal system and connective tissue (Z87.39)	0 (0.00)	120,809 (67.42)
US region, n (%)		
Northeast	13,709 (23.48)	38,202 (21.32)
Southeast	16,985 (29.09)	49,190 (27.45)
Midwest	12,101 (20.73)	44,261 (24.70)
Southwest	6,284 (10.76)	17,635 (9.84)
West	9,300 (15.93)	29,900 (16.69)

* CREST, calcinosis, Raynaud phenomenon, esophageal dysmotility, sclerodactyly, telangiectasias; ICD, *International Classification of Diseases*.

ArcGIS Online. Additional maps found in Supplemental Figures 1 and 2 display the incidence of cohorts 1 and 2 diagnosis codes in relation to SSc treatment centers and Superfund sites, respectively.

Moran's Index. Moran's Index (Moran's I) of global spatial autocorrelation was conducted on the identified cohorts in the United States in Geoda (version 1.22.0.4). Global Moran's I provides an easily interpretable measure of whether data exhibit true clustering. A positive value for global Moran's I indicates that the data are spatially clustered, whereas a negative value suggests a dispersed pattern in the data; random distributions result in values near 0. Likewise, negative Z scores indicate more dispersion, whereas positive Z scores indicate more clustering. A positive global Moran's I is indicative of a less than 1% likelihood that clustering occurs as the result of random chance. We conducted local Moran's I analysis in ArcGIS Online. Local Moran's I indicates which zip codes are clusters of high or low incidences. Zip code boundaries were colored by local Moran's I cluster assignments, which were conducted in ArcGIS Online, with inverse distance weighting for K nearest neighbors based on a chordal approximation to geodetic Euclidean distance. To identify specific hot or cold clusters of disease in cohorts 1 and 2, an optimized hot spot analysis was conducted using the Getis Ord Gi* statistic in ArcGIS Online.

RESULTS

Geographic distribution of cohort 1. We assembled a cohort of 58,379 Medicare beneficiaries, heretofore referred to as cohort 1, carrying at least one ICD-9 and/or 10 code associated with scleroderma (Table 1). Most beneficiaries carried a diagnosis of "systemic sclerosis" (16.15%), "progressive systemic sclerosis" (0.79%), "pulmonary hypertension secondary to scleroderma" (51.26%), or "renal involvement in scleroderma" (18.98%). This cohort of beneficiaries was 67.70% female, 84.05% White, and predominantly from the Northeast, Southeast, and Midwest United States, with a female-to-male ratio of

2.09:1 (Table 1). A detailed summary of cohort 1 distribution by sex, age, and race is reported in Supplemental Tables 1 and 2.

The incidence rate of cohort 1 diagnosis codes in each zip code in the United States was then plotted. As shown in Figure 1A and Supplemental Figure 3A, high incidences of cohort 1 diagnosis codes were noted in the northeast region of Mississippi, the northeast and southern regions of Louisiana, and in New Jersey and Maryland. Global spatial autocorrelation using

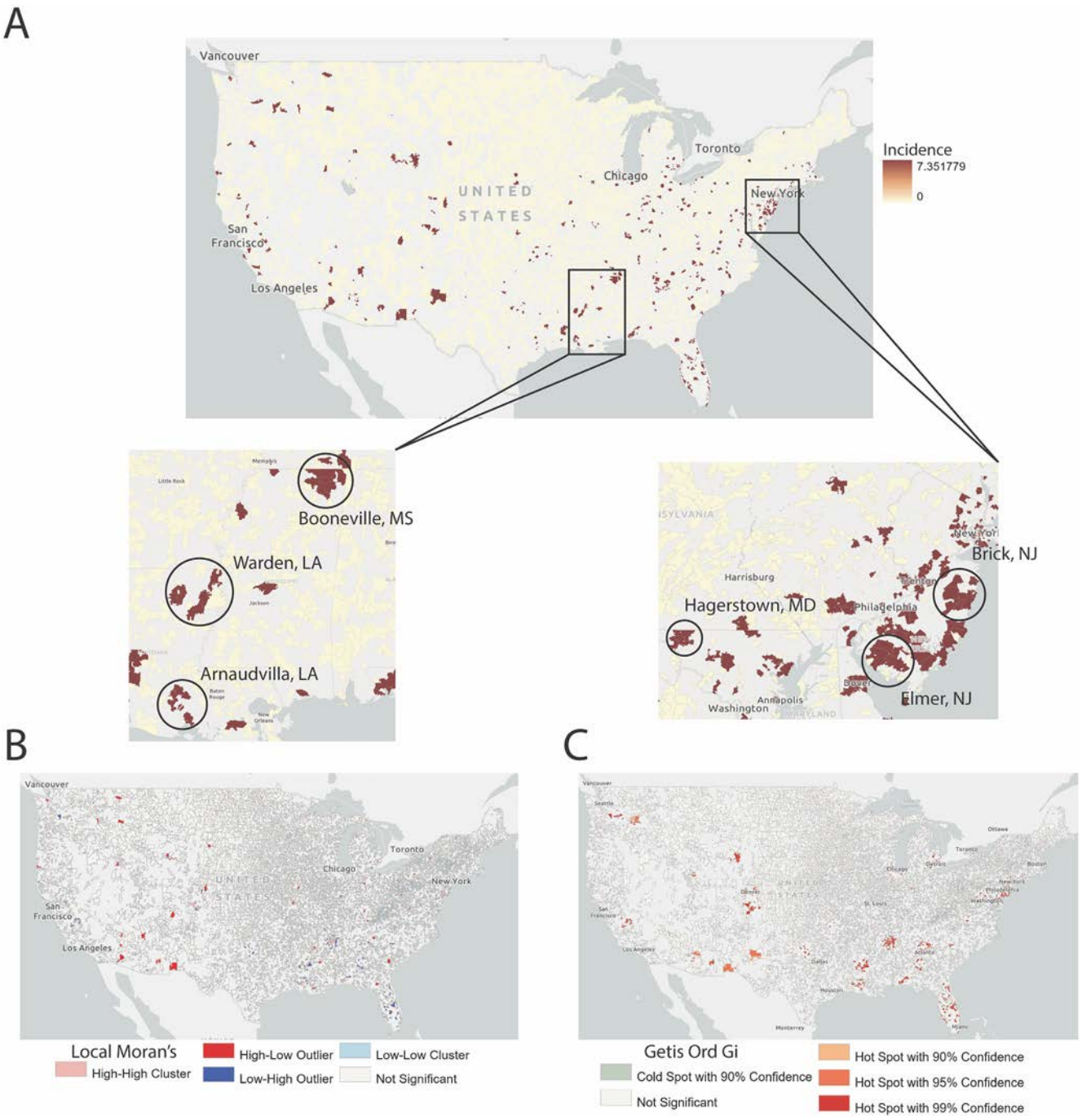

Figure 1. (A) Distribution of the incidence rate of Medicare beneficiaries carrying at least one of the diagnosis codes for cohort 1 (see Table 1) by zip code between the years 2014 and 2018. Zip codes with high incidences are indicated in red and dark orange, whereas zip codes with low incidences are shown in yellow or light orange. Statistically significant clusters were noted in Mississippi, Maryland, and New Jersey. (B) Results of Moran's Index local spatial autocorrelation on cohort 1 geographic data. Zip codes with high incidences surrounded by other zip codes of high incidences are indicated in pink. Dark red indicates zip codes of high incidences surrounded by areas of low incidences. Light blue and dark blue indicate areas of low incidences. (C) Results of optimized hot spot analysis on cohort 1 geographic data. Hot spots of high incidences are shown in varying shades of red, whereas cold spots of low incidences are shown in varying shades of blue.

Moran's I was calculated for cohort 1, resulting in a Moran's I of 0.5880, a Z score of 74.3258, and a *P* value of 0.00 (Supplemental Figure 4). The distribution of cohort 1 beneficiaries was clustered in this cohort and did not exhibit a random dispersion. Cohort 1 clusters were not consistently found at or near treatment center zip codes. Scleroderma specialty treatment centers are notably absent in Mississippi, where distinct clusters in rural areas were located (Supplemental Figure 1A).

To confirm these results, local spatial autocorrelation and optimized hot spot analysis was performed. Zip codes with pink, red, dark blue, or light blue coloration in Figures 1B and C had *P* values below 0.05. As demonstrated in Figure 1B and Supplemental Figure 3B, local spatial autocorrelation identified clusters in New Jersey, Louisiana, and Mississippi. Optimized hot spot analysis confirmed the presence of hot spots in Louisiana and Mississippi (Figure 1C, Supplemental Figure 3C).

To identify potential connections between SSc and Superfund site localization, the incidence of cohort 1 diagnosis codes was plotted in zip codes that contained at least one Superfund site (Figure 2A, Supplemental Figure 3D). Using these parameters, a high incidence of diagnosis codes was noted in zip codes containing at least one Superfund site in New Jersey and Maryland (Figure 2B). These sites included Central Chemical, Upper Valley Deerfield Township Sanitary Landfill, Nascolite Corp, Vineland Chemical Company, and Brick Township Landfill (Figure 2B). Based on the calculated national incidence of cohort 1 diagnosis codes, which was 0.50%, the risks of having an SSc diagnosis code for individuals residing in zip codes containing each of these Superfund sites were 5.66, 3.72, 3.84, 3.44, and

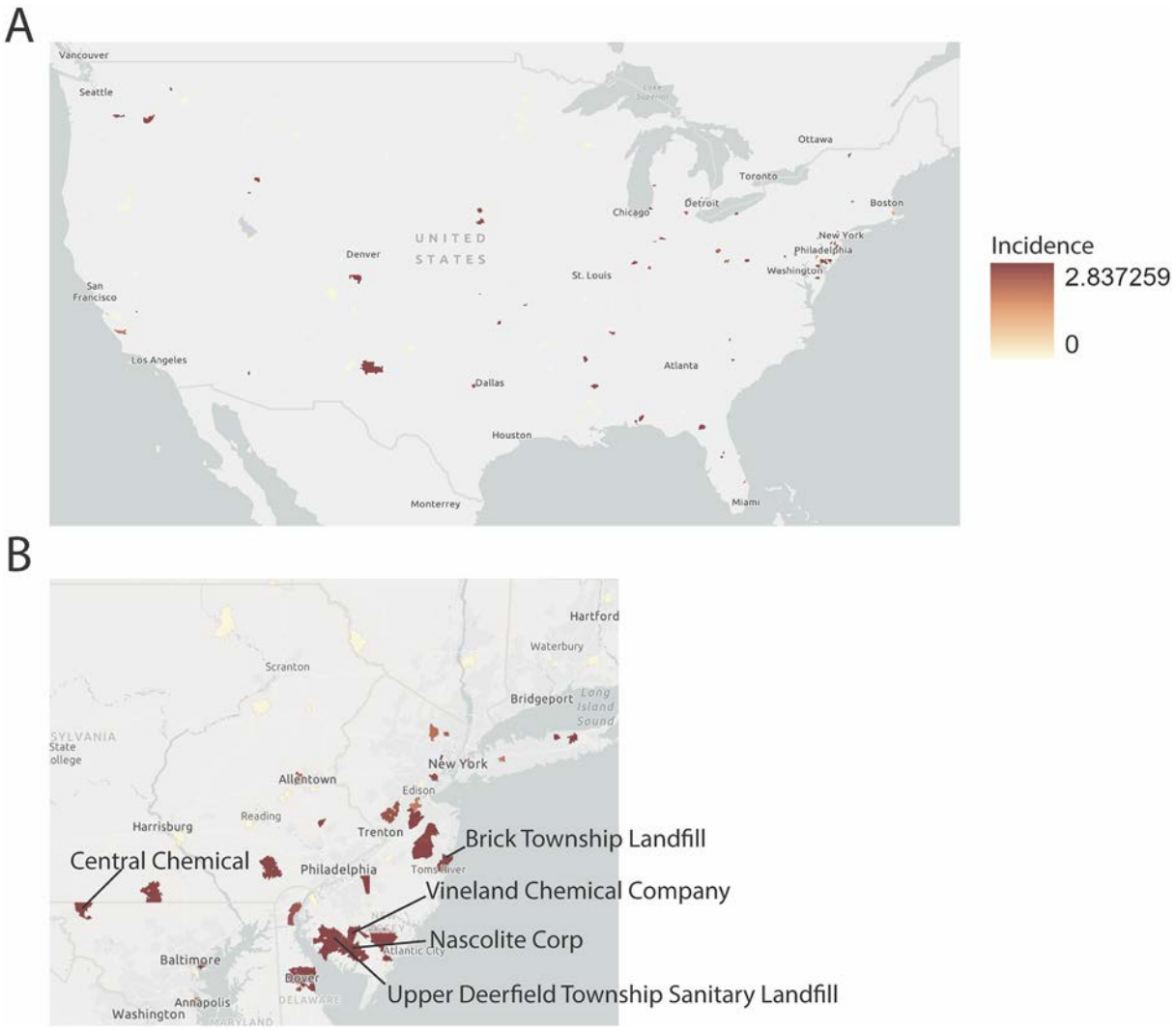

Figure 2. (A) Distribution of the incidence rate of Medicare beneficiaries carrying at least one of the diagnosis codes for cohort 1 (see Table 1) by zip code containing at least one Superfund site between the years 2014 and 2018. Zip codes with high incidences are indicated in red and dark orange, whereas zip codes with low incidences are shown in yellow and light orange. (B) Cohort 1 clusters were noted near Central Chemical, Brick Township Landfill, Vineland Chemical Company, Nascolite Corp, and Upper Deerfield Township Sanitary Landfill.

4.9 times greater, respectively, than the national risk (Supplemental Table 3). A map displaying incidence rates for cohort 1 and locations of Superfund sites can be found in Supplemental Figure 2A.

Distribution of SSc and history of SSc by age, sex, race, and diagnosis code. Cohort 1 does not contain a diagnosis code capturing beneficiaries with a history of SSc. Here we define cohort 2, which includes beneficiaries carrying the diagnosis codes captured in cohort 1 with the addition of the code Z87.39. This code encompasses a history of SSc and related conditions. This code was included in cohort 2 to capture beneficiaries who came into the cohort who had received an SSc diagnosis code before being enrolled in Medicare. These beneficiaries would not have been provided an additional SSc diagnosis code during the catchment period, rather they would be given the history of SSc code. In cohort 2 we identified 179,188 Medicare beneficiaries carrying at least one of the diagnosis codes for this cohort between the years 2014 and 2018, 114,974 beneficiaries were female and 64,214 were male (Table 1). The ratio of women to men was approximately 1.79:1. A detailed breakdown of cohort 2 demographic distribution by sex, age, and race is reported in Supplemental Tables 4 and 5. The majority of the cohort 2 beneficiaries (87,971, 49.09%) were between the ages 65 and 74. The cohort included 6,496 beneficiaries (3.63%) that turned 65 during the 2014 to 2018 period (Table 1).

The racial distribution of this cohort was predominantly White (85.56%). The remaining beneficiaries identified as Black (8.37%), unknown (1.27%), other (1.43%), Asian (1.55%), Hispanic (1.40%), and North American Native (0.42%) (Supplemental Table 5). Thus, cohort 2 was composed of primarily White and Black beneficiaries, which is consistent with previously published literature.[28,29]

Geographic distribution of cohort 2. The geographic distribution of cohort 2 across the United States was assessed for geographic clusters to test the hypothesis that SSc incidence is nonrandom and associated with environmental factors that contribute to disease presentation and progression. Accordingly, the incidence rate for cohort 2 diagnosis codes in each zip code in the United States was plotted (Figure 3A). Examples of locations with high incidences are shown in Figure 3A. These include a location in Georgia; three locations in Mississippi; and distinct singular locations in Ohio, the southeast corner of Wisconsin, and New York. To stratify locations based on relative risk, rates in individual zip codes were compared to the national average. The national incidence for cohort 2 diagnosis codes between 2014 and 2018 was calculated as 1.56%. The relative risks of cohort 2 diagnosis codes in Fond du Lac (Wisconsin), Rootstown (Ohio), Endicott (New York), Milledgeville (Georgia), Booneville (Mississippi), Prentiss (Mississippi), and Columbia (Mississippi) are 4.08, 4.49, 3.36, 7.5, 5.22, 6.09, and 6.19, respectively (Supplemental Table 6).

To determine whether the clusters of high incidences were statistically significant, Moran's I global spatial autocorrelation analysis was conducted using Geoda. Analysis of the clusters identified in Figure 3A resulted in a Moran's I of 0.5212, a Z score of 59.4272, and a *P* value of 0.00 (Supplemental Figure 5). Cohort 2 clustering was highly statistically significant. These results indicated the observed clusters were specific with a nonrandom distribution across the United States. Cohort 2 clusters were not consistently found at or near treatment center zip codes (Supplemental Figure 1B).

Statistically significant low or high clusters were identified using local spatial autocorrelation and optimized hot spot analysis. Pink, light blue, red, or dark blue colors denote statistically significant high or low clusters ($P \leq 0.05$; Figure 3B). The results of the local spatial autocorrelation identified statistically significant clusters of diagnosis codes in Wisconsin, Michigan, Ohio, New York, North Carolina, Georgia, and Mississippi (Figure 3B). Because noise was detected in the local spatial autocorrelation, an optimized hot spot analysis was performed to confirm that these were truly significant clusters. This analysis not only confirmed the presence of hot spots in the states noted previously, but identified additional hot spots in Oklahoma, Louisiana, Indiana, Pennsylvania, and New Hampshire (Figure 3C).

Because environmental factors may influence the development and/or progression of rheumatologic and connective tissue diseases, potential cluster localization near Superfund sites was investigated. A total of 15,551 Medicare beneficiaries carrying a cohort 2 diagnosis code resided in the same zip code as a Superfund site in the years 2014 to 2018. As shown in Figure 4A–D, high rates of cohort 2 diagnosis codes were identified at the Newsome Brothers/Old Reichhold Chemicals, Inc (9.66%), Endicott Village Well Field (5.24%), Picayune Wood Treating (5.17%), Mid-South Wood Products (4.21%), and Refuse Hideaway Landfill (3.81%) Superfund sites. Comparison of calculated risk ratios with national incidence demonstrated that the risks of carrying a cohort 2 diagnosis code at Newsome Brothers/Old Reichhold Chemicals, Inc, Picayune Wood Treating, and Mid-South Wood Products were 9.66, 3.31, and 2.70 times greater, respectively, than the national risk in the Medicare beneficiary population between 2014 and 2018 (Figure 4B–D). A map displaying incidence rates for cohort 2 and locations of Superfund sites can be found in Supplemental Figure 2B.

DISCUSSION

Consistent with many autoimmune diseases, there is an increased incidence of SSc in women compared with men; previous studies have reported female-to-male ratios[9,30,31] of 3:1 to 7.8:1. In these study cohorts (Table 1), female-to-male ratios of 2.09:1 in cohort 1 and 1.79:1 in cohort 2 were observed. Notably,

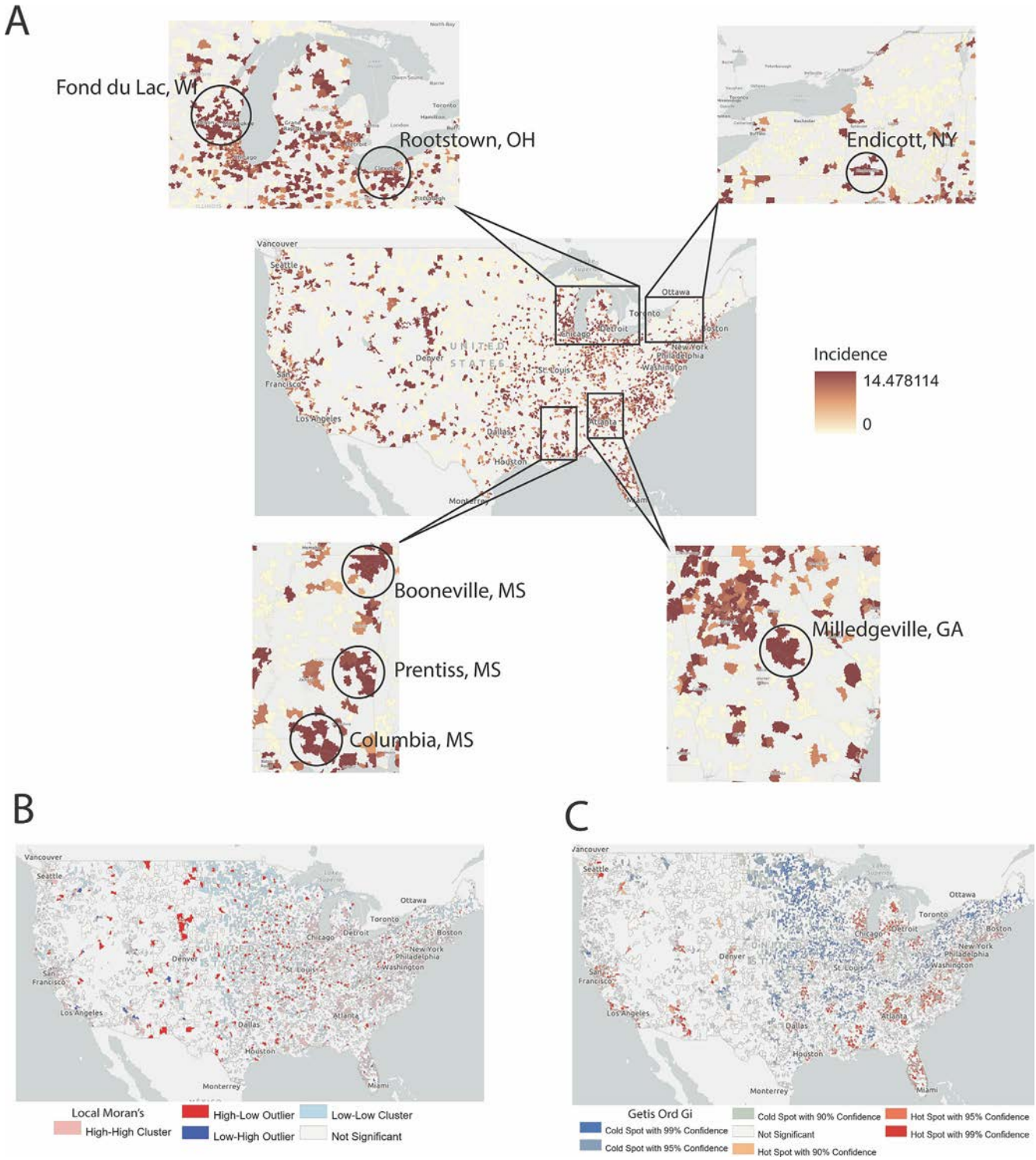

Figure 3. (A) Distribution of the incidence rate of Medicare beneficiaries carrying at least one diagnosis code for cohort 2 (see Table 1) by zip code between the years 2014 and 2018. Lower incidence rates are indicated in yellow and orange, whereas higher incidence rates are indicated in darker orange and red. Statistically significant clusters were identified in Wisconsin, Michigan, Ohio, New York, Georgia, Mississippi, Nebraska, and Kansas. (B) Results of the Moran's Index local spatial autocorrelation on cohort 2 geographic distribution. Zip codes with high incidences surrounded by other zip codes of high incidences are indicated in pink. Dark red indicates zip codes of high incidences surrounded by areas of low incidences. Light blue and dark blue indicate areas of low incidences. (C) Results of optimized hot spot analysis on cohort 2 geographic data. Hot spots of high incidence are shown in varying shades of red, whereas cold spots of low incidences are indicated in varying shades of blue.

Figure 4. (A) Distribution of the incidence rate of Medicare beneficiaries carrying a diagnosis of cohort 2 by zip code containing at least one Superfund site between the years 2014 and 2018. Zip codes with high incidences are colored in a red or orange, whereas zip codes with low incidences are colored in tan and light orange. (B–D) Clusters of cohort 2 were detected in New York, Arkansas, Mississippi, and Wisconsin.

these analyses culled Medicare data from a large sample size that spanned the geographic entirety of the United States, and beneficiaries included in these analyses were between the ages of 65 and 74. Patients with severe chronic illnesses that cause significant disability, including renal disease, can qualify for Medicare before the age of 65, which speaks to the severe disability patients with SSc experience. Therefore, we did not include individuals under the age of 65 in our analysis.

The geographic distribution of cohort 1 beneficiaries occurred in clustered, nonrandom patterns in the United States. We identified several zip codes within clusters with high cohort 1 incidence in northeast Mississippi, Louisiana, New Jersey, and Maryland. Although potential sources of contamination that may account for the increased incidence of cohort 1 diagnosis codes in northeast Mississippi and Louisiana were not apparent, clusters of cohort 1 diagnosis codes were identified in and around New

Jersey and Maryland zip codes that contained at least one Superfund site. Zip codes that encompassed the Central Chemical, Brick Township Landfill, Vineland Chemical Company, Nascolite Corp, and Upper Deerfield Township Sanitary Landfill Superfund sites demonstrated elevated risk ratios of 5.66, 3.72, 3.84, 3.44, and 4.90, respectively, compared to the national risk of cohort 1 diagnosis codes between the years 2014 and 2018.

New Jersey has the greatest number of Superfund sites in the United States,[32] many of which processed vinyl chloride, a compound that has been associated with the development of SSc.[24] Interestingly, 1,2 dichloroethane, which is used in the production of vinyl chloride, is found at 62 of the 114 National Priorities List (NPL) sites in New Jersey.[33] 1,2 dichloroethane is particularly dangerous when it is absorbed into ground water and then used for showering, as inhalation of shower-generated vapors may induce immune activation and result in damage to the nervous system, liver, kidneys, and lungs.[34] Although vinyl chloride has been associated with liver, dermal, and kidney fibrosis,[35–37] the role of 1,2 dichloroethane in fibrotic activation is unclear. The effect of both of these chemicals on skin fibrosis have not been established and warrants further investigation.

Similar to cohort 1, the geographic distribution of cohort 2 beneficiaries occurred in clustered, nonrandom patterns in the United States. Of note, southeast Wisconsin exhibited a high incidence of cohort 2 diagnosis codes in many zip codes in and around the Milwaukee metropolitan area, home to lumber processing plants, dairies, breweries, and agriculture and manufacturing industries.[38] Given this history of industrial activity, it is difficult to pinpoint any one pollutant with fibrotic or inflammatory-inducing properties. However, landfill Superfund sites in this area may contribute to disease development through exposure to synthetic chemicals and heavy metals. For example, Refuse Hideaway Landfill, located in southeast Wisconsin outside of Madison, has reported the presence of TCE,[39] which has been implicated in SSc presentation and progression.[40,41] In addition, heavy metals, including cadmium, cobalt, copper, iron, lead, and zinc, have been found in the 38-acre City Disposal Corp Landfill, a Superfund site in Dunn, Wisconsin.[42] Exposure to cadmium, lead, and zinc has been associated with SSc development.[43,44]

Similarly, Sauble Township in Lake County, Michigan, had a high burden of cohort 2 diagnosis codes. Lake County is home to the Wash King Laundry Superfund site, where perchloroethylene was used in dry cleaning. This site remains on the NPL, a government-curated list of Superfund sites, and in situ thermal remediation has been recommended to clean up the impacted area. Exposure to perchloroethylene and TCE has been associated with diffuse SSc onset.[45]

Clusters of cohort 2 diagnosis codes were identified around Prentiss, Marion, and Lauderdale counties in Mississippi. Notably, Marion County was home to the Newsome Brothers/Old Reichold Chemicals Inc Superfund site in Columbia, Mississippi. This site was initially used as a wood treating facility before its conversion to a chemical plant where pentachlorophenol, a potential human carcinogen,[46] was mixed with diesel oil. Although the EPA removed this site from the NPL in 2000 and it is now deemed safe, many individuals exposed to chemicals before remediation may still reside in the area, potentially accounting for the increased incidence reported here.

Moreover, a high incidence cluster of cohort 2 diagnosis codes were noted in and around Endicott, New York. Endicott is home to the Endicott Village Well Field Superfund site, which contains a Ranney well and accompanying zone of influence on area groundwater. The EPA detected vinyl chloride, a potential trigger for SSc onset,[23,24] and other volatile organic compounds in the water from the Ranney well. The source of well contamination was identified as the neighboring Endicott Landfill.

Wood processing Superfund sites including the Mid-South Wood Products in Arkansas, Newsome Brothers/Old Reichhold Chemicals in Mississippi, Picayune Wood Treating in Mississippi, and Georgia-Pacific Corp Timber Company were associated with high rates of cohort 2 diagnosis codes. Read counts for *Rhodotorula glutinis*, a red yeast that thrives in wet lumber, are significantly higher in skin from individuals with SSc compared to healthy controls.[47,48] Given the role of innate immune activation in SSc, it is possible that exposure to yeast may induce inflammation driven by microbial engagement of Toll-like receptors (TLRs). In support of this potential mechanism, TLR2, which binds fungal ligands and regulates fibrotic and inflammatory activation, is upregulated in SSc.[49,50] Therefore, it is possible that *R glutinis* found at Superfund sites associated with lumber processing plays a role in the induction and/or progression of fibrotic signaling in SSc[8]).

In addition to regional clustering, high cohort 2 incidence was detected in isolated zip codes irrespective of localization near Superfund sites. For example, the incidences of cohort 2 diagnosis codes in WaKeeney, Kansas, and Hartington, Nebraska, were 6.55% and 6.18%, respectively. One potential explanation for this observation is that genetically predisposed family members living in close proximity to one another may have been exposed to the same or similar environmental triggers, leading to one or more of the diagnoses included in cohort 2. In this regard, epidemiologic studies demonstrate that the risk of developing SSc and other autoimmune conditions is increased in family members of patients with SSc. Although familial relations account for two-thirds of the phenotypic variance of disease, one-third remains unexplained.[51] Therefore, environmental exposures may account, at least in part, for the remaining risk, as has been noted in cancer.[52] Future studies may focus on the identification of environmental and occupational exposures among geographic clusters of familial SSc to evaluate associations between these factors and potential risk of disease development.

Although high incidence clusters of cohorts 1 and 2 were identified in and around Booneville, Mississippi, these were not associated with Superfund sites, and obvious sources of environmental contamination were not apparent. In light of this, future

work may be directed at enrolling participants in a cohort study to gather information on occupational exposures and lifestyle factors that may contribute to these increased cohort 1 and 2 incidences.

The cohort 2 analysis in this study included individuals with "personal history of other diseases of the musculoskeletal system and connective tissues" to capture SSc beneficiaries with a history of SSc. Notably, individuals with a diagnosis of "personal history of other diseases of the musculoskeletal system and connective tissues" constitute the largest group in the analyzed cohort. Early SSc symptoms and features overlap with other autoimmune conditions, further confounding accurate diagnosis. As a result, physicians may provide patients with broader diagnosis codes until their diseases differentiate.

Strengths of this study include the large sample size (n = 179,188), broad geographic distribution, comprehensive lists of diagnosis codes, environmental contaminant information, and zip code–level beneficiary residence. This is significantly different from previous analyses, which evaluated smaller sample sizes or are limited by sex or geospatial region.[6,10,53,54] One limitation of this study is that levels of synthetic chemicals were not directly measured at geographic sites of increased disease incidence. Nonetheless, these results may generate hypotheses about potential disease triggers that may be assessed quantitatively. It should be noted that this study was limited to the disease definitions described in the ICD-9 and 10 codes available in the Medicare database. Such constraints may result in a failure to adequately describe a beneficiary's disease using such codes and could inadvertently exclude SSc beneficiaries from our cohorts. In addition, because this study was not restricted to a single geographic region or by sex, it may fail to adequately capture differences in disease incidence associated with sex or race. To limit potential confounders associated with age, the study cohort was restricted to individuals over the age of 65. Although there is a small fraction of beneficiaries under the age of 65 that met diagnostic code criteria, this younger population is likely to differ clinically and by sex.[55,56] In this regard, men are often diagnosed with SSc at a young age, with more severe disease and substantial organ involvement. A separate study of this younger population may be warranted to elucidate the environmental triggers responsible for increased disease severity.[30,57] One further limitation of this study is that the Medicare data to which we have access do not provide information about the diagnosing physician, and therefore we cannot conduct additional analyses to discern if expert centers diagnosed beneficiaries differently than clinics at other locations.

To our knowledge, this study is the first to investigate the distribution of SSc and SSc-related diagnosis codes across the United States in a population of Medicare beneficiaries over the age of 65 between the years 2014 and 2018. We examined the geographic distribution of 179,188 beneficiaries with at least one of the diagnosis codes, reported in Table 1, and discovered that cohorts 1 and 2 exhibited a nonrandom, clustered distribution pattern across the United States. Many of these clusters showed a significant association with Superfund sites, areas the US government has deemed hazardous to human and environmental health due to contamination by pollutants. These results provide the foundation for future investigation of the relationship between environmental pollutants and chronic inflammatory and fibrotic illnesses.

ACKNOWLEDGMENTS

The authors would like to thank Monica Adams-Foster and Sukdith Punjasthitkul for their data management and administrative contributions to this study. The authors would like to thank Drs Jonathan Chipman and Xun Shi for their indispensable guidance concerning geospatial statistics.

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 Whitfield 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. Allanore Y, Simms R, Distler O, et al. Systemic sclerosis. Nat Rev Dis Primers 2015;1(1):15002.
2. 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.
3. Bairkdar M, Rossides M, Westerlind H, et al. Incidence and prevalence of systemic sclerosis globally: a comprehensive systematic review and meta-analysis. Rheumatology (Oxford) 2021;60(7):3121–3133.
4. Tian J, Kang S, Zhang D, et al. Global, regional, and national incidence and prevalence of systemic sclerosis. Clin Immunol 2023;248:109267.
5. Bergamasco A, Hartmann N, Wallace L, et al. Epidemiology of systemic sclerosis and systemic sclerosis-associated interstitial lung disease. Clin Epidemiol 2019;11:257–273.
6. Knarborg M, Hyldgaard C, Bendstrup E, et al. Incidence, prevalence and regional distribution of systemic sclerosis and related interstitial lung disease: a nationwide retrospective cohort study. Chron Respir Dis 2022;19:14799731221125559.
7. Coi A, Barsotti S, Santoro M, et al; Rare Diseases Working Group. Epidemiology of systemic sclerosis: a multi-database population-based study in Tuscany (Italy). Orphanet J Rare Dis 2021;16(1):90.
8. Royle JG, Lanyon PC, Grainge MJ, et al. The incidence, prevalence, and survival of systemic sclerosis in the UK Clinical Practice Research Datalink. Clin Rheumatol 2018;37(8):2103–2111.
9. Ferri C, Valentini G, Cozzi F, et al; Systemic Sclerosis Study Group of the Italian Society of Rheumatology (SIR-GSSSc). Systemic sclerosis: demographic, clinical, and serologic features and survival in 1,012 Italian patients. Medicine (Baltimore) 2002;81(2):139–153.

10. Kassamali B, Kassamali AA, Muntyanu A, et al. Geographic distribution and environmental triggers of systemic sclerosis cases from 2 large academic tertiary centers in Massachusetts. J Am Acad Dermatol 2022;86(4):925–927.
11. Mayes MD, Lacey JV Jr, Beebe-Dimmer J, et al. Prevalence, incidence, survival, and disease characteristics of systemic sclerosis in a large US population. Arthritis Rheum 2003;48(8):2246–2255.
12. Rodriguez-Pla A, Simms RW. Geographic disparity in systemic sclerosis mortality in the United States: 1999–2017. J Scleroderma Relat Disord 2021;6(2):139–145.
13. Fan Y, Bender S, Shi W, et al. Incidence and prevalence of systemic sclerosis and systemic sclerosis with interstitial lung disease in the United States. J Manag Care Spec Pharm 2020;26(12):1539–1547.
14. Lescoat A, Ballerie A, Lecureur V, et al. The neglected association of crystalline silica exposure and systemic sclerosis. Rheumatology (Oxford) 2020;59(12):3587–3588.
15. Boudigaard SH, Schlünssen V, Vestergaard JM, et al. Occupational exposure to respirable crystalline silica and risk of autoimmune rheumatic diseases: a nationwide cohort study. Int J Epidemiol 2021;50(4):1213–1226.
16. Ferri C, Artoni E, Sighinolfi GL, et al. High serum levels of silica nanoparticles in systemic sclerosis patients with occupational exposure: possible pathogenetic role in disease phenotypes. Semin Arthritis Rheum 2018;48(3):475–481.
17. Jain S, Joshi V, Rathore YS, et al. Erasmus syndrome: silicosis and systemic sclerosis. Indian J Occup Environ Med 2017;21(2):94–96.
18. Kim JY, Do SY, Moon YH, et al. Systemic sclerosis due to crystalline silica exposure among jewelry workers in Korea: two case reports. Ann Occup Environ Med 2017;29:18.
19. Yáñez Díaz S, Morán M, Unamuno P, et al. Silica and trichloroethylene-induced progressive systemic sclerosis. Dermatology 1992;184(2):98–102.
20. Lockey JE, Kelly CR, Cannon GW, et al. Progressive systemic sclerosis associated with exposure to trichloroethylene. J Occup Med 1987;29(6):493–496.
21. Forkert PG, Forkert L. Trichloroethylene induces pulmonary fibrosis in mice. Can J Physiol Pharmacol 1994;72(3):205–210.
22. Diot E, Lesire V, Guilmot J, et al. Systemic sclerosis and occupational risk factors: a case-control study. Occup Environ Med 2002;59(8):545–549.
23. Ostlere LS, Harris D, Buckley C, et al. Atypical systemic sclerosis following exposure to vinyl chloride monomer. A case report and review of the cutaneous aspects of vinyl chloride disease. Clin Exp Dermatol 1992;17(3):208–210.
24. Black CM, Welsh KI, Walker AE, et al. Genetic susceptibility to scleroderma-like syndrome induced by vinyl chloride. Lancet Lond Engl 1983;1(8314):53–55.
25. Amin R, Nelson A, McDougall S. A spatial study of the location of Superfund sites and associated cancer risk. Stat Public Policy (Phila) 2018;5(1):1–9.
26. Budnick LD, Logue JN, Fox JM, et al. Cancer and birth defects near the Drake Superfund site, Pennsylvania. Arch Environ Health Int J 1984;39(6):409–413.
27. Clapp RW, Hoffman K. Cancer mortality in IBM Endicott plant workers, 1969–2001: an update on a NY production plant. Environ Health 2008;7(1):13.
28. Gelber AC, Manno RL, Shah AA, et al. Race and association with disease manifestations and mortality in scleroderma: a 20-year experience at the Johns Hopkins Scleroderma Center and review of the literature. Medicine (Baltimore) 2013;92(4):191–205.
29. Reveille JD. Ethnicity and race and systemic sclerosis: how it affects susceptibility, severity, antibody genetics, and clinical manifestations. Curr Rheumatol Rep 2003;5(2):160–167.
30. Steen VD, Medsger TA. Changes in causes of death in systemic sclerosis, 1972–2002. Ann Rheum Dis 2007;66(7):940–944.
31. Steen VD, Oddis CV, Conte CG, et al. TA. Incidence of systemic sclerosis in Allegheny County, Pennsylvania. A twenty-year study of hospital-diagnosed cases, 1963-1982. Arthritis Rheum 1997;40(3):441–445.
32. US Environmental Protection Agency. National Priorities List (NPL) Sites - by State. August 14, 2015. Accessed April 11, 2022. https://www.epa.gov/superfund/national-priorities-list-npl-sites-state
33. Agency for Toxic Substances and Disease Registry. 1,2-Dichloroethane. February 10, 2021. Accessed April 11, 2022. https://wwwn.cdc.gov/TSP/substances/ToxSubstance.aspx?toxid=110
34. National Research Council (US) Committee on Toxicology. An assessment of the health risks of seven pesticides used for termite control. Washington (DC): National Academies Press (US); 1982. PENTACHLOROPHENOL. Accessed April 22, 2022. https://www.ncbi.nlm.nih.gov/books/NBK217616/
35. Goodman RS, Mittal L, Parker ER. Public health risks, dermatological manifestations, and environmental justice associated with vinyl chloride exposure: narrative review. JMIR Dermatol 2023;6(1):e48998.
36. Fedeli U, Girardi P, Mastrangelo G. Occupational exposure to vinyl chloride and liver diseases. World J Gastroenterol 2019;25(33):4885–4891.
37. Hsu YH, Chuang HC, Lee YH, et al. Induction of fibrosis and autophagy in kidney cells by vinyl chloride. Cells 2019;8(6):601.
38. Institute for Reforming Government. Wisconsin Leads The Way In Manufacturing. October 3, 2019. Accessed May 16, 2022. https://reforminggovernment.org/wisconsin-leads-the-way-in-manufacturing/
39. US Environmental Protection Agency. Refuse Hideaway Landfill. Middleton, WI. Accessed March 27, 2024. https://cumulis.epa.gov/supercpad/SiteProfiles/index.cfm?fuseaction=second.contams&id=0505114
40. Zhao JH, Duan Y, Wang YJ, et al. The influence of different solvents on systemic sclerosis: an updated meta-analysis of 14 case-control studies. J Clin Rheumatol 2016;22(5):253–259.
41. Garabrant DH, Lacey JV Jr, Laing TJ, et al. Scleroderma and solvent exposure among women. Am J Epidemiol 2003;157(6):493–500.
42. US Environmental Protection Agency. City Disposal Corp. Landfill. Dunn, WI. Accessed March 27, 2024. https://cumulis.epa.gov/supercpad/SiteProfiles/index.cfm?fuseaction=second.contams&id=0505117
43. Forte G, Fadda C, Bocca B, et al. Association between exposure to heavy metals and systemic sclerosis: the levels of Al, Cd, Hg, and Pb in blood and urine of patients. Biol Trace Elem Res 2019;190(1):1–10.
44. Marie I, Gehanno JF, Bubenheim M, et al. Systemic sclerosis and exposure to heavy metals: a case control study of 100 patients and 300 controls. Autoimmun Rev 2017;16(3):223–230.
45. Pralong P, Cavailhes A, Balme B, et al. [Diffuse systemic sclerosis after occupational exposure to trichloroethylene and perchloroethylene]. Ann Dermatol Venereol 2009;136(10):713–717.
46. US Environmental Protection Agency. Search Superfund Site Information. Accessed January 12, 2022. https://cumulis.epa.gov/supercpad/CurSites/srchsites.cfm
47. Arron ST, Dimon MT, Li Z, et al. High Rhodotorula sequences in skin transcriptome of patients with diffuse systemic sclerosis. J Invest Dermatol 2014;134(8):2138–2145.
48. Johnson ME, Franks JM, Cai G, et al. Microbiome dysbiosis is associated with disease duration and increased inflammatory gene expression in systemic sclerosis skin. Arthritis Res Ther 2019;21(1):49.

49. Henderson J, Bhattacharyya S, Varga J, et al. Targeting TLRs and the inflammasome in systemic sclerosis. Pharmacol Ther 2018;192: 163–169.
50. Fang F, Marangoni RG, Zhou X, et al. Toll-like receptor 9 signaling is augmented in systemic sclerosis and elicits transforming growth factor β-dependent fibroblast activation. Arthritis Rheumatol 2016;68(8): 1989–2002.
51. Kuo CF, Luo SF, Yu KH, et al. Familial risk of systemic sclerosis and co-aggregation of autoimmune diseases in affected families. Arthritis Res Ther 2016;18(1):231.
52. Simchoni S, Friedman E, Kaufman B, et al. Familial clustering of site-specific cancer risks associated with BRCA1 and BRCA2 mutations in the Ashkenazi Jewish population. Proc Natl Acad Sci USA 2006; 103(10):3770–3774.
53. Medsger TA, Masi AT. The epidemiology of systemic sclerosis (scleroderma) among male U.S. veterans. J Chronic Dis 1978;31(2):73–85.
54. Frech TM, Murtaugh MA, Amuan M, et al. The frequency of Raynaud's phenomenon, very early diagnosis of systemic sclerosis, and systemic sclerosis in a large Veteran Health Administration database. BMC Rheumatol 2021;5(1):42.
55. Peoples C, Medsger TA Jr, Lucas M, et al. Gender differences in systemic sclerosis: relationship to clinical features, serologic status and outcomes. J Scleroderma Relat Disord 2016;1(2):177–240.
56. Hughes M, Pauling JD, Armstrong-James L, et al. Gender-related differences in systemic sclerosis. Autoimmun Rev 2020;19(4):102494.
57. Gaultier JB, Hot A, Cathébras P, et al. [Systemic sclerosis in men]. Rev Med Interne 2008;29(3):181–186.

Arthritis Care & Research
Vol. 77, No. 7, July 2025, pp 867–872
DOI 10.1002/acr.25513

Association Between Metabolic Syndrome and Radiographic Changes in Psoriatic Arthritis: A Cohort Study

Fadi Kharouf,[1] Shangyi Gao,[2] S. Ercan Tunc,[1] Justine Y. Ye,[2] Daniel Pereira,[2] Dafna D. Gladman,[1] and Vinod Chandran[1]

Objective. Metabolic syndrome (MetS) is a known comorbidity of psoriatic arthritis (PsA) and is associated with PsA disease activity. We aimed to explore the association between MetS and radiographic features (peripheral and axial) in PsA.

Methods. We included patients with PsA followed at our prospective observational cohort for the period between 1978 and 2024. We identified patients with MetS on longitudinal follow-up and used generalized estimating equations (GEE) analysis to define the radiographic features independently associated with MetS, adjusting for age, sex, PsA disease duration, calendar decade, and use of targeted disease-modifying antirheumatic drugs.

Results. The study population consisted of 1,422 patients, out of which 400 (28.1%) had MetS at baseline (clinic entry) and 836 (58.79%) had a record of MetS (per the harmonized definition) over a median follow-up duration of 10.59 (interquartile range 4.52–18.28) years. The mean (SD) age of our cohort at baseline was 44.43 (12.98) years, with 789 patients (55.5%) identifying as men. Mean (SD) body mass index was 28.79 (6.36) kg/m^2. In the GEE analysis, MetS was not significantly associated with axial disease or radiographic damage to peripheral joints, assessed as the presence of syndesmophytes or sacroiliitis and the radiographic damaged joint count, respectively. On the other hand, MetS was significantly associated with calcaneal spurs, diffuse idiopathic skeletal hyperostosis, and degenerative disc disease.

Conclusion. MetS is associated with degenerative and metabolic changes in the spine and entheses but not with radiographic damage in PsA.

INTRODUCTION

Psoriatic arthritis (PsA) is a complex inflammatory disease with heterogeneous clinical features, which complicates psoriasis in 24% of patients.[1] Musculoskeletal manifestations of PsA include peripheral arthritis, enthesitis, dactylitis, spondylitis, and sacroiliitis.[2] Obesity, type 2 diabetes mellitus (DM), hypertension, metabolic syndrome (MetS), fatty liver, and an increased risk of cardiovascular events are all associated comorbidities.[3]

Psoriatic disease activity, radiographic damage, and mortality rates are all interrelated. It is well accepted that inflamed joints result in radiographic damage.[4] We have previously shown that patients with PsA are at an increased risk of death compared with the general population, with evidence of previous disease activity and radiographic damage being prognostic indicators.[5,6] Given the importance of structural changes and their association with functional ability, quality of life, and survival, preventing radiographic damage remains a crucial goal of PsA therapy.[7]

Although it has several definitions, MetS generally refers to the co-occurrence of known cardiovascular risk factors, including insulin resistance, obesity, atherogenic dyslipidemia, and hypertension.[8] These are closely related and share common mediators, pathways, and pathophysiologic mechanisms. MetS and its components are significantly overrepresented in patients with PsA compared with the general population[9]; 23.5% to 62.9% of patients with PsA have MetS,[10] which contributes to adverse cardiovascular outcomes,[11] besides being associated with a chronic low-grade inflammatory state, disease activity, and less response

The Gladman Krembil Psoriatic Arthritis Research Program is funded by the Schroeder Arthritis Institute and a grant from the Krembil Foundation.

[1]Fadi Kharouf, MD, S. Ercan Tunc, MD, Dafna D. Gladman, MD, FRCPC, Vinod Chandran, MBBS, MD, DM, PhD, FRCPC: Schroeder Arthritis Institute, Krembil Research Institute, University Health Network and University of Toronto, Toronto, Ontario, Canada; [2]Shangyi Gao, MSc, Justine Y. Ye, MSc, Daniel Pereira, BSc: Schroeder Arthritis Institute, Krembil Research Institute, University Health Network, Toronto, Ontario, Canada.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25513).

Author disclosures and graphical abstract are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25513.

Address correspondence to Vinod Chandran, MBBS, MD, DM, PhD, FRCPC, at vinod.chandran@uhn.ca.

Submitted for publication August 14, 2024; accepted in revised form January 29, 2025.

SIGNIFICANCE & INNOVATIONS

- Despite being associated with disease activity, metabolic syndrome is not associated with axial or peripheral radiographic damage in psoriatic arthritis.
- Metabolic syndrome is associated with degenerative changes and diffuse idiopathic skeletal hyperostosis.

to treatment.[12–14] Because inflammation is a mediator of joint damage in PsA,[4] one may hypothesize that patients with MetS accrue greater radiographic progression. However, only a few studies have investigated this relationship, with inconclusive results.[15,16] This analysis aimed to explore the association between MetS and radiographic changes (peripheral and axial) in an observational PsA cohort.

PATIENTS AND METHODS

Setting. The University of Toronto PsA clinic has recruited and prospectively followed patients with PsA since 1978. Patients are enrolled if they have psoriasis and inflammatory arthritis. Over 99% fulfill the classification criteria for PsA.[17] Patients are evaluated at the time of recruitment into the clinic and every 6 to 12 months according to a standard protocol that includes a detailed history, physical examination, and laboratory assessment. Radiographs of peripheral joints and spine are obtained at baseline (clinic entry) and every 2 years and are scored according to the modified Steinbrocker method for peripheral joints[18] and the New York criteria for sacroiliac joints.[19] The presence of syndesmophytes, atlantoaxial subluxation, degenerative disc disease (DDD) (cervical and/or thoracolumbar), calcaneal spurs (Achilles and/or plantar), and diffuse idiopathic skeletal hyperostosis (DISH) is also recorded. Radiographs are scored by at least two rheumatologists by consensus. All information is collected via a web portal and stored in a database. We have previously shown that the methods of clinical and radiographic evaluation in our cohort are reliable, with no systematic bias with regard to disease severity and follow-up.[20]

Data collection. From our program database, we retrieved data (between 1978 and 2024) at the patients' first evaluation in the clinic and at 6 to 12 month intervals, including age, sex, ethnicity, age at the diagnosis of psoriasis and PsA, duration of PsA, waist circumference, body mass index (BMI), comorbidities (ever smoking, fibromyalgia, DM, hypertension, dyslipidemia, and cardiovascular disease), and the presence of mechanical and/or inflammatory back pain. Patients were classified as having MetS at their last clinic visit if they satisfied the harmonized definition by Alberti et al.[21] According to this definition, the condition is diagnosed when any three of the five following risk factors are present:

1. Elevated waist circumference with sex, population, and country-specific criteria.
2. Elevated triglycerides, defined as ≥150 mg/dL or its treatment.
3. Decreased high-density lipoprotein cholesterol, defined as <40 mg/dL in men and <50 mg/dL in women or its treatment.
4. Elevated blood pressure, defined as systolic blood pressure ≥130 mm Hg or diastolic blood pressure ≥85 mm Hg or its treatment.
5. Elevated fasting glucose, defined as blood glucose >100 mg/dL or its treatment.

We assessed musculoskeletal disease activity using the swollen joint count, tender joint count, actively inflamed (swollen or tender) joint count, presence of enthesitis (using the Spondyloarthritis Research Consortium of Canada enthesitis index[22]), dactylitis, and the Disease Activity Index for PsA score.[23] We summarized skin disease severity using the Psoriasis Area and Severity Index. We also retrieved other important disease-related measures including the presence of nail disease. We included the following radiographic features: sacroiliitis (bilateral grade 2 or unilateral ≥3), syndesmophytes, radiographic damaged joint count (number of joints with erosions scored according to the modified Steinbrocker score[18]), cervical and/or thoracolumbar DDD, atlantoaxial subluxation, calcaneal spurs (Achilles and/or plantar), and DISH. We defined the latter as the presence of flowing calcifications and ossifications along the anterolateral aspect of at least four contiguous vertebral bodies with or without associated localized pointed excrescences at the intervening vertebral body-intervertebral disc junctions.[24]

Laboratory data included erythrocyte sedimentation rate, C-reactive protein (CRP), and HLA-B*27 status. We included the following patient-reported outcome (PRO) measures: the Health Assessment Questionnaire, Medical Outcomes Study 36-Item Short Form Health Survey (SF-36) physical component summary score, SF-36 mental component summary score, a pain severity scale, and a patient global assessment of disease scale. We also retrieved information on current treatments with conventional synthetic disease-modifying antirheumatic drugs (DMARDs) (methotrexate, sulfasalazine, and leflunomide) and targeted DMARDs, including biologics (tumor necrosis factor inhibitors [TNFi] and interleukin [IL]-12/23i, IL-17i, and IL-23i) and synthetic DMARDs (apremilast, tofacitinib, and upadacitinib).

Statistical analysis. We described the baseline demographic and disease-related characteristics of patients with and without MetS. Continuous variables are expressed as the mean (SD) or median (interquartile range [IQR]), whereas categorical variables are presented as frequency (%).

A directed acyclic graph (DAG) was constructed to identify potential confounders and factors that may influence both the exposure (radiographic changes) and outcome (MetS) (see Supplementary Figure 1). Based on the DAG, the sufficient adjustment set for estimating the total effect of radiographic features on MetS included age, sex, PsA disease duration, calendar decade

Table 1. Baseline demographic and disease-related features of patients with PsA included in the study*

Variable	All patients (n = 1,422)[a]	Non-MetS (n = 586)	MetS (n = 836)
Demographic features			
Age, mean (SD), years	44.43 (12.98)	41.44 (13.10)	46.52 (12.49)
Age at diagnosis of psoriasis, mean (SD), years	28.71 (14.68)	26.67 (14.02)	30.14 (14.97)
Age at diagnosis of PsA, mean (SD), years	38.56 (13.70)	35.86 (13.68)	40.45 (13.40)
PsA disease duration, mean (SD), years	5.83 (7.82)	5.55 (7.88)	6.02 (7.77)
Sex, men, n (%)	789 (55.50)	303 (51.70)	486 (58.10)
White, n (%)	1,188 (83.90)	498 (85.40)	690 (82.80)
Waist circumference, mean (SD), cm	96.05 (15.63)	87.47 (15.06)	101.69 (13.26)
BMI, mean (SD), kg/m^2	28.79 (6.36)	25.78 (5.16)	30.37 (6.36)
Comorbidities			
Fibromyalgia, n (%)	121 (11.00)	34 (7.60)	87 (13.20)
Ever a smoker, n (%)	598 (43.80)	215 (39.80)	383 (46.40)
Diabetes mellitus, n (%)	96 (6.80)	4 (0.7)	92 (11.0)
Hypertension, n (%)	206 (14.50)	18 (3.1)	188 (22.5)
Dyslipidemia, n (%)	122 (8.60)	13 (2.2)	109 (13.1)
Cardiovascular disease, n (%)	241 (16.90)	33 (5.6)	208 (24.9)
Clinical features			
Inflammatory back pain, n (%)	285 (20.10)	103 (17.6)	182 (21.8)
Mechanical back pain, n (%)	160 (11.30)	59 (10.1)	101 (12.1)
SJC, median (IQR)	2.00 (0.00–5.00)	1.00 (0.00–4.00)	2.00 (0.00–5.00)
TJC, median (IQR)	4.00 (1.00–9.00)	3.00 (1.00–8.00)	4.00 (1.00–10.25)
AJC, median (IQR)	5.00 (2.00–12.00)	4.00 (1.00–11.00)	6.00 (2.00–13.00)
Dactylitis, n (%)	360 (25.50)	151 (25.90)	209 (25.20)
Enthesitis, n (%)	298 (21.20)	114 (19.60)	184 (22.30)
Enthesitis SPARCC score, median (min, max)	0.00 (0.00, 16.00)	0.00 (0.00, 14.00)	0.00 (0.00, 16.00)
PASI, median (IQR)	1.50 (0.00–4.50)	1.00 (0.00–3.90)	1.80 (0.00–4.90)
Nail disease, n (%)	774 (54.40)	310 (52.90)	464 (55.50)
Radiographic features			
Bilateral grade 2 or any ≥3 sacroiliitis, n (%)	286 (22.00)	115 (21.90)	171 (22.00)
Syndesmophytes, n (%)	157 (12.00)	56 (10.70)	101 (13.00)
Radiographic damaged joint count, median (IQR)	0.00 (0.00–4.00)	0.00 (0.00–3.00)	1.00 (0.00–4.00)
Atlantoaxial subluxation, n (%)	31 (2.40)	13 (2.50)	18 (2.30)
Plantar and/or Achilles calcaneal spurs, n (%)	588 (45.20)	172 (33.00)	416 (53.40)
DISH, n (%)	66 (5.30)	7 (1.50)	59 (7.60)
Cervical disc disease, n (%)	401 (32.30)	109 (23.30)	292 (37.80)
Thoracolumbar disc disease, n (%)	432 (34.80)	117 (25.00)	315 (40.80)
Laboratory features			
HLA-B*27 positive, n (%)	181 (16.50)	80 (19.90)	101 (14.50)
ESR, mean (SD), mm/h	21.93 (20.12)	21.56 (21.43)	22.18 (19.23)
CRP, mean (SD), mg/L	12.98 (19.26)	12.37 (21.21)	13.30 (18.18)
Patient-reported outcomes			
HAQ, median (IQR)	0.50 (0.12–1.12)	0.50 (0.00–1.00)	0.62 (0.12–1.25)
Global assessment of disease activity, median (IQR)	4.00 (2.00–6.00)	4.00 (2.00–6.00)	5.00 (2.00–7.00)
Pain severity, median (IQR)	4.00 (2.00–7.00)	3.00 (2.00–6.00)	5.00 (2.00–7.00)
SF-36 MCS, median (IQR)	47.56 (36.48–55.47)	48.38 (39.41–55.67)	46.46 (35.72–55.36)
SF-36 PCS, median (IQR)	35.92 (27.50–45.86)	37.43 (29.20–47.80)	34.89 (26.53–45.12)
Composite disease activity measure			
DAPSA, median (min, max)	18.30 (10.15, 32.00)	16.00 (9.00, 28.28)	19.20 (11.00, 34.00)
Pharmacotherapy			
Conventional synthetic DMARDs, n (%)	518 (36.40)	174 (29.70)	344 (41.10)
Targeted (biologic or synthetic) DMARDs, n (%)	156 (11.00)	60 (10.20)	96 (11.50)

* AJC, actively inflamed (tender/swollen) joint count; BMI, body mass index; CRP, C-reactive protein; DAPSA, Disease Activity Index for Psoriatic Arthritis; DMARD, disease-modifying antirheumatic drug; DISH, diffuse idiopathic skeletal hyperostosis; ESR, erythrocyte sedimentation rate; HAQ, Health Assessment Questionnaire; IQR, interquartile range; max, maximum, MetS, metabolic syndrome; min, minimum; Non-MetS, does not meet criteria for metabolic syndrome; PASI, Psoriasis Area Severity Index; PsA, psoriatic arthritis; SF-36 MCS, Medical Outcomes Study 36-Item Short Form Health Survey mental component summary score; SF-36 PCS, Medical Outcomes Study 36-Item Short Form Health Survey physical component summary score; SJC, swollen joint count; SPARCC, Spondyloarthritis Research Consortium of Canada; TJC, tender joint count.

[a] Four hundred patients (28.1%) had MetS at baseline.

(1978–1988, 1989–1999, 2000–2010, and 2011–2024), and targeted DMARD use.

Generalized estimating equations (GEE) were employed to account for the longitudinal nature of the data. Univariate and multivariate GEE models were developed to identify the radiographic features associated with MetS in patients with PsA, adjusting for age, sex, PsA disease duration, calendar decade, and targeted DMARD use. Odds ratios (ORs) with 95% confidence intervals (CIs) are reported. The adequacy, linearity, and independence of the model were thoroughly assessed and found to be valid. All analyses were performed using R version 9.4 (SAS Institute).

Ethics. We obtained ethics review and approval from the University Health Network Research Ethics Board (08-0630). We collected informed consent from all patients at enrollment. The data underlying this article will be shared upon reasonable request to the corresponding author.

RESULTS

Patient characteristics. The study population consisted of 1,422 patients, out of which 400 (28.1%) had MetS at baseline and 836 (58.79%) had a record of MetS over a median (IQR) follow-up duration of 10.59 (4.52–18.28) years. In patients who developed MetS after clinic enrollment, the median time to the first record of MetS was 6.85 (2.78–14.43) years. The mean (SD) age of our cohort at baseline was 44.43 (12.98) years, with 789 patients (55.5%) identifying as men. Mean (SD) BMI was 28.79 (6.36) kg/m^2. The median (IQR) radiographic damaged joint count was 0.00 (0.00–4.00), and 286 (22.0%) patients had sacroiliitis per the New York criteria.[19] One hundred fifty-six (11.0%) received targeted (biologic or synthetic) DMARDs at the time of clinic enrollment. Details of the demographic and disease characteristics at study entry are provided in Table 1.

Factors associated with MetS. We used GEE analysis to examine the radiographic features associated with MetS, adjusted for age, sex, PsA disease duration, calendar decade, and use of targeted DMARDs. The following variables were significantly associated with MetS in the univariate model (Table 2): plantar and/or Achilles calcaneal spurs (OR 2.58, 95% CI 1.90–3.51), DISH (OR 2.23, 95% CI 1.45–3.42), and DDD (OR 1.65, 95% CI 1.22–2.23). In the multivariate model, the following variables were significantly associated with MetS (Table 2): plantar and/or Achilles calcaneal spurs (OR 2.32, 95% CI 1.69–3.18), DISH (OR 2.00, 95% CI 1.28–3.13), and DDD (OR 1.51, 95% CI 1.11–2.04).There was no significant association with peripheral damage, assessed as the radiographic damaged joint count (OR 1.00, 95% CI 0.99–1.02), or axial damage, assessed as the presence of sacroiliitis (OR 0.79, 95% CI 0.58–1.08) or syndesmophytes (OR 1.16, 95% CI 0.81–1.65).

DISCUSSION

In our large longitudinal study, we have found that around 59% of the patients met the criteria for MetS over a median (IQR) follow-up duration of 10.59 (4.52–18.28) years. MetS was not associated with radiographic damage in PsA but correlated with degenerative and metabolic changes in the spine and entheses.

Only a few other analyses have investigated this relationship to date.[15,16] Xue et al[16] analyzed blood samples of healthy volunteers and patients with psoriasis and PsA to investigate the relation among adipokines and osteoclast precursors, radiographic damage scores, and disease activity indices. They concluded that there is an abnormal expression of soluble mediators of osteoclastogenesis and adipokines in PsA. Only TNF, however, correlated with radiographic damage scores. Haroon et al[15] studied a PsA cohort of 273 patients, 44% of whom had MetS. They demonstrated a significant association between insulin resistance and more severe PsA, which they defined as the presence of one or more of the PsA-related radiographic damage features

Table 2. Results of the generalized estimating equations analysis for radiographic features associated with metabolic syndrome, adjusted for age, sex, psoriatic arthritis duration, calendar decade, and use of targeted DMARDs*

Variable	Univariate model			Multivariate model		
	OR	95% CI	*P* value	OR	95% CI	*P* value
Syndesmophytes	1.27	0.91–1.77	0.17	1.16	0.81–1.65	0.43
Sacroiliitis	0.86	0.64–1.15	0.30	0.79	0.58–1.08	0.14
Radiographic damaged joint count[a]	1.01	0.99–1.03	0.22	1.00	0.99–1.02	0.65
Plantar and/or Achilles calcaneal spurs	2.58	1.90–3.51	<0.01	2.32	1.69–3.18	**<0.01**
DISH	2.23	1.45–3.42	<0.01	2.00	1.28–3.13	**<0.01**
Degenerative disc disease	1.65	1.22–2.23	<0.01	1.51	1.11–2.04	**0.01**
Atlantoaxial subluxation	1.21	0.61–2.38	0.58	1.12	0.55–2.25	0.76
HLA-B*27 positive	0.69	0.47–1.02	0.06	0.72	0.48–1.09	0.12
Conventional synthetic DMARDs	1.21	0.95–1.55	0.13	1.18	0.91–1.52	0.21

* Bold values indicate statistical significance. CI, confidence interval; DISH, diffuse idiopathic skeletal hyperostosis; DMARD, disease-modifying antirheumatic drug; OR, odds ratio.
[a] One-unit increase.

(peripheral joint erosions, osteolysis, or sacroiliitis) plus the use of TNFi. When only radiographic damage was used as the outcome, statistical significance was not maintained.[15]

Patients with obesity, an integral component but not synonymous with MetS, generally have higher PsA activity scores, worse PROs, lower probability of sustained remission, more frequent disease relapses, and less optimal responses to therapy (most notably TNFi).[9,25–28] This is explained, at least partly, by the low-grade systemic inflammatory state in individuals with obesity, commonly reflected by higher CRP concentrations.[29] In fact, obesity not only influences disease activity but may also be associated with different genetic profiles and disease patterns. We have previously shown a link between obesity and late-onset psoriasis and PsA,[30] a clinical subtype that is characterized by more active disease and greater radiographic damage.[31] On the other hand, the frequency of HLA-B*27, a marker of axial involvement (which in turn correlates with more severe peripheral disease[32]), tends to be lower in patients with higher BMI.[30] With all these interactions, exploring the association between MetS and radiographic damage in PsA is a reasonable research question.

Multivariate GEE analysis identified plantar and/or Achilles calcaneal spurs, DISH, and DDD (cervical and thoracolumbar) as radiographic features independently associated with MetS in patients with PsA but not radiographic damage to peripheral joints, sacroiliitis, or syndesmophytes. This is not unexpected because these radiographic findings share with MetS the close relation to aging, degenerative disease, and metabolic changes.[33–36] Although not characteristic radiographic disease features of PsA, both DISH and DDD may occur in patients with psoriatic disease.[37,38] In a previous study, we demonstrated an association between DISH and the presence of radiographic damage to peripheral joints.[38] Calcaneal spurs, although commonly secondary to degenerative disease, may also represent enthesophytic manifestation of PsA.[39] Thus, although MetS is associated with higher PsA disease activity, it is not associated with higher PsA-related radiographic damage to axial or peripheral joints.

It should be emphasized that factors associated with MetS, such as higher BMI and lower physical activity, are modifiable[40]; tight control and healthy lifestyle changes can result in favorable health-related outcomes, especially in the cardiovascular domain. In PsA, weight loss was shown to be associated with a higher rate of minimal disease activity in patients with overweight or obesity taking TNFi.[41] Klingberg et al[42] found that short-term weight loss treatment with a very low energy diet correlated favorably with disease activity in joints, entheses, and skin in patients with PsA and obesity.

Our study has several strengths, including the large number of patients, the organized and well-documented nature of the follow-up, and detailed clinical and radiographic data. We do acknowledge some limitations, including the retrospective design and that it was conducted at a single center.

In conclusion, MetS was not associated with more severe peripheral or axial radiographic damage in our study. On the other hand, it was found to be associated with metabolic and degenerative changes in the spine and entheses, most likely due to shared underlying pathomechanisms. These imaging changes, although not being inherent radiographic features of PsA, may have untoward impacts on patient symptoms and quality of life. The relationship between MetS and PsA, and potential mitigating strategies, merit further investigation.

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 Chandran 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. Alinaghi F, Calov M, Kristensen LE, et al. Prevalence of psoriatic arthritis in patients with psoriasis: a systematic review and meta-analysis of observational and clinical studies. J Am Acad Dermatol 2019;80(1):251–265.e19.
2. FitzGerald O, Ogdie A, Chandran V, et al. Psoriatic arthritis. Nat Rev Dis Primers 2021;7(1):59.
3. Ogdie A, Schwartzman S, Husni ME. Recognizing and managing comorbidities in psoriatic arthritis. Curr Opin Rheumatol 2015;27(2):118–126.
4. Bond SJ, Farewell VT, Schentag CT, et al. Predictors for radiological damage in psoriatic arthritis: results from a single centre. Ann Rheum Dis 2007;66(3):370–376.
5. Gladman DD, Farewell VT, Wong K, et al. Mortality studies in psoriatic arthritis: results from a single outpatient center. II. Prognostic indicators for death. Arthritis Rheum 1998;41(6):1103–1110.
6. Wong K, Gladman DD, Husted J, et al. Mortality studies in psoriatic arthritis: results from a single outpatient clinic. I. Causes and risk of death. Arthritis Rheum 1997;40(10):1868–1872.
7. Garcia-Leal M, Reyes-Soto MA, Hernandez-Galarza I, et al. Does current evidence on disease-modifying antirheumatic drugs for psoriatic arthritis reinforce an effect on radiographic progression? Results from a systematic review and meta-analysis. Clin Rheumatol 2021;40(9):3499–3510.
8. Huang PL. A comprehensive definition for metabolic syndrome. Dis Model Mech 2009;2(5–6):231–237.
9. Karmacharya P, Ogdie A, Eder L. Psoriatic arthritis and the association with cardiometabolic disease: a narrative review. Ther Adv Musculoskelet Dis 2021;13:1759720X21998279.
10. Urruticoechea-Arana A, Castañeda S, Otón T, et al. Prevalence of metabolic syndrome in psoriatic arthritis: systematic literature review and results from the CARMA cohort. J Clin Rheumatol 2022;28(2):e388–e396.
11. Guembe MJ, Fernandez-Lazaro CI, Sayon-Orea C, et al; RIVANA Study Investigators. Risk for cardiovascular disease associated with metabolic syndrome and its components: a 13-year prospective study in the RIVANA cohort. Cardiovasc Diabetol 2020;19(1):195.

12. Tylutka A, Morawin B, Walas Ł, et al. Assessment of metabolic syndrome predictors in relation to inflammation and visceral fat tissue in older adults. Sci Rep 2023;13(1):89.
13. Azevedo S, Santos-Faria D, Leite Silva J, et al. Obesity, metabolic syndrome and other comorbidities in rheumatoid arthritis and psoriatic arthritis: influence on disease activity and quality of life. Acta Reumatol Port 2019;44(4):322–324.
14. Costa L, Caso F, Ramonda R, et al. Metabolic syndrome and its relationship with the achievement of minimal disease activity state in psoriatic arthritis patients: an observational study. Immunol Res 2015; 61(1-2):147–153.
15. Haroon M, Gallagher P, Heffernan E, et al. High prevalence of metabolic syndrome and of insulin resistance in psoriatic arthritis is associated with the severity of underlying disease. J Rheumatol 2014;41(7): 1357–1365.
16. Xue Y, Jiang L, Cheng Q, et al. Adipokines in psoriatic arthritis patients: the correlations with osteoclast precursors and bone erosions. PLoS One 2012;7(10):e46740.
17. Taylor W, Gladman D, Helliwell P, et al; CASPAR Study Group. Classification criteria for psoriatic arthritis: development of new criteria from a large international study. Arthritis Rheum 2006;54(8):2665–2673.
18. Rahman P, Gladman DD, Cook RJ, et al. Radiological assessment in psoriatic arthritis. Br J Rheumatol 1998;37(7):760–765.
19. van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. Arthritis Rheum 1984;27(4):361–368.
20. Gladman DD, Chandran V. Observational cohort studies: lessons learnt from the University of Toronto Psoriatic Arthritis Program. Rheumatology (Oxford) 2011;50(1):25–31.
21. Alberti KGMM, Eckel RH, Grundy SM, et al; International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; International Association for the Study of Obesity. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. Circulation 2009;120(16): 1640–1645.
22. Maksymowych WP, Mallon C, Morrow S, et al. Development and validation of the Spondyloarthritis Research Consortium of Canada (SPARCC) Enthesitis Index. Ann Rheum Dis 2009;68(6):948–953.
23. Schoels MM, Aletaha D, Alasti F, et al. Disease activity in psoriatic arthritis (PsA): defining remission and treatment success using the DAPSA score. Ann Rheum Dis 2016;75(5):811–818.
24. Resnick D, Niwayama G. Radiographic and pathologic features of spinal involvement in diffuse idiopathic skeletal hyperostosis (DISH). Radiology 1976;119(3):559–568.
25. di Minno MND, Peluso R, Iervolino S, et al. Obesity and the prediction of minimal disease activity: a prospective study in psoriatic arthritis. Arthritis Care Res (Hoboken) 2013;65(1):141–147.
26. Eder L, Thavaneswaran A, Chandran V, et al. Obesity is associated with a lower probability of achieving sustained minimal disease activity state among patients with psoriatic arthritis. Ann Rheum Dis 2015; 74(5):813–817.
27. Højgaard P, Glintborg B, Kristensen LE, et al. The influence of obesity on response to tumour necrosis factor-α inhibitors in psoriatic arthritis: results from the DANBIO and ICEBIO registries. Rheumatology (Oxford) 2016;55(12):2191–2199.
28. Singh S, Facciorusso A, Singh AG, et al. Obesity and response to anti-tumor necrosis factor-α agents in patients with select immune-mediated inflammatory diseases: a systematic review and meta-analysis. PLoS One 2018;13(5):e0195123.
29. Visser M, Bouter LM, McQuillan GM, et al. Elevated C-reactive protein levels in overweight and obese adults. JAMA 1999;282(22):2131–2135.
30. Eder L, Abji F, Rosen CF, et al. The association between obesity and clinical features of psoriatic arthritis: a case-control study. J Rheumatol 2017;44(4):437–443.
31. Polachek A, Al-Johani R, Li S, et al. Late onset psoriatic arthritis in a longitudinal cohort: disease presentation, activity over time and prognosis. Semin Arthritis Rheum 2019;48(5):834–839.
32. Chandran V, Tolusso DC, Cook RJ, et al. Risk factors for axial inflammatory arthritis in patients with psoriatic arthritis. J Rheumatol 2010; 37(4):809–815.
33. Hirode G, Wong RJ. Trends in the prevalence of metabolic syndrome in the United States, 2011-2016. JAMA 2020;323(24):2526–2528.
34. Kos N, Gradisnik L, Velnar T. A brief review of the degenerative intervertebral disc disease. Med Arch 2019;73(6):421–424.
35. Mader R, Verlaan JJ, Eshed I, et al. Diffuse idiopathic skeletal hyperostosis (DISH): where we are now and where to go next. RMD Open 2017;3(1):e000472.
36. Weiss E. Calcaneal spurs: examining etiology using prehistoric skeletal remains to understand present day heel pain. Foot (Edinb) 2012; 22(3):125–129.
37. Ayan G, Sadic A, Kilic L, et al. Degenerative and inflammatory osteoproliferations in lumbar radiographs in psoriatic arthritis patients. J Clin Med 2022;11(7):2009.
38. Haddad A, Thavaneswaran A, Toloza S, et al. Diffuse idiopathic skeletal hyperostosis in psoriatic arthritis. J Rheumatol 2013;40(8):1367–1373.
39. Gladman DD, Abufayyah M, Salonen D, et al. Radiological characteristics of the calcaneal spurs in psoriatic arthritis. Clin Exp Rheumatol 2014;32(3):401–403.
40. Garcia KC, Confortin SC, Meneghini V, et al. Metabolic syndrome and its association with changes in modifiable risk factors: Epifloripa Aging Study. J Diabetes Metab Disord 2022;21(1):77–84.
41. Di Minno MND, Peluso R, Iervolino S, et al; CaRRDs Study Group. Weight loss and achievement of minimal disease activity in patients with psoriatic arthritis starting treatment with tumour necrosis factor α blockers. Ann Rheum Dis 2014;73(6):1157–1162.
42. Klingberg E, Bilberg A, Björkman S, et al. Weight loss improves disease activity in patients with psoriatic arthritis and obesity: an interventional study. Arthritis Res Ther 2019;21(1):17.

Arthritis Care & Research
Vol. 77, No. 7, July 2025, pp 873–880
DOI 10.1002/acr.25506

American College of Rheumatology
Empowering Rheumatology Professionals

Prevalence and Clinical Characteristics of Vasculitis in the Alaska Native and American Indian Peoples of Alaska

Ben A. Henderson, Vivek R. Mehta, Peter Holck, Tammy L. Choromanski, Amy Wilson, Flora Lee, and Elizabeth D. Ferucci

Objective. Our objective was to determine the prevalence and clinical characteristics of vasculitis in Alaska Native and American Indian (AN/AI) peoples of Alaska.

Methods. We queried the electronic health records of participating tribal health organizations within the Alaska Tribal Health System (ATHS) to identify adults with diagnostic codes related to vasculitis. Medical record abstraction was performed for all adults with potential vasculitis to confirm fulfillment of inclusion criteria, subtype, and clinical characteristics. The denominator for prevalence was the 2019 ATHS user population ≥ 18 (except giant cell arteritis [GCA], defined for persons ≥ 50).

Results. The age-adjusted prevalence per 1,000,000 AN/AI adults was 752 (95% confidence interval [CI] 581–959) for all vasculitis, with systemic vasculitis being the most common at 518 (95% CI 379–695). The most prevalent types of systemic vasculitis were antineutrophil cytoplasmic antibody–associated vasculitis (AAV) at 340 per million adults (95% CI 230–488) and GCA at 28 per 100,000 ≥ 50 (95% CI 12–56). The most prevalent subtype of AAV was granulomatosis with polyangiitis (GPA) at 244 per million adults (95% CI 148–380). AAV was diagnosed at a mean age of 54.2 years (SD 17), often with high markers of inflammation and renal involvement. GCA was diagnosed at a mean age of 69.6 years (SD 9.2).

Conclusion. The prevalence of AAV (especially GPA) in AN/AI peoples is high. GCA prevalence is lower than White populations, but higher than many other populations. AN/AI peoples with AAV and GCA may present at younger ages with more severe disease than other populations.

INTRODUCTION

Vasculitis describes a group of rare autoimmune diseases that involve inflammation of blood vessel wall.[1] Systemic vasculitis is frequently associated with high morbidity and mortality.[2] Epidemiologic studies of vasculitis have been limited by disease rarity, changing classification criteria,[3] and varying study methodologies. Limited studies have found significant differences in the prevalence and manifestations of vasculitis between different geographic regions and populations.[3] Despite the differences in populations found in the literature, most studies in the United States have reported on regions with predominately White populations. The prevalence and clinical manifestations of many subtypes of systemic vasculitis are unreported in the Indigenous North American peoples.

Some subtypes of systemic vasculitis have been studied more than others, including giant cell arteritis (GCA) and antineutrophil cytoplasmic antibody (ANCA)–associated vasculitis (AAV). GCA is a subtype of large-vessel vasculitis with a global estimated prevalence of 51.74 per 100,000 people aged ≥ 50 years old, with significant variation by region.[4] The global annual incidence of GCA is approximately 10 cases per 100,000 people aged ≥ 50 years old, with highest incidence in populations of Scandinavian descent.[4] A prior study of GCA in the Alaska Native peoples reported an annual incidence from 1983 to 2003 of approximately one per 100,000 people aged ≥ 50 years old.[5] This study required a positive temporal artery biopsy for inclusion, which may increase specificity but decrease sensitivity for GCA classification compared with studies using broader criteria, which have reported higher incidence rates.[4,6–9] AAV is a category of small-

Supported by the National Institute on Minority Health and Health Disparities, NIH (award R01-MD-014664).

Ben A. Henderson, BS, Vivek R. Mehta, MD, Peter Holck, PhD, MPH, Tammy L. Choromanski, MPH, Amy Wilson, PhD, RN, Flora Lee, RN, Elizabeth D. Ferucci, MD, MPH: Alaska Native Tribal Health Consortium, Anchorage, Alaska.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25506.

Address correspondence via email to Elizabeth D. Ferucci, MD, MPH, at edferucci@anthc.org.

Submitted for publication September 3, 2024; accepted in revised form February 3, 2025.

SIGNIFICANCE & INNOVATIONS

- This study is the first to report on the prevalence and manifestations of many types of vasculitis in the Alaska Native and American Indian peoples.
- The prevalence of antineutrophil cytoplasmic antibody–associated vasculitis (AAV) in the Alaska Native and American Indian peoples of Alaska, especially granulomatosis with polyangiitis, is high compared with other studied populations. The prevalence of giant cell arteritis (GCA) is lower than the global pooled prevalence.
- Alaska Native and American Indian adults with AAV and GCA may present with more severe clinical features and at younger ages than other studied populations.
- Although prior studies have found high rates of hepatitis B (HBV)–associated vasculitis, prevalence is now low, with no individuals with HBV-associated vasculitis identified during the study period.

vessel vasculitis with prevalence in Olmsted County, Minnesota, of 421 per 1,000,000 adults (aged ≥ 18 years) as of January 1, 2015,[10] which is significantly higher than the global pooled prevalence of AAV of 198 per 1,000,000 persons.[11] Prior studies have identified regional differences in the ratios of granulomatosis with polyangiitis (GPA) and microscopic polyangiitis (MPA) cases,[3,11] increasing incidence at higher latitudes, and significant differences among different populations living in nearby regions.[12] More studies are needed to identify the etiology of these differences, particularly in populations underrepresented in the current data.[11] Other than GCA, the only epidemiologic study of vasculitis in Alaska Native and American Indian (AN/AI) adults comes from a 1974 to 1978 report of the annual incidence of hepatitis B (HBV)–associated vasculitis in the Alaska Native peoples living in Southwest Alaska. This study was performed when HBV was endemic and reported an annual incidence of HBV-associated vasculitis of 150 per million people, which was 75 times higher than the incidence reported in Michigan and the highest in the world at the time.[13] A follow-up study in the same region after a mass HBV immunization effort did not report any new cases of HBV-associated vasculitis from 1989 until the end of the study period in 2004.[14]

Although most vasculitis subtypes have either never been studied in the Indigenous North American peoples or were studied before updated classification criteria, health disparities of other rheumatic diseases have been well-documented in the Indigenous North American peoples.[15–21] The objective of this study was to determine the prevalence of vasculitis in the AN/AI peoples of Alaska. Secondary objectives were to report the prevalence of the individual types of vasculitis according to the 2012 Chapel Hill Consensus Conference nomenclature[1] and to document the associated clinical characteristics and treatment patterns to allow for comparison with other populations and identification of potential health disparities in the population.

PATIENTS AND METHODS

Ethics approval and consent to participate. This observational study was approved by the Alaska Area Institutional Review Board (AAIRB) as expedited research (AAIRB# 2019-03-021) with a waiver of informed consent. In addition to AAIRB approval, tribal approval and privacy consults were obtained from the Alaska Native Tribal Health Consortium (ANTHC), the Southcentral Foundation Board of Directors, and other participating regional tribal health organizations (THOs) prior to starting the study. This manuscript was also reviewed and approved by ANTHC, Southcentral Foundation, and other participating THOs before journal submission.

Study population and clinical services. The Alaska Tribal Health System (ATHS) is a network of regional THOs operating health and health-related programs under a self-governance compact agreement with the US Indian Health Service (IHS) that serves all of the state of Alaska. Within the ATHS, rheumatology specialty services have been available for more than 35 years through the Alaska Native Medical Center (ANMC)–the tertiary care center for AN/AI peoples statewide. It has been standard practice for providers in the ATHS to refer people with suspected rheumatic diseases to ANMC or 1 of the 12 regional field clinics, with telemedicine follow-up when appropriate. Most THOs within the ATHS have adopted a shared Cerner platform as their electronic health record (EHR), allowing medical record access across the ATHS. This study was developed using existing data within the EHR collected for nonresearch purposes (ie, medical services).

Case ascertainment. For participating THOs using the shared Cerner EHR platform, we used queries of the EHR to identify adults with a potential diagnosis of vasculitis by using International Classification of Diseases, Ninth Revision (ICD-9) codes related to vasculitis for visits occurring from January 1, 2012, to September 30, 2015, and ICD-10 codes related to vasculitis for visits occurring from October 1, 2015, to December 31, 2019. This study was funded starting in 2019 and the end date of December 31, 2019, was selected at that time in order to allow the necessary time for institutional review board and tribal approvals, data collection and compilation from multiple sources, and data analysis. These queries excluded individuals who were not AN/AI, defined as "non-Indian beneficiaries" by the IHS and individuals who were not aged ≥ 18 years old on December 31, 2019. ICD-9 codes queried included 446.0 (polyarteritis nodosa [PAN]), 446.2 (hypersensitivity angiitis), 446.4 (GPA), 446.5 (GCA), 446.7 (Takayasu arteritis), and 447.6 (arteritis, nonspecified). ICD-10 codes queried included D69.0 (allergic

purpura), M31.0 (hypersensitivity angiitis), M30.0 (PAN), M31.3 (GPA), M31.4 (Takayasu arteritis), M31.5 (GCA with polymyalgia rheumatica), M31.6 (other GCA), M31.7 (MPA), M31.8 (other specified necrotizing vasculopathies), and M31.9 (necrotizing vasculopathy, unspecified). For collaborating THOs not using the shared Cerner platform, we used the same methods for locally performed EHR queries. All data from queries of the shared Cerner platforms and regional THOs were combined into one database using the REDCap platform[22,23] hosted on a secure ANTHC server for abstraction, review, and deidentification before data analysis.

The denominator used for the 2019-point prevalence of all categories and subtypes except GCA was the 2019 ATHS adult user population (aged ≥ 18 years old) of participating THOs, which is defined by the IHS as the number of AN/AI individuals who received medical or dental care within the ATHS at least once during the previous three fiscal years. The denominator used for the 2019-point prevalence of GCA was the 2019 ATHS user population aged ≥ 50 years old to reflect reporting used in other studies[3,4,9,24,25] and the classification criteria including[6] or requiring[7,26] age ≥ 50 years for classification.

Medical record abstraction. Medical records for all adults with a potential diagnosis of vasculitis were abstracted using a standardized data abstraction form in REDCap. Data collected included patient demographics, laboratory findings at diagnosis, histopathology, imaging, medications ever used for treatment, and the clinical manifestations ever attributed to vasculitis until December 31, 2019. A research nurse abstractor trained by the Principal Investigator (EDF) performed the initial abstraction. This training included an overview of vasculitis subtypes,[1] clinical manifestations, classification criteria, and detailed instructions and examples for each element in the standard data abstraction form, following a provided data dictionary. Clinical manifestations abstracted included those found in the Birmingham Vasculitis Activity Score (BVAS), version 3.[27] After abstraction, two rheumatologists (VRM and EDF) reviewed all abstracted records for quality assurance to confirm the presence or absence of diagnosed vasculitis. An additional abstractor (BAH) reviewed medical records of all adults with a validated diagnosis of vasculitis to ensure completeness of the abstracted demographics, diagnostic testing, medications ever used for treatment, and clinical characteristics.

Case definitions. The primary case definition for each type of vasculitis was a confirmed diagnosis by a rheumatologist or a confirmed diagnosis by a provider that met classification criteria including documentation excluding potential vasculitis mimics. Adults with a potential diagnosis of leukocytoclastic vasculitis (LCV) limited to the skin not diagnosed by a rheumatologist were included if there was a positive biopsy or a diagnosis from a dermatologist with supportive clinical findings and exclusion of alternative diagnoses.[28]

Statistical analysis. The 2019-point prevalence for vasculitis overall, systemic vasculitis overall, and each subtype of systemic vasculitis was calculated by dividing the number of adults with vasculitis meeting our case definition prevalent on December 31, 2019, by the 2019 adult ATHS user population of participating THOs (or the user population aged ≥ 50 years for GCA). Prevalence was age-adjusted to the 2,000 projected US population.[29] The 95% confidence intervals (95% CIs) were calculated using the method of Clopper and Pearson[30] for unadjusted prevalence and the method of Fay and Kim[31] for age-adjusted prevalence. For vasculitis subtypes present in less than five individuals, the unadjusted prevalence was censored and clinical characteristics were not reported to protect confidentiality, per our agreement with the institutional review board and reviewing THOs. Differences between characteristics of people with GPA or MPA were assessed using two-tailed *t*-tests and Fisher's exact tests for continuous and categorical variables, respectively. *P* values ≤ 0.05 were considered statistically significant. Missing values for laboratory or pathology results were excluded from the analyses and noted in the tables when not available. Statistical analyses were performed using R 4.2.3.[32]

RESULTS

The flow chart for the inclusion of adults with potential vasculitis is presented in Figure 1. The 2019 ATHS user population aged ≥ 18 years of participating THOs was 93,720. Of the 152 adults with potential vasculitis diagnosis identified from the electronic medical record using ICD-9 and ICD-10 codes, 74 were validated to have vasculitis. All 74 adults with validated diagnosis of vasculitis were included in reporting clinical characteristics, whereas only the 63 alive on December 31, 2019, were included in the 2019-point prevalence. Subtypes with less than five adults were grouped to protect confidentiality. There were no individuals with a diagnosis of PAN, HBV-associated vasculitis, or hepatitis C–associated vasculitis identified during the study period.

The 2019-point prevalence of vasculitis in AN/AI adults by major category and individual subtype is presented in Table 1. The age-adjusted prevalence per 1,000,000 adults was 752 (95% CI 581–959) for all vasculitis, 518 (95% CI 379–695) for primary systemic vasculitis, 189 (95% CI 112–303) for LCV limited to the skin, and 44 (95% CI 14–112) for drug-associated vasculitis. The age-adjusted prevalence of primary systemic vasculitis subtypes per 1,000,000 adults was 340 (95% CI 230–488) for AAV, 11 (95% CI 0–62) for Takayasu arteritis, 49 (95% CI 18–116) for Henoch-Schönlein purpura, 10 (95% CI 0–59) for vasculitis associated with systemic disease, and 28 (95% CI 12–56) per 100,000 aged ≥ 50 years for GCA. The age-adjusted prevalence per 1,000,000 adults of each AAV subtype was

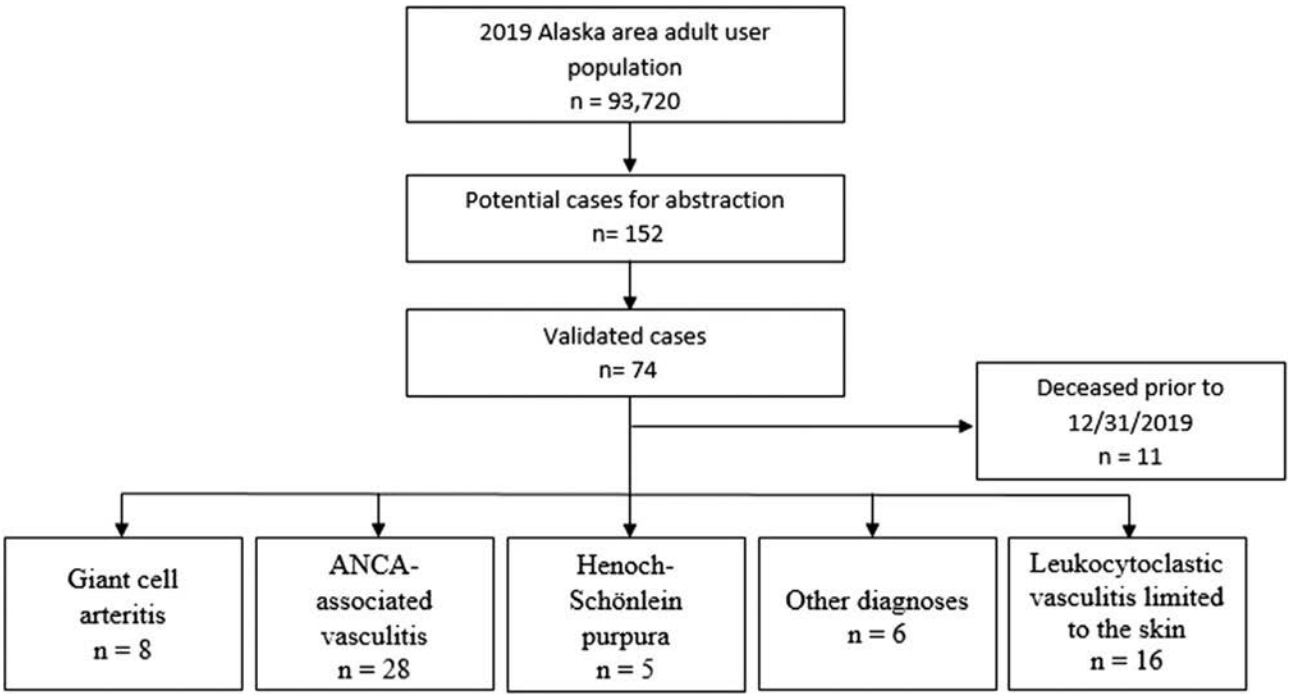

Figure 1. Flowchart for the inclusion of adults with vasculitis in the 2019 prevalence calculation. Adults with a potential diagnosis of vasculitis were identified based on International Classification of Disease codes or recorded diagnoses. Medical records were abstracted and validated by the authors to confirm fulfillment of the case definition. Subtypes with <5 adults were grouped to protect confidentiality. ANCA, antineutrophil cytoplasmic antibody.

244 (95% CI 148–380) for GPA, 86 (95% CI 37–172) for MPA, and 11 (95% CI 0–62) for eosinophilic granulomatosis with polyangiitis (EGPA).

The demographics, treatment patterns, and laboratory findings at diagnosis for people with AAV subtypes are summarized in Table 2, excluding EGPA to protect confidentiality because the sample size for this subtype was less than five. The mean age of diagnosis was 52.2 years for GPA (SD 15.8) and 59.7 years for MPA (SD 18.5). Approximately 57% of people with AAV (excluding EGPA) were female sex, with no significant difference between GPA and MPA. All people with AAV were treated with systemic corticosteroids at some point in their disease course. After corticosteroids, the most commonly prescribed treatments were rituximab (72% of GPA and 40% of MPA) and cyclophosphamide (52% of GPA and 70% of MPA). Markers of inflammation were high at the time of diagnosis for both GPA and MPA,

Table 1. Unadjusted and age-adjusted 2019-point prevalence of vasculitis in Alaska Native and American Indian adults*

Category	2019 prevalent, n	2019 unadjusted (95% CI)	2019 age-adjusted (95% CI)
All vasculitis	63	672 (517–860)	752 (581–959)
Primary systemic vasculitis	43	459 (332–618)	518 (379–695)
All AAV	28	299 (199–432)	340 (230–488)
GPA	20	213 (130–330)	244 (148–380)
MPA	7	75 (30–154)	86 (37–172)
EGPA	<5	—	11 (0–62)
Giant cell arteritis[a]	8	25 (11–49)	28 (12–56)
Takayasu arteritis	<5	—	11 (0–62)
Henoch-Schönlein purpura	5	53 (17–125)	49 (18–116)
Vasculitis associated with systemic disease	<5	—	10 (0–59)
LCV limited to the skin	16	171 (98–277)	189 (112–303)
Drug-associated vasculitis	<5	—	44 (14–112)

* Per 1,000,000 adults aged ≥ 18 years old, unless otherwise specified. 95% CI, 95% confidence interval; AAV, ANCA-associated vasculitis; EGPA, eosinophilic granulomatosis with polyangiitis; GPA, granulomatosis with polyangiitis; LCV, leukocytoclastic vasculitis; MPA, microscopic polyangiitis.
[a] Per 100,000 adults aged ≥ 50 years old. The subtypes of primary systemic vasculitis were AAV, giant cell arteritis, Takayasu arteritis, Henoch-Schönlein purpura, and vasculitis associated with systemic disease. For subtypes with <5 individuals, the number is reported as <5 and unadjusted prevalence is not reported to protect confidentiality.

Table 2. Demographics, treatment patterns, and laboratory findings at diagnosis of Alaska Native and American Indian adults with ANCA-associated vasculitis*

Characteristic	GPA (n = 25)	MPA (n = 10)	*P* value (two-tailed)
Age at diagnosis, mean (SD), y	52.2 (15.8)	59.7 (18.5)	0.235
Female sex	14 (56)	6 (60)	1.00
Corticosteroids ever	25 (100)	10 (100)	1.00
Rituximab ever	18 (72)	4 (40)	0.123
Cyclophosphamide ever	13 (52)	7 (70)	0.458
Methotrexate ever	8 (32)	2 (20)	0.686
Azathioprine ever	5 (20)	2 (20)	1.00
ESR (mm/hr), mean (SD)[a]	72.6 (33.3)	81.9 (32.0)	0.506
CRP (mg/dL), mean (SD)[a]	6.81 (6.57)	11.79 (8.54)	0.095
Creatinine (mg/dL), mean (SD)	3.0 (3.6)	3.2 (2.0)	0.870
Creatinine ≥ 1.41 mg/dL	13 (52)	9 (90)	0.055
Abnormal urinalysis[a]	18/23 (78)	8/9 (89)	0.648
Sediment	13/23 (57)	7/9 (78)	0.422
Proteinuria >1+	14/23 (61)	8/9 (89)	0.210
RBCs ≥ 3/hpf	17/23 (74)	7/9 (78)	1.00
Leukocytosis (WBCs > 11,000/mm^3)	11 (44)	5 (50)	1.00
ANCA positive	24 (96)	10 (100)	1.00
p-ANCA (IF)	3 (12)	7 (70)	**<0.01**
MPO-ANCA (ELISA)	5 (20)	9 (90)	**<0.01**
c-ANCA (IF)	14 (56)	0 (0)	**<0.01**
PR3-ANCA (ELISA)	18 (72)	0 (0)	**<0.01**
Patients biopsied[a]	21/24 (88)	8/9 (89)	1.00
Kidney	11/21 (52)	7/8 (88)	0.110
ENT or pulmonary	8/21 (38)	0/9 (0)	0.067
Other/undetermined organ	2/21 (10)	1/8 (13)	1.00
Biopsies consistent with vasculitis, when performed	19/21 (91)	8/8 (100)	1.00

* Values are the number (%) unless indicated otherwise. Significant differences ($P < 0.05$) are bolded. ANCA, antineutrophil cytoplasmic antibody; c-ANCA, cytoplasmic-ANCA; CRP, C-reactive protein; ELISA, enzyme-linked immunosorbent assay; ENT, ear, nose, and throat; ESR, erythrocyte sedimentation rate; GPA, granulomatosis with polyangiitis; hpf, high-powered field; IF, immunofluorescence; MPA, microscopic polyangiitis; MPO-ANCA, myeloperoxidase-ANCA; p-ANCA, perinuclear-ANCA; PR3-ANCA, proteinase-3-ANCA; RBC, red blood cell; WBC, white blood cell.

[a] Results not available in electronic health records were excluded from the data, including findings for five ESRs, one CRP, two urinalyses, and one biopsy for GPA; and findings for two ESRs, two CRPs, one urinalysis, and one biopsy for MPA.

including high mean erythrocyte sedimentation rate (ESR) (mm/hr) (72.6 [SD 33.3] and 81.9 [SD 32.0], respectively) and high mean C-reactive protein (CRP) (mg/dL) (6.81 [SD 6.57] and 11.79 [SD 8.54], respectively). Renal involvement was significant in both GPA and MPA, including abnormal urinalysis (78% and 89%, respectively), high mean serum creatinine (mg/dL) (3.0 [SD 3.6] and 3.2 [SD 2.0], respectively), and a serum creatinine ≥ 1.41 mg/dL in 52% of people with GPA and 90% in people with MPA. Although MPA showed higher ESR, CRP, and creatinine than GPA, the differences were not statistically significant. The majority of people with GPA were proteinase3 (PR3)-ANCA positive (72%) and/or cytoplasmic-ANCA positive (56%), and the majority of people with MPA were myeloperoxidase-ANCA positive (90%) and/or perinuclear-ANCA positive (70%). Approximately 88% of people with GPA and MPA underwent any biopsy, and 93% of those who underwent a biopsy had positive results.

The prevalent clinical manifestations of people with GPA (n = 25) and MPA (n = 10) by BVAS categories are presented in Table 3. The most common systems affected in GPA were ear, nose, and throat (ENT) (80%), renal (80%), respiratory (72%), and general (72%). The most common ENT manifestations in GPA were bloody nasal discharge, crusting, or granuloma (52%), paranasal sinus involvement (40%), conductive and/or sensorineural hearing loss (32%), and subglottic stenosis (12%). Although not included in the BVAS, 16% of people with GPA had saddle nose deformity documented. The most common systems affected in MPA were renal (90%), respiratory (50%), and general (50%). Nervous system manifestations for GPA and MPA were less common (32% and 40%, respectively), including cranial nerve palsies (14.3%), sensory neuropathy (14.3%), and headache (14.3%). Less common manifestations included mononeuritis multiplex and stroke.

The demographics, treatment patterns, clinical manifestations, and diagnostic findings of people with GCA (n = 9) are presented in Table 4. The mean age at diagnosis was 69.6 years (SD 9.2) and 67% were female. All people with GCA were treated with corticosteroids and 44% were also treated with tocilizumab. The most common clinical manifestations were temporal headache (100%), myalgia and/or arthralgia (100%), blurred vision and/or sudden vision loss (44%), and jaw pain and/or claudication

Table 3. Clinical characteristics of adults with ANCA-associated vasculitis *

System involvement, at least one symptom	GPA (n = 25)	MPA (n = 10)	*P* value (two-tailed)
General	18 (72)	5 (50)	0.258
Cutaneous	3 (12)	1 (10)	1.00
Mucous membranes/eyes	8 (32)	1 (10)	0.235
Ear, nose, and throat	20 (80)	4 (40)	**0.041**
Respiratory	18 (72)	5 (50)	0.258
Cardiovascular	1 (4)	1 (10)	0.496
Abdominal	0 (0)	0 (0)	1.00
Renal	20 (80)	9 (90)	0.649
Nervous system	8 (32)	4 (40)	0.706

* Values are the number (%) of patients with at least one manifestation ever attributed to vasculitis within each category defined by Birmingham Vasculitis Activity Score. Specific symptoms within categories are not presented to protect confidentiality because of small sample sizes (n < 5). Significant differences ($P < 0.05$) are bolded. ANCA, antineutrophil cytoplasmic antibody; GPA, granulomatosis with polyangiitis; MPA, microscopic polyangiitis.

(44%). Approximately 89% had temporal artery biopsies performed, compared with only 22% who had vascular imaging performed. Biopsies and imaging were each positive in 50% of people who had them performed. The mean ESR was high at diagnosis at 77.3 mm/hr (SD 36.4). CRP was not performed as often as ESR (67% vs 89%, respectively), but the mean was also high at 2.89 mg/dL (SD 3.11).

DISCUSSION

The 2019 age-adjusted prevalence of all forms of vasculitis in Alaska Native and American Indian adults was 752 per 1,000,000 (95% CI 581–959). Primary systemic vasculitis was the most common category (consisting mostly of AAV and GCA), followed by LCV limited to the skin. The prevalence of HBV-associated vasculitis and PAN in the AN/AI peoples is very low, with no individuals with these diagnoses identified during the study period. AN/AI peoples with GPA and MPA were often diagnosed at relatively young ages, had manifestations of significant disease activity (high mean ESR, CRP, and creatinine), and had significant renal involvement (80% and 90% had BVAS renal manifestations). Similarly, AN/AI peoples with GCA were often diagnosed at relatively young ages, had signs of significant disease activity (high mean ESR and CRP), and had significant clinical manifestations (100% had headache, 44% had jaw pain/claudication, and 44% had blurred or loss of vision). Only 50% of people with GCA who had a temporal artery biopsy performed had positive results.

Table 4. Clinical characteristics of Alaska Native and American Indian peoples with GCA*

Characteristic	Findings (n = 9)
Age at diagnosis, mean (SD), y	69.6 (9.2)
Female sex	6 (67)
Corticosteroids only	5 (56)
Tocilizumab only	0 (0)
Prednisone and tocilizumab	4 (44)
Temporal headache	9 (100)
Myalgia and/or arthralgia	9 (100)
Blurred vision and/or sudden vision loss	4 (44)
Jaw pain and/or claudication	4 (44)
Tender, palpable, or swollen temporal artery	2 (22)
Temporal artery biopsied	8 (89)
Biopsy positive, when performed	4/8 (50)
ESR (mm/hr), mean (SD)[a]	77.3 (36.4)
CRP (mg/dL), mean (SD)[a]	2.89 (3.11)
Leukocyte count (1,000/mm^3), mean (SD)[a]	7.67 (2.42)

* Values are the number (%) unless otherwise specified. CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; GCA, giant cell arteritis.
[a] Data were not available for all patient procedures including one ESR, three CRPs, and one leukocyte count.

The age-adjusted prevalence of AAV per 1,000,000 AN/AI adults of 340 (95% CI 230–488) is significantly higher than the global pooled prevalence of 198 (95% CI 187–210),[11] although this global estimate included a few studies that defined adults as aged ≥ 16 years. The age-adjusted prevalence of AAV in AN/AI adults is the second highest in the world after the age- and sex-adjusted prevalence in Olmsted County[10] of 421 (95% CI 296–546), although the difference is not statistically significant. We report the highest estimated prevalence of GPA in the world in AN/AI adults at 244 (95% CI 148–380), although the difference from the prevalence in Olmsted County of 218 (95% CI 129–308) is not statistically significant. In addition, AN/AI peoples with AAV may present at younger ages and with more severe clinical features than other studied populations. Compared with the mean values at diagnosis reported in Olmsted County,[10] AN/AI peoples with GPA were 3.9 years younger, had an ESR 2.3 times higher, a CRP 1.9 times higher, and a serum creatinine 1.8 times higher. Similarly, AN/AI peoples with MPA were 8.0 years younger and had an ESR 1.4 times higher, a CRP 2.4 times higher, and a serum creatinine 1.3 times higher than in Olmsted County.[10] Given the small sample sizes in both studies, it is important to note that differences may be due to chance.

The age-adjusted prevalence estimate of GCA in the AN/AI peoples was 28 (95% CI 12–56) per 100,000 aged ≥ 50 years and was lower than the global pooled prevalence of 51.74 (95% CI 42.04–61.43),[4] although the difference is not statistically significant. Of note, approximately 45% of the weight of this global estimate comes from studies of majority White populations of

European and/or Scandinavian descent, which have been shown to have the highest prevalence in the world. Although prevalence in our study is lower than in the majority of White populations, it is higher than studies of different racial and ethnic populations included in the global estimate, including 7 (95% CI 5.71–8.29) in Tunisia,[24] 1.47 (95% CI 1.27–1.67) in Japan,[4] and 12.2 (95% CI 7.95–16.45) in Spain.[33] Although prior studies of GCA in AN/AI peoples in Alaska reported very low incidence,[4] our study seems to indicate that it is more common than previously reported. This may be partially attributed to the use of more sensitive updated criteria that do not require a positive temporal artery biopsy for diagnosis.[8] For our study specifically, only 44% of the adults confirmed to have GCA had positive biopsies that would have met the inclusion criteria of the 1983 to 2003 study.[5] AN/AI peoples with GCA were on average 6.7 years younger at diagnosis than people diagnosed in Olmsted County.[34] AN/AI peoples with GCA appeared to present with more severe clinical characteristics than in Olmsted County,[34] including headache (100% vs 73%), blurred vision (33% vs 18%), and transient vision loss (11% vs 5%), with other findings being comparable. Notably, we identified few cases (n = 9) and as such, these differences in clinical characteristics may be due to chance.

The reasons for disparities in the development and severity of vasculitis in the AN/AI peoples compared with other populations are likely multifaceted. An environmental and genetic component may explain some of our findings including high prevalence of AAV, low prevalence of GCA, a higher ratio of GPA to MPA prevalence (~3:1) compared with Olmsted County (~1.2:1), and younger ages of diagnosis for AAV and GCA. The younger ages of diagnosis seen in this study may be partially explained by the young age-distribution of the AN/AI population[35] compared with the overall US population. Growing evidence has implicated social determinants of health more in health disparities than genetic factors,[36] and future research into disparities in severity of disease should consider these factors. Although delays in diagnosis are possible given Alaska's vast geography, they are unlikely to fully explain these disparities given the increased prevalence and the availability of rheumatology coverage within the ATHS.

This study has some limitations. First, data collection was limited to what was documented in the EHR. In any study using existing data from medical records, there is a risk of incomplete data capture and underestimation of the prevalence or the frequency of specific clinical characteristics. This would be more likely for people diagnosed before the adoption of the shared EHR platform beginning in 2011. Secondly, because of the small sample sizes, the precision of our estimates and ability to establish significant differences is limited. This limitation is inherent in studies of rare diseases and in small populations, especially Indigenous North American populations, and should not preclude studies of these diseases or populations. Third, only AN/AI peoples who were seen within the ATHS from January 1, 2012, to December 31, 2019, were included in this study. If any AN/AI peoples did not receive any ATHS-provided medical services during the study period (ie, received care entirely outside of the ATHS), they would not be identified as having vasculitis. This risk was mitigated by restricting the denominator to individuals receiving medical care through the ATHS (referred to as "user population"). The ATHS is the sole provider of IHS services in a state geographically separated from the continental United States, and recent data from the Alaska Area IHS and State of Alaska show that approximately 94% of all AN/AI individuals in Alaska are included in the "user population."[35,37] Lastly, this study reports on the aggregate data from AN/AI peoples in Alaska and may not be generalizable to all Indigenous populations.

To our knowledge, this study is the first to report the prevalence of many subtypes of vasculitis in a population of Indigenous North American peoples and one of the first studies in any population to report the prevalence of vasculitis as a whole, by major categories, and individual subtypes. In addition, this is the first study to document the clinical manifestations, laboratory data, treatment patterns, and diagnostic findings of AN/AI peoples with GCA and AAV to allow for comparison with other populations studied using similar methodologies. Our data suggest that the prevalence of AAV in the AN/AI peoples of Alaska is higher than most populations and similar to Olmsted County, which previously reported the highest prevalence in the world. GCA appears to be more common in the AN/AI peoples than previously thought, with higher prevalence than many racial and ethnic populations, but lower prevalence than White populations. The AN/AI peoples may present at younger ages and with more severe clinical features in AAV and GCA, which is consistent with disparities found in other rheumatic diseases,[18–20] but sample sizes are small. While this study will help clinicians within the ATHS to practice more informed care of AN/AI peoples with vasculitis, more research is needed to identify and address reasons for health disparities, risk and protective factors for developing vasculitis, and clinically relevant differences in clinical manifestations. The findings of this study suggest that studies of vasculitis in other Indigenous populations are warranted.

ACKNOWLEDGMENTS

We would like to thank the Alaska Native Tribal Health Organizations statewide for their time reviewing our study and manuscripts.

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 Ferucci 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. Jennette JC, Falk RJ, Bacon PA, et al. 2012 Revised International Chapel Hill Consensus Conference nomenclature of vasculitides. Arthritis Rheum 2013;65:1–11.
2. Aitken M, Basu N. Improving quality of life in vasculitis patients. Rheumatology (Oxford) 2020;59(suppl 3):iii132–iii135.
3. Watts RA, Hatemi G, Burns JC, et al. Global epidemiology of vasculitis. Nat Rev Rheumatol 2022;18:22–34.
4. Li KJ, Semenov D, Turk M, et al. A meta-analysis of the epidemiology of giant cell arteritis across time and space. Arthritis Res Ther 2021; 23:82.
5. Mader TH, Werner RP, Chamberlain DG, et al. Giant cell arteritis in Alaska natives. Can J Ophthalmol 2009;44:53–56.
6. Hunder GG, Bloch DA, Michel BA, et al. The American College of Rheumatology 1990 criteria for the classification of giant cell arteritis. Arthritis Rheum 1990;33:1122–1128.
7. Ponte C, Grayson PC, Robson JC, et al; DCVAS Study Group. 2022 American College of Rheumatology/EULAR classification criteria for giant cell arteritis. Arthritis Rheumatol 2022;74:1881–1889.
8. Dua AB, Husainat NM, Kalot MA, et al. Giant cell arteritis: a systematic review and meta-analysis of test accuracy and benefits and harms of common treatments. ACR Open Rheumatol 2021;3:429–441.
9. Yates M, Graham K, Watts RA, et al. The prevalence of giant cell arteritis and polymyalgia rheumatica in a UK primary care population. BMC Musculoskelet Disord 2016;17:285.
10. Berti A, Cornec D, Crowson SC, et al. The epidemiology of antineutrophil cytoplasmic autoantibody–associated vasculitis in Olmsted County, Minnesota: a twenty-year US population-based study. Arthritis Rheumatol 2017;69:2338–2350.
11. Redondo-Rodriguez R, Mena-Vázquez N, Cabezas-Lucena MA, et al. Systematic review and metaanalysis of worldwide incidence and prevalence of antineutrophil cytoplasmic antibody (ANCA) associated vasculitis. J Clin Med 2022;11:2573.
12. Faurschou M, Helleberg M, Obel N, et al. Incidence of granulomatosis with polyangiitis (Wegener's) in Greenland and the Faroe Islands: epidemiology of an ANCA-associated vasculitic syndrome in two ethnically distinct populations in the North Atlantic area. Clin Exp Rheumatol 2013;31(suppl 75):S52–S55.
13. McMahon BJ, Bender TR, Templin DW, et al. Vasculitis in Eskimos living in an area hyperendemic for hepatitis B. JAMA 1980;244:2180–2182.
14. Hurlburt KJ, McMahon BJ, Simonetti JP, et al. Hepatitis B-associated vasculitis in Alaska natives: viral genotype, clinical and serologic outcome. Liver Int 2007;27:627–632.
15. Ferucci ED. Understanding the disproportionate burden of rheumatic diseases in Indigenous North American populations. Rheum Dis Clin North Am 2020;46:651–660.
16. Barnabe C, Jones CA, Bernatsky S, et al. Inflammatory arthritis prevalence and health services use in the First Nations and non–First Nations populations of Alberta, Canada. Arthritis Care Res (Hoboken) 2017;69:467–474.
17. McDougall C, Hurd K, Barnabe C. Systematic review of rheumatic disease epidemiology in the Indigenous populations of Canada, the United States, Australia, and New Zealand. Semin Arthritis Rheum 2017;46:675–686.
18. Ferucci ED, Templin DW, Lanier AP. Rheumatoid arthritis in American Indians and Alaska Natives: a review of the literature. Semin Arthritis Rheum 2005;34:662–667.
19. Ferucci ED, Johnston JM, Gaddy JR, et al. Prevalence and incidence of systemic lupus erythematosus in a population-based registry of American Indian and Alaska Native people, 2007-2009. Arthritis Rheumatol 2014;66:2494–2502.
20. Scofield RH, Sharma R, Pezant N, et al. American Indians have a higher risk of Sjögren's syndrome and more disease activity than European Americans and African Americans. Arthritis Care Res (Hoboken) 2020;72:1049–1056.
21. Wei J, Ketner E, Mammen AL. Increased risk of statin-associated autoimmune myopathy among American Indians. Arthritis Rheumatol 2022;74:1602–1603.
22. Harris PA, Taylor R, Thielke R, et al. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42:377–381.
23. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019;95:103208.
24. Khalifa M, Karmani M, Jaafoura NG, et al. Study group of GCAiT. Epidemiological and clinical features of giant cell arteritis in Tunisia. Eur J Intern Med 2009;20:208–212.
25. Crowson CS, Matteson EL. Contemporary prevalence estimates for giant cell arteritis and polymyalgia rheumatica, 2015. Semin Arthritis Rheum 2017;47:253–256.
26. Sait RM, Lepore M, Kwasnicki R, et al. The 2016 revised ACR criteria for diagnosis of giant cell arteritis – our case series: can this avoid unnecessary temporal artery biopsies? Int J Surg Open 2017;9: 19–23.
27. Mukhtyar C, Lee R, Brown D, et al. Modification and validation of the Birmingham Vasculitis Activity Score (version 3). Ann Rheum Dis 2009;68:1827–1832.
28. Crowson AN, Mihm MCJ Jr, Magro CM. Cutaneous vasculitis: a review. J Cutan Pathol 2003;30:161–173.
29. Klein RJ, Schoenborn CA. Age adjustment using the 2000 projected U.S. population. Healthy People 2010 Stat Notes 2001:1–10.
30. Clopper CJ, Pearson ES. The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika 1934;26:404–413.
31. Fay MP, Kim S. Confidence intervals for directly standardized rates using mid-p gamma intervals. Biom J 2016;59:377–387.
32. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing; 2023.
33. Romero-Gómez C, Aguilar-García JA, García-de-Lucas MD, et al. Epidemiological study of primary systemic vasculitides among adults in Southern Spain and review of the main epidemiological studies. Clin Exp Rheumatol 2015;33(suppl 89):S11–S18.
34. Garvey TD, Koster MJ, Crowson CS, et al. Incidence, survival, and diagnostic trends in GCA across seven decades in a North American population-based cohort. Semin Arthritis Rheum 2021;51:1193–1199.
35. Alaska Native Epidemiology Center. Alaska Native health status report: third edition. Anchorage, AK: Alaska Native Tribal Health Consortium 2021; pg. 10. Accessed February 8, 2024. https://epi.anthc.org/publications/
36. Sankar P, Cho MK, Condit CM, et al. Genetic research and health disparities. JAMA 2004;291:2985–2989.
37. Alaska Department of Labor and Workforce Development; Research and Analysis Section. Alaska population projections; 2023-2050. The State of Alaska 2024; pg. 5, 20-23. Accessed December 30, 2024. https://live.laborstats.alaska.gov/pop/projections/pub/popproj.pdf

Arthritis Care & Research
Vol. 77, No. 7, July 2025, pp 881–891
DOI 10.1002/acr.25508

Patient Perceptions of Medication Therapy for Prevention of Posttraumatic Osteoarthritis Following Anterior Cruciate Ligament Injury: A Qualitative Content Analysis

Lily M. Waddell,[1] Donald P. Mitchener,[1] Kelly C. Frier,[1] Morgan H. Jones,[2] Elena Losina,[2] Nick Bansback,[3] Liana Fraenkel,[4] Jason S. Kim,[5] Jeffrey N. Katz,[2] Faith Selzer,[2] and Adam Easterbrook[3]

Objective. Posttraumatic osteoarthritis (PTOA) accounts for nearly 12% of osteoarthritis incidences and often occurs after anterior cruciate ligament (ACL) tear. Ensuring the uptake of preventive treatments for PTOA requires that investigators and clinicians understand factors influencing patients to seek preventive therapies. This qualitative, descriptive study aimed to assess individuals' willingness to adopt a medication therapy for PTOA prevention following ACL injury.

Methods. We enrolled participants who had an ACL tear within two years of enrollment. Study individuals participated in a semistructured interview or focus group. We reviewed audio transcriptions for accuracy, and then organized the data inductively, beginning with open coding of audio transcriptions using NVivo 12. Finally, using a qualitative content analysis approach, we identified, revised, and constructed themes and subthemes.

Results. Twenty-five individuals (mean age 25 years, 60% women) participated. Participants were an average of 10 months after injury (mean 310 days, 95% confidence interval [CI] 249–371) and reported a mean Knee Injury and Osteoarthritis Outcome Score pain score of 80.3 (95% CI 74.5–86.2). We identified three main themes related to general treatment for PTOA (eg, unwanted side effects), medication treatment for PTOA (eg, concern about pill size and dose frequency), and clinical trial attributes (eg, time commitment).

Conclusion. Although participants expressed great interest in trying medication therapy for PTOA prevention, there was variability in which components of treatment mattered to them. Our results stress the importance of using qualitative approaches such as this one to inform the design of trials and treatments that real-world patients will pursue with enthusiasm.

INTRODUCTION

Osteoarthritis (OA) is a common, disabling condition that affects more than 12% of adults in the United States.[1] Posttraumatic OA (PTOA) accounts for nearly 12% of all cases of OA and occurs disproportionately in younger individuals.[2–4] PTOA occurs as a result of trauma, such as anterior cruciate ligament (ACL) rupture and subsequent surgery. Patients who experience traumatic knee injury have a three- to four-fold higher risk of developing OA compared to their uninjured peers.[5] Sustaining ACL injury early in adulthood is associated with increased utilization of total knee replacement, especially if the injury is accompanied by meniscal tear.[4] Approximately 50% of persons who undergo ACL reconstruction (ACLR) develop PTOA within 10 to 15 years of surgery.[6–8] Given its onset in early to middle adulthood, PTOA often compromises productivity and quality of life.

The earlier age at onset of PTOA compared to idiopathic OA emphasizes the need to develop treatments that prevent or delay its development following injury. Interventions that reduce the risk of PTOA could include oral and/or intraarticular medications, rehabilitation strategies, and exercise programs.[9] Metformin, a common drug used to treat type II diabetes, has been identified

Supported by the Arthritis Foundation.
[1]Lily M. Waddell, BA, Donald P. Mitchener, BS, Kelly C. Frier, BS: Brigham and Women's Hospital, Boston, Massachusetts; [2]Morgan H. Jones, MD, MPH, Elena Losina, PhD, Jeffrey N. Katz, MD, MSc, Faith Selzer, PhD: Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts; [3]Nick Bansback, MSc, PhD, Adam Easterbrook, PhD: University of British Columbia, Vancouver, Canada; [4]Liana Fraenkel, MD, MPH: Yale University, New Haven, Connecticut, and Berkshire Medical Center, Pittsfield, Massachusetts; [5]Jason S. Kim, PhD: Arthritis Foundation, Atlanta, Georgia.

Additional supplementary information cited in this article can be found online in the Supporting Information section (http://onlinelibrary.wiley.com/doi/10.1002/acr.25508).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25508.

Address correspondence via email to Adam Easterbrook, PhD, at aeasterbrook@advancinghealth.ubc.ca.

Submitted for publication March 25, 2024; accepted in revised form February 4, 2025.

SIGNIFICANCE & INNOVATIONS

- We help address the relative paucity of research exploring patient perceptions of their long-term future following anterior cruciate ligament injury, surgery, and recovery, including the elevated risk of p[22–24]osttraumatic osteoarthritis (PTOA).
- Compared to previous studies, we focus on younger individuals given that preventive measures should occur earlier and there is a gap in our understanding of the trade-offs between risk and treatment efficacy for younger people.
- To our knowledge, the present study is the first to examine patient preferences for interventions to prevent PTOA, with the goal of informing future clinical trials and clinical care for patients at risk of developing PTOA.

as a potential medication therapy to prevent early PTOA. This protection against PTOA may be attributed to its anti-inflammatory effects and role in chondrocyte metabolism, homeostasis, and autophagy.[10–18] Previous studies in animal models indicate that metformin effectively mitigates trauma-induced OA symptoms, including synovitis and cartilage degradation.[10,11,13,15,19–21] Clinical studies have shown decreased pain and inflammation in individuals with knee OA who were treated with metformin.[22–24] This qualitative study sought to inform the development of a multicenter randomized controlled trial evaluating the efficacy of metformin in delaying the onset of PTOA in individuals undergoing ACLR (the Preventing Injured Knees from Osteoarthritis: Severity Outcomes Trial).

Before designing trials of interventions to prevent or delay PTOA in persons who have sustained knee injury, researchers should assess whether individuals who would be eligible for the trial are willing to engage with and adhere to the intervention. Assessing the potential uptake of the intervention will inform the success of trial enrollment and adherence to trial protocols. Previous studies of participant preferences for the treatment of OA have focused on older populations. For example, Fraenkel et al administered a conjoint analysis survey to 304 participants (median age 57 years, recruited regardless of OA diagnosis status or level of knee pain) to estimate preferences for treatment based on risks and benefits, route of administration, and cost.[25] They determined that 59% of participants might be willing to accept a moderate degree of risk to prevent worsening of OA.[25] However, studies targeting PTOA prevention would involve younger persons who do not have OA, for whom the trade-offs between risk and treatment efficacy are not well understood.

A growing body of qualitative research examines the biopsychosocial impacts of ACL injury, including the factors influencing recovery. However, previous qualitative research on ACL injury has generally focused on return to sport and the potential for reinjury. There has been limited research on patient perceptions of longer-term risk of PTOA following ACL injury, and attendant compromise in quality of life. One qualitative study assessed perceptions around physical activity and PTOA in Canadian young adults with an intra-articular knee injury sustained 3 to 10 years before.[26] It found that many individuals accepted future PTOA as inevitable and felt they lacked control over preventing it.[26] Our study builds on these previous findings by investigating patient perceptions of PTOA risk and their willingness to adopt preventive treatments, including the willingness to participate in a hypothetical clinical trial of medication therapy to prevent PTOA.

PATIENTS AND METHODS

Participant recruitment. This study was approved by both the Mass General Brigham (MGB) institutional review board (IRB) and the University of British Columbia's (UBC) behavioral research ethics board (MGB IRB no. 2022P003225; UBC BREB no. H22-03720). Participants were recruited via purposive sampling from orthopedic practices at MGB, an academic medical center and tertiary referral center. Individuals were aged 18 to 35 years, English-speaking, and had an ACL tear documented on magnetic resonance imaging (MRI) within two years of enrollment. We excluded individuals with a history of ACL injury on either knee before two years before enrollment (to avoid recruitment of patients who may already have signs or symptoms of PTOA) or moderate to severe radiographic OA (Kellgren-Lawrence grade[27] 3 or 4) and individuals who were unable to provide informed consent.

Data collection. Participants completed either a semi-structured individual interview or small focus group (up to three participants) based on their preference for group versus individual interviewing, as well as scheduling availability. All interviews and focus groups were conducted virtually, using Zoom (Zoom Video Communications, Inc). The MGB IRB considered logging onto the Zoom meeting as implied consent to participate in this minimal-risk study. To increase comfort and safety, all focus groups except for one were composed of people with the same self-reported gender identity. Interviews lasted approximately 45 minutes, and focus groups ranged between 60 and 90 minutes. Interviews were moderated jointly by an experienced qualitative researcher and sociologist (AE), and one research assistant (LMW or KCF). We conducted one pilot interview with a 30-year-old White male relative of a study coinvestigator (with a history of ACL tear eight months before the pilot interview but no other knowledge of the research topic) to refine the flow and ensure the comprehensibility of the interview guide. This pilot interview was not included in the analysis. We conducted interviews and focus groups between March and July 2023 and stopped when saturation was met (ie, when additional interviews and focus groups stopped yielding new ideas or experiences).

The semistructured interview guide is provided as supplementary material.

Data elements. All participants completed a brief questionnaire before or after their interview. Questions asked about participants' date of injury and ACLR, current level of pain and physical activity, health care utilization (use of medications, physical therapy, and regular visits with a primary care provider), and anthropomorphic and sociodemographic factors (height, weight, age, biologic sex, gender identity, race, ethnicity, and educational attainment). With the exception of height, weight, and age, participants selected their responses to these questions from a fixed set of categories.

Data analysis. We used an inductive qualitative content analysis approach guided by the process of analysis described by Hsieh and Shannon[28] and outlined by Elo and Kyngäs.[29] This process divides the analysis into three distinct phases: preparation, organization, and reporting. In the preparation phase, research assistants (LMW, DPM, and KCF) downloaded audio transcriptions of the interviews and focus groups from Zoom and reviewed transcriptions for accuracy. We then organized the transcriptions using an inductive approach, beginning with open coding using NVivo 12 (QSR International).[29] After identifying initial codes, we identified, revised, and ultimately constructed broad themes and more specific subthemes. The development of themes incorporated a priori questions and topics addressed in the interview guide, as well as novel ideas and patterns introduced by the interview participants. Finally, in the reporting phase, we reviewed the codes again for content and classified them into binary, categorical, or ordinal groups depending on the subtheme. We assigned perception of risk to participants based on their coded responses to questions regarding their level of concern about developing PTOA (Table 1). We followed the 32-item Consolidated criteria for reporting qualitative research checklist in reporting our research, and the completed checklist can be found[30] in Supplementary Table 1.

RESULTS

Figure 1 presents participant flow throughout the study. In total, 195 individuals were prescreened, 130 were identified as eligible, and ultimately 25 individuals participated (mean age 25 years, 15 [60%] female) across four focus groups and 15 individual interviews. The number of participants who chose focus groups versus individual interviews was similar across sexes and age groups. On average, focus groups and interviews occurred 10 months after individuals' ACL injury (mean 310 days, 95%

Table 1. Perception of risk assignments and related codes*

Participant	Code	Assigned perception of risk
5	I think my risk of developing knee OA in the next 10 years is low or no more than average for my age group.	Low
9	For my surgery, I went with the option that would reduce my risk of OA later in life.	Low
13	I'm hopeful that my OA risk is lower because (1) I already feel better than I had been and (2) they didn't have to do a graft for my surgery, which should lessen the risk of complications.	Low
108	Because I work in the mental health field and I've been an athlete for most of my life, I'm not worried about my higher risk for developing OA.	Low
142	While I don't have the exact percentage, I think my risk for developing OA in the next 10 years is low.	Low
21	My surgeon told me I might develop OA, and I've been thinking about it since.	Medium
27	My injury makes me aware that OA is a real possibility, and I'm grateful I can prepare for it now.	Medium
28	I think my risk of developing knee OA in my lifetime is average.	Medium
70	I think my risk of developing knee OA in my lifetime is average.	Medium
113	With aging, I expect to have some degenerative changes, but with an active lifestyle, I think I can lower my risk.	Medium
140	Part of me feels that I won't get OA, but the other part of me feels I will.	Medium
166	I don't think I really have control over my risk for OA, so I want to keep doing what I've been doing.	Medium
168	I've been doing 12-hour shifts as an EMT since surgery and I haven't noticed much pain, so I don't think my risk is as bad as I initially thought.	Medium
171	I feel more at ease knowing that my calculated OA risk is 14% rather than the 60% I had initially predicted.	Medium
177	I don't know. I'm just trying to think of my parents and grandparents, maybe fifties? (In response to the following question: When do you think you'll start developing symptoms of OA?)	Medium
8	I think that during a person's lifetime, the average risk for OA is 35% and mine is 70%, because of my injury.	High
30	I know that I am at higher risk for OA than the average person, but I'm not entirely sure what my risk is.	High
34	I'm expecting to develop OA in the next 15 years.	High
54	My injury, recovery, and knowing that I'm at higher risk for OA—it all sucks.	High
74	I've come to terms with my higher-than-average risk of OA.	High
87	After learning about my increased OA risk, it would probably help if I lost weight.	High
102	I'm always thinking about my risk of OA because I know ACL tears can cause it.	High
110	I think my risk for OA is higher than average because I can already feel some pain after exercising.	High
149	I think my risk of developing OA in my injured knee in the next 10 years is high.	High
174	I'm fairly sure I'll be getting OA in the next 10–15 years.	High

* ACL, anterior cruciate ligament; EMT, emergency medical technician; OA, osteoarthritis.

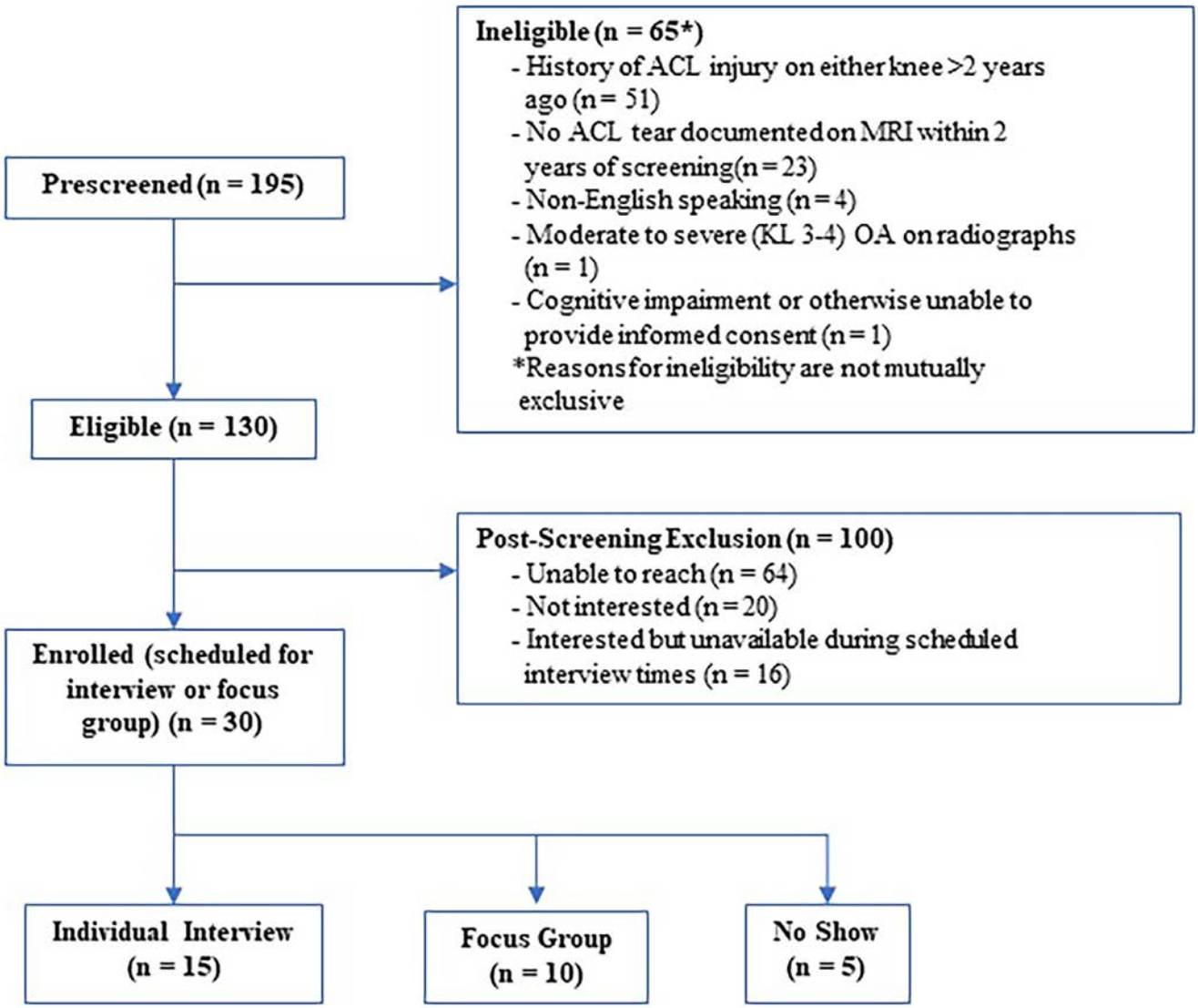

Figure 1. Participant flow diagram. ACL, anterior cruciate ligament; KL, Kellgren-Lawrence; MRI, magnetic resonance imaging; OA, osteoarthritis.

confidence interval [CI] 249–371), at which point participants reported a mean Knee Injury and Osteoarthritis Outcome Score pain score of 80.3 (95% CI 74.5–86.2, from a scale of 0–100; 100 is best). Table 2 presents a summary of participant demographics.

We identified three main themes from the interviews and focus groups: general treatment to prevent PTOA, clinical trial attributes, and medication (pill) treatment for PTOA. Within these broad themes, we identified 16 subthemes reflecting discrete characteristics of a treatment or trial that participants felt would influence their willingness to try the treatment or take part in the clinical trial. For each subtheme, we categorized participant codes into binary (more likely to participate vs less likely to participate), ordinal (small, medium, large), or categorical (headache vs nausea vs mood swings) response options. Illustrative quotes for each theme and corresponding subthemes are provided in Tables 3, 4, and 5. Supplementary Table 2 presents an exploratory tabulation of these themes, subthemes, and response options, stratified by age, gender identity, and perception of risk, to attempt to identify whether these factors may mediate patient choices about treatments and trials.

Theme 1: general treatment to prevent PTOA. Participants discussed the benefits and detriments of pursuing general treatments that might prevent PTOA, assuming that preventive treatments currently existed. Within the benefits subtheme, participants discussed the amount of reduction in the risk of developing PTOA they would expect in order to try a treatment. For detriments to engaging in a treatment, participants mentioned financial burden and perceived risk of side effects.

Table 2. Demographic characteristics of interview and focus group participants*

Variable	Value
Age at time of interview, mean (SD)	25.5 (4.1)
Gender identity, n (%)	
Woman	15 (60)
Man	10 (40)
Race, n (%)	
White	18 (72)
Asian American	2 (8)
Multiracial	2 (8)
Prefer not to state	2 (8)
Black or African American	1 (4)
Ethnicity, n (%)	
Not Hispanic or Latino	21 (84)
Hispanic or Latino	4 (16)
Days since injury, mean (SD)	310 (141)
KOOS pain, mean (SD) (0–100, 100 = best)	80.3 (14.1)

* KOOS, Knee Injury and Osteoarthritis Outcome Score.

Table 3 presents illustrative quotes for each subtheme and response option under theme 1.

Perceived benefit of a preventive treatment for PTOA. Eighteen of 25 participants (72%) discussed the level of benefit they would expect in order to choose to receive a treatment. We sorted these into expecting either small, medium, or large benefit. Six participants (33%) expressed that they would be willing to accept a "small" benefit. These participants described a desire to secure themselves any amount of pain- and impairment-free movement they could. Eight participants (44%) discussed expecting a "medium" benefit. These participants typically discussed wanting to return to a level of risk comparable to their non-injured peers. Four participants (22%) expressed wanting a "large" benefit. For these individuals, they desired an almost absolute elimination of PTOA risk, expressing that they would be unwilling to settle for anything less than treatment that almost guaranteed PTOA prevention. Of the participants who discussed treatment benefits, 83% of individuals aged 18 to 29 years required a medium or large benefit to consider adopting a preventive treatment, compared to only 33% of those aged 30 years or older.

Out-of-pocket treatment cost. Out-of-pocket costs were one of the most important factors influencing whether to try a treatment. The treatments presented were hypothetical and treatment efficacy was not quantified, such that participants often struggled to name a specific dollar amount. Only 8 of 25 participants (32%) were able to pinpoint a price range for preventive PTOA treatment. Of those who did, they were grouped into those willing to pay less than $50 per month for treatment (two

Table 3. Illustrative quotes under theme one: general treatment to prevent posttraumatic osteoarthritis*

Subtheme and response option	Illustrative quote
Benefits	
Small benefit	Even if it was just a year [delay], I'd be like, 'Okay, that's a whole other year that I can do things I love after being injured.' It's like, 'Oh, I would kill for a year to play soccer.' (Participant 149)
Medium benefit	I would want to see a reduction to something that's pretty close to what your average person experiences. (Participant 30)
Large benefit	Yeah. Personally, I would say far, far below 25%. (Participant 34) LMW: Okay, you want to be at 0%? Yeah, yeah, definitely. (Participant 34)
Cost	
<$50 per month	If it was something I had to take monthly, I would probably be annoyed if it was anything more than 30 to 40 dollars. (Participant 171)
$50–150 per month	I would pay up to $50 to $80 a month for it. (Participant 140)
>$150 per month	It's so hard to say, but I would spend probably 500 bucks a month if you could tell me I would not have knee arthritis…That would be worth it to me to sacrifice in other aspects of life. (Participant 27)
Side effects or perceived risks	
GI (nausea, diarrhea, constipation)	The big one that I don't particularly enjoy is constipation. (Participant 142)
Increased pain or stiffness	I wouldn't wanna give up any mobility in the knee so maybe if it causes stiffness or something. (Participant 113)
Headache	When I have a headache, I just don't feel like doing anything and having a few weeks from not doing anything would be really bad. So, yeah. (Participant 8)
Mental status, mood, energy, or sleep	I mean, I don't know if it was affecting my appetite…or just mental state, or like, if I felt lethargic, maybe, or just changes, in my normal bodily function or feeling, I guess that would be kind of concerning to me. (Participant 13)
Metabolism, cancer, other organ systems	I guess any systemic symptom I wouldn't be okay with. (Participant 28)
Sexual dysfunction	If you told me, 'Hey, there's a chance you're going to have to, you know, start taking Viagra going forward,' I'd probably [shy] away from this arthritis pill just because I don't want to have [that happen]. (Participant 74)
Hair loss	Honestly if you even said hair loss and like I was gonna go bald taking it…That's aesthetic but something like that would potentially drive me away. (Participant 54)

* GI, gastrointestinal.

Table 4. Illustrative quotes under theme two: attributes of a clinical trial to prevent posttraumatic osteoarthritis*

Subtheme and response option	Illustrative quote
Compensation	
Monetary	I think covering travel, and I mean as long as the pill and MRI and all that is free, I would say, maybe $100 per visit. (Participant 34)
Nonmonetary	[Referring to desire for study to offer support groups] Yeah, I think…an opportunity to share with others how things are going…to combat the loneliness of being a part of something new that no one else is going through. (Participant 108)
Placebo group	
Does not affect willingness to participate	If I do get the placebo, I'm getting money for just taking a bunch of sugar pills. So that seems reasonable. (Participant 9)
Does affect willingness to participate	I'd almost feel silly going in [for study visits]. I don't know…In my head, I'd kind of be…like, 'I did all that for nothing.' (Participant 149)
Adherence monitoring	
Phone app	I think I'd prefer the app, where you can go into it and click it. I have that right now with my PT stuff. It just vibrated that I have to do my workouts for the day! (Participant 8)
Bluetooth bottle cap	Definitely the bottle because that way, I don't have to be [on] my phone all the time. (Participant 87)
Views on MRI	
No negative effect on willingness to participate	Honestly, I kind of want to get an MRI. I'm curious how [my knee is] doing. So that sounds like an added compensation benefit, if anything. (Participant 27)
Distance	
Less than 30 mi/hr	I think once…my commute starts pushing greater than like a 30 minute drive. I think that's where it might become an inconvenience. (Participant 74)
More than 30 mi/hr	If it was just to the study…probably an hour, maybe an hour and a half. (Participant 102)
Method of recruitment	
In person	Most certainly it would definitely be more convincing if I was talking to my doctor. If someone called me out of the blue I would probably hang up immediately. (Participant 142)
Over the phone	I'd probably also want a phone call. (Participant 74)
No preference	[Re: would it affect your decision to participate if the study was presented to you in person, or if someone called you?] I think either would be fine. (Participant 28)
Person recruiting	
Member of the health care team	Not to sound snobby, but especially in the landscape of America today. In hierarchical order, my surgeon, my physical therapist, and then probably the person who I consider biased in the research study. (Participant 27)
Not a member of the health care team	I actually think I may prefer a research assistant, because I feel like they take more time to explain things. (Participant 54)
No preference	[In response to question asking if they have a preference for care team member vs research team member] Not really. As long as it doesn't set off the fraud bells in my head. (Participant 9)
Study communications	
Study website	Like a specific web page that we can access with everything there. (Participant 70)
Email	I think I'd probably read the emails more often than the website, just because it'd be in my inbox and I read my emails in the morning while I'm having a coffee. (Participant 8)

* MRI, magnetic resonance imaging; PT, physical therapy.

participants, 25%), $50 to $150 per month (two participants, 25%), or more than $150 per month (four participants, 50%) for treatment. However, especially for those in the more than $150 per month group, willingness to pay more for treatment was conditional on the treatment offering great risk reduction.

Potential side effects of treatment. Eleven of 25 participants (44%) named specific side effects that they were concerned about. Participants could name more than one side effect they were concerned about, resulting in 17 unique responses. We grouped these responses into gastrointestinal symptoms (four participants, 24%); increased pain or stiffness (three participants, 18%); headache (one participant, 6%); mental status, mood, energy, or sleep alterations (three participants, 18%); metabolic or organ dysfunction (including cancer) (three participants, 18%); sexual dysfunction (two participants, 12%); and hair loss (one participant, 6%). There was variability in the side effects that participants would, and would not, accept in a treatment. For instance, one participant indicated that they have no problem with headaches, whereas another felt that headaches were a side effect they could not tolerate. Personal and family medical history were related to which side effects participants were willing to tolerate. For example, one participant with Charcot-Marie-Tooth disease described being unfazed by muscle cramps because of her diagnosis.

Theme 2: attributes of a clinical trial to prevent PTOA. When presented with participation in a hypothetical clinical trial, participants discussed potential advantages or barriers

Table 5. Illustrative quotes under theme three: medication therapy to prevent posttraumatic osteoarthritis

Subtheme and response option	Illustrative quote
Amount (number of pills)	
Number of pills matters	I think less is more…I take a daily pill right now, and I struggle to even remember the one so three? [It's] kind of a lot versus one is more manageable. (Participant 113)
Number of pills does not matter	Oh, three in one moment. Yeah, yeah, I can do that…I would just take it with my other medication in the morning. That's it. (Participant 87)
Duration	
Less than 6 mo	Let's say, if I'm feeling really, really cooperative, and I'm just so excited about it, I would say… probably three to four months. (Participant 108)
6–12 mo	Maybe six months…a year might be pushing it. But…I wouldn't want to be taking something for life, like as a "maybe protective" effect. (Participant 113)
>12 mo or indefinitely	I think we'd be probably talking at least years because I don't know that months would be worth it…In terms of how many years, ideally like five. (Participant 113)
Size	
Prefer smaller	I know my roommate used to have to take a giant pill, and she hated it—it was the worst part of her day. (Participant 149)
No preference	I mean, when I was younger I would have a hard time swallowing pills…But I think now size wouldn't matter that much. (Participant 102)
Dose frequency	
Prefer once a day	I feel like if it's once daily or a couple of times a week, that'll be okay. But if it's something I have to do multiple times a day, and have to keep reminding myself to do it, it might be less nice. (Participant 5)
Safety profile of drug	
Prefer well-established safety and efficacy	But yeah, efficacy would play a huge role. Because if there wasn't really any evidence, I think I would be a little wary of it. (Participant 113)

to entry into a trial (compensation, MRI, distance to trial sites, method of recruitment, and person responsible for recruitment), as well as logistical considerations during the trial (presence of a placebo control group, adherence monitoring, and study communication preferences) (Table 4).

Desired compensation for participation in a clinical trial. Twenty-one participants (84%) discussed compensation in relation to participation in a clinical trial. However, especially for participants who had not participated or worked in clinical research before, it was often difficult to pinpoint a reasonable amount of financial compensation for trial participation. For this reason, only 8 of 21 participants (38%) who discussed compensation were able to identify a numeric amount of desired financial compensation (ranging from "at least minimum wage" to $50 per hour). Although exact dollar amounts varied, participants demonstrated a wish for compensation to be "fair" and convey a level of respect for the valuable commitment research participants make to helping move science forward, as well as counter the perceived risks and effort that participation would entail. Other than monetary compensation, 13 participants (62%) introduced discussed other forms of compensation. For example, several participants independently described wanting a trial to incorporate opportunities to meet with other people recovering from ACL injury, in an open forum or support group format. Similar to monetary compensation, participants felt that these kinds of nonmonetary compensation helped to make study participants feel like valued members of the community and important contributors to science.

Presence of a placebo control group. Sixteen participants (64%) discussed their perceptions of potentially receiving a placebo instead of the active treatment. Of these, 11 participants (69%) accepted a placebo as a critical part of the clinical research process and felt that the other benefits of participating in the study (eg, financial compensation, free MRIs, etc) made up for potentially not receiving the active treatment. Five participants (31%) felt strongly that the presence of a placebo group soured the research experience, evoking feelings of frustration, embarrassment at having to come in for continued study visits despite taking a placebo, and futility.

Adherence monitoring for a clinical trial protocol. Participants were asked to discuss three options for medication adherence monitoring: pen and paper diary, reminder app, and pill bottle monitor that tracks how often the bottle is opened. Seventeen participants (68%) discussed adherence monitoring. No participants discussed preferring the traditional drug diary.

> "No! Oh, no, like I'm a graphic designer. I do a lot of things on [the] computer. I only write on paper whenever it's [my] grocery list or my to-do list." (Selina, a 30-year-old woman; first names, when used, are pseudonyms to protect participant confidentiality)

Instead, participants desired a tech-friendly approach to medication adherence monitoring, with nine participants (53%) indicating they would prefer using a phone app to track adherence, and another eight participants (47%) opting for an electronic adherence monitoring device, such as Medication Event Monitoring System caps, which take the burden of tracking adherence out of the hands of the participant.[31]

MRI as a trial component. Fourteen participants (52%) discussed their comfort with completing an MRI scan as part of the study. None of these 14 participants stated that MRI would make them less likely to participate. In fact, several participants were enthusiastic about the prospect of undergoing an MRI scan, particularly when the cost is covered by the research study and a physician is available to discuss the results.

Distance to clinical trial site. Participants discussed distance and time to the study site as factors that might influence participation. Of the 11 participants (44%) who spoke about distance as a factor, five participants (45%) stated they would not be willing to travel more than an hour to the study site (approximately 30 miles with traffic). Six participants (55%) stated they would be willing to travel more than one hour or 30 miles, with one participant even offering to fly up to Boston from Georgia for study visits.

Method of recruitment into the trial. Eight participants (32%) discussed their preferred way to be introduced to the study. Five participants (63%) preferred being approached about the study during a scheduled in-person visit as opposed to over the phone. One participant (13%) preferred being contacted about the study over the phone. The two remaining participants (25%) had no preference between in-person and phone recruitment.

Person who initiates recruitment into the trial. Nineteen participants (76%) discussed how they may or may not be influenced to participate based on who initiated recruitment. Eleven participants (58%) discussed preferring to be introduced to potential research studies by their surgeon, physical therapist, or another member of their clinical care team. Seven participants (37%) had no preference between members of their existing care team and nonmembers (eg, a research assistant). Only one participant (5%) preferred to be presented with research studies by a research team member as opposed to one of their existing providers. Despite differences in opinion between participants about their preference for the trial recruiter, participants universally described a desire to be presented with information about the trial by someone they deem trustworthy and unbiased, whether that person is a member of their existing care team or not.

Study communication preferences. Six of 25 participants (24%) explicitly described their preference for mode of communication with the study team. Five interviewees (71%) preferred electronic modes of communication, such as a study website or emails, as opposed to traditional hard copy mail updates. Two participants (29%) described combining emails and a website to send timely reminders to participants (email) while also ensuring that they have up-to-date information always at their fingertips (website).

Theme 3: medication therapy to prevent PTOA. *Number of pills per dose.* Participants discussed the influence of features of the pill itself (number and size) as well as of medication treatment in general (dose frequency, duration of treatment, and safety profile) that influence their decision-making around treatment (Table 5). Twelve participants (48%) described the influence of pill quantity (taken at one time) on treatment desirability (as either a negative influence or not a consideration). For 9 of 12 respondents (75%), quantity was not a consideration. However, three participants (25%) expressed that having to take multiple pills, even at one time, created confusion and potential for incorrect dosing.

Size of pill. Often, medication doses can be given as one larger pill or several smaller ones. Seventeen of 25 participants (68%) discussed the size of the medication. Six participants (35%) discussed preferring smaller pills, even if it required taking more than one in a sitting. Eleven participants (65%) felt that pill size was not an important consideration. Several participants pointed to age as a potential influence on being able to tolerate larger pills but identified opposite age groups as being less likely to tolerate larger pills.

> "When I was younger, I would have a hard time swallowing pills, so maybe if I was a younger participant, that would be difficult." (Maddie, a 20-year-old woman)

> "I've noticed that it's a big thing for older generations. Like my grandpa, he'd be like, 'No, I don't want to take those pills,' or 'I have to take too many' or stuff like that." (Jose, a 24-year-old man)

Dose frequency. Of the seven participants (28% of the total sample size) who described dose frequency as a treatment consideration, all felt that a once-daily medication is optimal. Participants expressed that dosing either more or less frequently than once daily conflicted with a lot of their other established routines, creating room for missed or incorrectly timed doses.

Duration of treatment and safety profile of medication. Nineteen of 25 participants (76%) discussed duration of treatment. Two of 19 respondents (11%) expressed being only willing to take a trial medication for a short-term period (<6 months). Six of 19 individuals (32%) felt that taking medication for six months to a year was acceptable. However, some individuals (11 of 19, 58%) expressed willingness to remain on treatment for longer than a year, and in some cases indefinitely (provided that cost was not prohibitive and side effects were nonserious). For all nine individuals (36%) who expressed medication-based safety concerns, experimental medications with relatively little supporting evidence were less appealing than medications with well-established safety profiles and a wealth of evidence to support the efficacy.

DISCUSSION

We identified three main themes based on participants' perceptions of preventive treatment and participation in a clinical trial: general preventive treatment for PTOA, attributes of a clinical trial to prevent PTOA, and medication therapy to prevent PTOA. Most participants expressed concern about their elevated risk for

developing PTOA but differed in their willingness to engage in potential preventive treatment. Generally, the higher a participant's perception of PTOA risk, the more agreeable they were about accepting some negative treatment components. Namely, they were more likely to accept smaller treatment benefit, were willing to stay on treatment for longer, and were willing to travel further to study appointments than participants with lower perceptions of risk. These findings make sense given that the more worried a participant is about developing PTOA, the greater lengths they are willing to go to prevent it.

We explored the intersection of age and gender identity with participants' experiences. With respect to age, participants who were 30 years or older were more willing to discuss trying treatments that had lower efficacy. It may be that as individuals age, they are more aware of the impact that injury and disease will have on their ability to engage in different activities, and thus they are more willing to settle for lower treatment efficacy. These participants also more frequently indicated that large pill size was a potential deterrent compared to those we interviewed who were 18 to 29 years old.

Participants' gender identities were often associated with feelings about and demonstration of internalized societal beliefs regarding what it means to be a man or woman. In turn, this influenced how they perceived treatment and their potential risk of developing PTOA.[32–36] For example, participants who identified as men seemed more concerned about how treatments influenced aspects related to "being manly," such as discussing fears of side effects to do with sexual dysfunction or hair loss, while also expressing feelings that they were doing something honorable by participating in clinical research (eg, one participant described feeling like he would be helping society if he were to be assigned to the placebo group).[35] In contrast, participants who identified as women were more likely to openly discuss a desire for emotional support through nonmonetary forms of compensation, such as support groups, and were more concerned about feeling silly or foolish if they found out they had been assigned to the placebo group. This perhaps reflects internalized anxiety among female participants about not being taken seriously in the American health care system.[32]

The results of the present study corroborate several previous qualitative studies that explored concerns among persons recovering from ACL injury about the development of PTOA. Piussi et al[37,38] identified "uncertainty about the future and osteoarthritis" and specifically found that participants, including those with satisfactory current knee function, expressed apprehension about future impairments due to OA in persons who both underwent or did not undergo ACLR. Similarly, Karlström et al[39] identified a category entitled "disrupted me" and noted that it included "obstructive thoughts about the future, including anxiety about further surgeries and need of knee replacement due to osteoarthritis." Additionally, Truong et al[40,41]conducted qualitative studies with Canadian young and middle adult participants enrolled in a randomized controlled trial evaluating a physical therapist–guided education and exercise therapy intervention. They identified a category called "regaining control of knee health" and found that intervention's education, goal setting, and social support helped change participants' previously negative beliefs about their knee health and promoted optimism in facing future knee changes.[40,41]

This study's findings regarding the impact of potential side effects on the willingness to use medication therapy for OA align with other qualitative studies among adults outside the United States. Selten et al[42] found that some middle- to older-aged adults with hip and/or knee OA in the Netherlands were hesitant to use pharmacologic treatments due to concerns about potential side effects, the risk of developing tolerance or dependence on medication, and the possibility of interfering with the body's natural signals that help prevent joint overuse. Similarly, Kamsan et al[43] reported that older adults with knee OA in Malaysia were often reluctant to use prescription medications to manage their symptoms, citing worries about long-term effects and negative previous experiences with medication side effects. However, many were open to considering traditional remedies, particularly those recommended by family or peers. Additionally, Townsend et al[44] found that many Canadian adults with rheumatoid arthritis depended on over-the-counter medications to alleviate their symptoms but were reluctant to use or preferred to limit their use of prescription medication due to side effects.

This study is strengthened by involving a multidisciplinary team with expertise in both orthopedics or arthritis research and qualitative analysis. Additionally, the sample size of this study (n = 25) allowed us to reach data saturation as determined by the authors and is in line with guidance for qualitative researchers.[45] However, we acknowledge that our quantitative results are limited because of the small sample and because percentages are derived from counts of comments in recorded conversations, rather than responses to explicit questions. Because of this, our conclusions should be interpreted cautiously. The present study is also limited by the lack of diversity within our study sample. The majority of the participants (72%) were White, which, although reflective of the overall patient population in the hospital system from which these participants were recruited, limits the generalizability of this study's findings.[46] Another limitation, and perhaps a call to change clinical practice, was a lack of standardized information on the risk of PTOA development after ACLR. Despite a wealth of research in the last decade on the elevated risk for PTOA after ACL tear, a surprising number of the interview participants indicated that this study interview was their first exposure to any kind of information about PTOA, whereas others had been thoroughly informed of the risks by their health care team, personal connections, or through independent research. These differences in exposure to trusted information about PTOA development meant that interviewed participants were not always on a level playing field. It remains unclear how

exactly receiving information about PTOA influences patient decision-making, but a number of participants expressed resentment or disappointment at not having been fully informed of their risks before participating in this interview study.

Altogether, participants expressed great interest in preventive treatments for PTOA. Further research in this area should focus not only on treatment development but on implementation—using qualitative analyses such as this one to inform the design of trials and treatments that real-world patients will pursue with enthusiasm. Future qualitative research should also expand on this work by further exploring the potential interactions between participant factors such as age, gender, race, ethnicity, and socioeconomic status and the willingness to adopt preventive strategies for PTOA.

Additionally, the results of this study stress the importance of clinicians adopting an individualized approach to recovery, rehabilitation, and subsequent care for patients following ACLR, including potential monitoring for and counseling about PTOA. Participants differed dramatically in the treatment or trial characteristics that have the greatest influence on their decision-making, and our results show that these differences may be partly rooted in differences in age, gender, and perception of risk. Providers should keep this in mind when discussing treatment choices with patients, understanding that there will never be a one-size-fits-all option that works for every patient.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Easterbrook had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Waddell, Jones, Losina, Bansback, Fraenkel, Kim, Katz, Selzer, Easterbrook.

Acquisition of data. Waddell, Mitchener, Frier, Easterbrook.

Analysis and interpretation of data. Waddell, Mitchener, Frier, Jones, Losina, Bansback, Fraenkel, Kim, Katz, Selzer, Easterbrook.

REFERENCES

1. Centers for Disease Control and Prevention. Osteoarthritis (OA). US Department of Health & Human Services; 2020. Accessed October 23, 2023. https://www.cdc.gov/arthritis/osteoarthritis/
2. Osteoarthritis Research Society International. (OARSI) ORSI. Osteoarthritis: A Serious Disease, Submitted to the U.S. Food and Drug Administration. 2016. Accessed October 23, 2023. https://oarsi.org/sites/oarsi/files/docs/2016/oarsi_white_paper_oa_serious_disease_121416_1.pdf
3. Thomas AC, Hubbard-Turner T, Wikstrom EA, et al. Epidemiology of posttraumatic osteoarthritis. J Athl Train 2017;52(6):491–496.
4. Brown TD, Johnston RC, Saltzman CL, et al. Posttraumatic osteoarthritis: a first estimate of incidence, prevalence, and burden of disease. J Orthop Trauma 2006;20(10):739–744.
5. Richmond SA, Fukuchi RK, Ezzat A, et al. Are joint injury, sport activity, physical activity, obesity, or occupational activities predictors for osteoarthritis? A systematic review. J Orthop Sports Phys Ther 2013;43(8):515–B19.
6. Barenius B, Ponzer S, Shalabi A, et al. Increased risk of osteoarthritis after anterior cruciate ligament reconstruction: a 14-year follow-up study of a randomized controlled trial. Am J Sports Med 2014;42(5):1049–1057.
7. Lohmander LS, Ostenberg A, Englund M, et al. High prevalence of knee osteoarthritis, pain, and functional limitations in female soccer players twelve years after anterior cruciate ligament injury. Arthritis Rheum 2004;50(10):3145–3152.
8. Luc B, Gribble PA, Pietrosimone BG. Osteoarthritis prevalence following anterior cruciate ligament reconstruction: a systematic review and numbers-needed-to-treat analysis. J Athl Train 2014;49(6):806–819.
9. Watt FE, Corp N, Kingsbury SR, et al. Towards prevention of post-traumatic osteoarthritis: report from an international expert working group on considerations for the design and conduct of interventional studies following acute knee injury. Osteoarthritis Cartilage 2019;27(1):23–33.
10. Li H, Ding X, Terkeltaub R, et al. Exploration of metformin as novel therapy for osteoarthritis: preventing cartilage degeneration and reducing pain behavior. Arthritis Res Ther 2020;22(1):34.
11. Li J, Zhang B, Liu WX, et al. Metformin limits osteoarthritis development and progression through activation of AMPK signalling. Ann Rheum Dis 2020;79(5):635–645.
12. Wang C, Yang Y, Zhang Y, et al. Protective effects of metformin against osteoarthritis through upregulation of SIRT3-mediated PINK1/Parkin-dependent mitophagy in primary chondrocytes. Biosci Trends 2019;12(6):605–612.
13. Wang C, Yao Z, Zhang Y, et al. Metformin mitigates cartilage degradation by activating AMPK/SIRT1-mediated autophagy in a mouse osteoarthritis model. Front Pharmacol 2020;11:1114.
14. Xu L, Ma F, Huang J, et al. Metformin hydrochloride encapsulation by alginate strontium hydrogel for cartilage regeneration by reliving cellular senescence. Biomacromolecules 2021;22(2):671–680.
15. Yan J, Ding D, Feng G, et al. Metformin reduces chondrocyte pyroptosis in an osteoarthritis mouse model by inhibiting NLRP3 inflammasome activation. Exp Ther Med 2022;23(3):222.
16. Zhang M, Liu Y, Huan Z, et al. Metformin protects chondrocytes against IL-1beta induced injury by regulation of the AMPK/NF-kappa B signaling pathway. Pharmazie 2020;75(12):632–636.
17. Sullivan JK, Huizinga J, Edwards RR, et al. Cost-effectiveness of duloxetine for knee OA subjects: the role of pain severity. Osteoarthritis Cartilage 2021;29(1):28–38.
18. Park MJ, Moon SJ, Baek JA, et al. Metformin augments anti-inflammatory and chondroprotective properties of mesenchymal stem cells in experimental osteoarthritis. J Immunol 2019;203(1):127–136.
19. Feng X, Pan J, Li J, et al. Metformin attenuates cartilage degeneration in an experimental osteoarthritis model by regulating AMPK/mTOR. Aging (Albany NY) 2020;12(2):1087–1103.
20. Yan J, Feng G, Ma L, et al. Metformin alleviates osteoarthritis in mice by inhibiting chondrocyte ferroptosis and improving subchondral osteosclerosis and angiogenesis. J Orthop Surg Res 2022;17(1):333.
21. Guo H, Ding D, Wang L, et al. Metformin attenuates osteoclast-mediated abnormal subchondral bone remodeling and alleviates osteoarthritis via AMPK/NF-kappaB/ERK signaling pathway. PLoS One 2021;16(12):e0261127.
22. Alimoradi N, Tahami M, Firouzabadi N, et al. Metformin attenuates symptoms of osteoarthritis: role of genetic diversity of Bcl2 and CXCL16 in OA. Arthritis Res Ther 2023;25(1):35.
23. Mohammed MM, Al-Shamma KJ, Jasim NA. Evaluation of the anti-inflammatory effect of metformin as adjuvant therapy to NSAID (meloxicam) in patients with knee osteoarthritis. Int J Sci Nat 2014;5(2):277–282.

24. Mohammed MM, Al-Shamma KJ, Jassim NA. Evaluation of the clinical use of metformin or pioglitazone in combination with meloxicam in patients with knee osteoarthritis; using knee injury and osteoarthritis outcome score. Iraqi J Pharm Sci 2014;23(2):13–23.
25. Fraenkel L, Suter L, Cunningham CE, Hawker G. Understanding preferences for disease-modifying drugs in osteoarthritis. Arthritis Care Res (Hoboken). 2014 Aug; 66(8):1186-92.
26. Ezzat AM, Brussoni M, Whittaker JL, et al. A qualitative investigation of the attitudes and beliefs about physical activity and post-traumatic osteoarthritis in young adults 3-10 years after an intra-articular knee injury. Phys Ther Sport 2018;32:98–108.
27. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis 1957;16(4):494–502.
28. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005;15(9):1277–1288.
29. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008;62(1):107–115.
30. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007;19(6):349–357.
31. McGrady ME, Holbein CE, Smith AW, et al. An independent evaluation of the accuracy and usability of electronic adherence monitoring devices. Ann Intern Med 2018;169(6):419–422.
32. West C, Zimmerman DH. Doing gender. Gender & Society 1987;1(2): 125–151.
33. Bannon I, Correia M, eds. The other half of gender: men's issues in development. World Bank Publications; 2006.
34. Connell RW, Messerschmidt JW. Hegemonic masculinity: rethinking the concept. Gender & Society 2005;19(6):829–859.
35. Kimmel MS. Guyland: The perilous world where boys become men. 1st ed. Harper; 2008.
36. Pascoe CJ. Dude, you're a fag: masculinity and sexuality in high school. University of California Press; 2007.
37. Piussi R, Magnusson C, Andersson S, et al. Some, but not all, patients experience full symptom resolution and a positive rehabilitation process after ACL reconstruction: an interview study. Knee Surg Sports Traumatol Arthrosc 2023;31(7):2927–2935.
38. Piussi R, Simonson R, Kjellander M, et al. When context creates uncertainty: experiences of patients who choose rehabilitation as a treatment after an ACL injury. BMJ Open Sport Exerc Med 2023;9 (1):e001501.
39. Karlström J, Wiklund M, Tengman E. Disrupted knee - disrupted me: a strenuous process of regaining balance in the aftermath of an anterior cruciate ligament injury. BMC Musculoskelet Disord 2022;23 (1):290.
40. Truong LK, Mosewich AD, Miciak M, et al. Social support and therapeutic relationships intertwine to influence exercise behavior in people with sport-related knee injuries. Physiother Theory Pract 2025;41(1): 139–152.
41. Truong LK, Mosewich AD, Miciak M, et al. "I feel I'm leading the charge." Experiences of a virtual physiotherapist-guided knee health program for persons at-risk of osteoarthritis after a sport-related knee injury. Osteoarthr Cartil Open 2023;5(1):100333.
42. Selten EM, Vriezekolk JE, Geenen R, et al. Reasons for treatment choices in knee and hip osteoarthritis: a qualitative study. Arthritis Care Res (Hoboken) 2016;68(9):1260–1267.
43. Kamsan SS, Singh DKA, Tan MP, et al. The knowledge and self-management educational needs of older adults with knee osteoarthritis: a qualitative study. PLoS One 2020;15(3):e0230318.
44. Townsend A, Backman CL, Adam P, et al. A qualitative interview study: patient accounts of medication use in early rheumatoid arthritis from symptom onset to early postdiagnosis. BMJ Open 2013;3(2): e002164.
45. Creswell JW, Poth CN. Qualitative inquiry and research design: choosing among five traditions. Sage Publications, Inc; 1998.
46. Valtis YK, Stevenson KE, Murphy EM, et al. Race and ethnicity and the utilization of security responses in a hospital setting. J Gen Intern Med 2023;38(1):30–35.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 77, No. 7, July 2025, pp 892–899
DOI 10.1002/acr.25511

Comparing Community-Level Social Determinants of Health With Patient Race in Total Hip Arthroplasty Outcomes

Bella Mehta,[1] Yi Yiyuan,[2] Diyu Pearce-Fisher,[3] Kaylee Ho,[2] Susan M. Goodman,[1] Michael L. Parks,[1] Fei Wang,[2] Mark A. Fontana,[1] Said Ibrahim,[4] Peter Cram,[5] and Rich Caruana[6]

Objective. Social determinants of health (SDOH), including race, have a key role in total hip arthroplasty (THA) disparities. We compared the collective influence of community-level SDOH to the influence of individual factors such as race, on THA outcomes.

Methods. This retrospective cohort study of the Pennsylvania Health Care Cost Containment Council Database (2012–2018) included 105,336 patients undergoing unilateral primary elective THA. We extracted "community" factors from the US census by geocoding patient zip codes, including walkability index, household income, foreign-born individuals, English proficiency, computer and internet access, unpaid family workers, those lacking health insurances, and education. We trained an explainable boosting machine, a modern form of generalized additive models, to predict 90-day readmission, 90-day mortality, one-year revision, and length of stay (LOS). Mean absolute scores were aggregated to measure variable importance (ie, variables that contributed most to the prediction).

Results. The rates of readmission, revision, and mortality were 8%, 1.5%, and 0.3%, respectively, with a median LOS of two days. Predictive performance measured by area under the receiver operating characteristic curve was 0.76 for mortality, 0.66 for readmission, and 0.57 for one-year revision. For LOS, the root mean squared error was 0.41 (R^2 = 0.2). The top three predictors of mortality were community, discharge location, and age; for readmission, they were discharge location, age, and comorbidities; for revision, they were community, discharge location, and comorbidities; and for LOS, they were discharge location, community, and comorbidities.

Conclusion. Community-level SDOH were significantly more important than individual race in contributing to the prediction of THA outcomes, especially for 90-day mortality.

INTRODUCTION

Elective total hip arthroplasty (THA) is a prototypical "preference-sensitive" procedure. THA utilization rates have grown rapidly worldwide and are expected to reach 572,000 annually in the United States by the year 2030.[1–3] Although use of THA has increased and outcomes have improved over the past decade, these benefits are not shared equally throughout the population. The disparity is largely mediated by social determinants of health (SDOH).[4–8] SDOH, such as race and ethnicity, have been extensively researched in THA use and outcomes. Multiple studies have shown that African American race is associated with decreased use and worse outcomes after THA compared to White race.[9–11] However, there is also evidence that in addition to race and ethnicity, community-level factors play an important role in a wide range of health care outcomes.[12–15] Although there are some studies on community-level SDOH describing worse outcomes in THA among patients from low socioeconomic status (SES) communities, many other community-level SDOH are not comprehensively studied in THA outcomes.[16–19]

Additionally, given the complex interplay of SDOH, it is difficult to study how both individual features and groups of features

Presented in part at the 2023 Osteoarthritis Research Society International World Congress and the 2023 EULAR European Congress of Rheumatology.

Supported by the Anna-Maria and Stephen Kellen Foundation Total Knee Improvement Program; the National Institute of Arthritis and Musculoskeletal and Skin Diseases, NIH (award 1K23-AR-082991-01A1); and Schmidt Sciences, LLC.

[1]Bella Mehta, MBBS, MS, MD, Susan M. Goodman, MD, Michael L. Parks, MD, Mark A. Fontana, PhD: Hospital for Special Surgery and Weill Cornell Medicine, New York, New York; [2]Yi Yiyuan, MS, Kaylee Ho, MS, Fei Wang, PhD: Weill Cornell Medicine, New York, New York; [3]Diyu Pearce-Fisher, BS: Hospital for Special Surgery, New York, New York; [4]Said Ibrahim, MD, MPH, MBA: Northwell Health, New Hyde Park, New York; [5]Peter Cram, MD, MBA: The University of Texas Medical Branch at Galveston; [6]Rich Caruana, PhD: Microsoft Research, Redmond, Washington.

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.25511).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25511.

Address correspondence via email to Bella Mehta, MBBS, MS, MD, at drbellamehta@gmail.com.

Submitted for publication May 30, 2024; accepted in revised form January 31, 2025.

SIGNIFICANCE & INNOVATIONS

- This study illustrates that prevalent community factors (eg, walkability index, household income, foreign-born individuals, English proficiency, computer and internet access, unpaid family workers, those lacking health insurance, and education, etc) are more important than race in contributing to prediction of all patient total hip arthroplasty (THA) outcomes.
- Our study enacts the use of explainable boosting machines, which are a modern form of generalized additive models to measure groups of features together by aggregating the importance of attributes and can account for interactions in community-level variables.
- Given the importance of community variables in predicting patient THA outcomes, this study underscores the urgency in addressing community-level variables when enacting policy changes, resource allocation, and social program implementation.

may all contribute to THA outcomes using traditional methods of statistics. Complex machine learning models can do so but sometimes at the expense of interpretability. Explainable boosting machine (EBM), a form of generalized additive models (GAMs), allows for more complex modeling of these intricate relationships but is easier to interpret than other machine learning models. We sought to compare the collective influence of community-level SDOH with the influence of individual factors, such as race, on THA outcomes using EBM models.

PATIENTS AND METHODS

Study cohort. We performed a retrospective analysis on the Pennsylvania Health Care Cost Containment Council (PHC4) database (2012–2018). The PHC4 data set includes demographics from all patients discharged from 170 nongovernmental acute care hospitals in the state of Pennsylvania known as Pennsylvania Assigned Facilities, excluding Veterans Administration and military hospitals. The state-run agency collects more than 4.5 million records each year, including deidentified patient demographic information, diagnostic and procedural codes, hospital information, and financial data.

We used the PHC4 data set to identify 140,092 patients who underwent elective primary THA using the *International Classification of Diseases, Ninth Revision, Clinical Modification* (ICD-9-CM) procedure code 81.51 for primary THA from 2012 to September 2015 and the ICD-10 procedure codes 0SR90xx or 0SRB0xx thereafter. These are validated codes from the American Joint Replacement Registry, with a sensitivity of 99%, a specificity of 91%, and a positive predictive value of 91%.[20,21] The study cohort and methodology have previously been described in detail.[22]

We excluded a total of 34,756 patients, including those whose diagnostic codes suggested inflammatory arthritis (rheumatoid arthritis, systemic lupus erythematosus, psoriatic arthritis, spondyloarthropathy), pathologic fracture, avascular necrosis, and metastatic and bone cancer. Other exclusion criteria included the following: nonelective admissions or interhospital transfers before THA, two or more past THAs (likely administrative data set error), bilateral THA during the index hospitalization, and revision during the index hospitalization. We excluded patients residing outside Pennsylvania as well as those missing key variables such as age, sex, residential zip or postal code, and community-level income.

We extracted information on important covariates, including patient- and facility-level variables, from the PHC4 data set. This included race, but complete data on ethnicity were not available. Medical comorbidities were identified using the Quan adaptation of the Elixhauser comorbidity index, as previously described.[23] The research project was determined by the Weill Cornell Medicine Institutional Review Board to meet exemption requirements at the Hospital for Special Surgery because it includes only research involving the collection or study of existing data, documents, records, pathologic specimens, or diagnostic specimens, which are publicly available and/or recorded by the investigator in such a manner that patients cannot be identified directly or through identifiers linked to the patients.

Extraction of community factors and individual-level factors. We geocoded the five-digit zip codes for each patient. For factors that were known to be relevant to or to correlate with THA outcomes, we extracted the corresponding census tract variables from the American Community Survey. These included the following: (1) The percentage of householders living alone in each zip code were used to represent social support, which is known to be important to THA recovery.[24–26] (2) The percentage of foreign-born individuals and (3) the percentage of speaking languages other than English were used to represent the immigrant population and acculturation, which were shown in prior studies to be important in joint replacement and other surgical outcomes.[27–29] (4) The percentage with computer access and (5) the percentage with internet access represented access to information as well as digital and health literacy.[30–33] (6) The median household income, (7) the percentage of unpaid family workers, (8) the percentage without health insurance and not in the labor force, and (9) education (percentage above high school level and percentage above college level) were identified to represent SES and SDOH and are known to have associations with THA outcomes.[5–7,16,34–36] Lastly, we included (10) the National Walkability Index because walkability has a considerable effect on osteoarthritis and postarthroplasty outcomes.[37–39] The US Environmental Protection Agency National Walkability Index tool was used to approximate regional walking ease.[40–42] It is ranked between 1 and 20, with 1 being the least probability

toward walking and 20 being the most.[43] All of these variables are subsequently combined as "community" factors as detailed in the Statistical analysis section. We included the following individual-level factors: demographics (age, sex, and race, categorized as Black, White, and "other" which included all races except Black and White), discharge location (home vs nonhome), and patient's comorbidity burden as measured by the Elixhauser comorbidity index.

Statistical analysis. Patient- and community-level demographics were described for 105,336 patients undergoing THA using the values no. (%) for categorical variables and median (interquartile range [IQR]) for continuous variables. Differences in patient- and community-level demographics were compared using the chi-square test for categorical variables or the Wilcoxon rank sum test for continuous variables based on 90-day readmission, 90-day mortality, and one-year revision status (yes compared to no) (Supplementary Table S1). Length of stay (LOS) was transformed by the equation to normalize distribution: log (LOS + 0.5). We developed EBM models for each of the outcomes, including demographics (age, sex, race) and discharge location, community-level predictors, and patient's comorbidity burden as measured by the Elixhauser score.

We used a modern form of GAMs called EBMs so that we could measure groups of features together by aggregating the importance of attributes.[44] GAMs are similar to highly interpretable models, but they have prediction accuracy comparable to tree-based machine learning models, such as random forests.[45] Many community factors are correlated, and the round-robin approach of the EBM model, along with a low learning rate, helps distribute predictive power among these features. This minimizes collinearity effects and prevents multiple counting of evidence, ensuring that the explanatory power of feature groups is accurately reflected in their weights and importance to predictions.[46] Because EBMs are additive models, it is easy to measure the contribution of each feature or group of features to each prediction.[47–49] We used EBMs because they have high accuracy with no loss of fidelity or approximations involved, are fully interpretable, and allow the importance of features to be measured both individually and aggregated as groups.[45,50,51]

On 70% of data stratified by each outcome, we trained an EBM model to predict risk for 90-day readmission, 90-day mortality, LOS, and one-year revision, controlling for demographics, comorbidities, and community-level variables.[47,52] The area under the receiver operating characteristic curve (AUROC) on test data (30% of the sample) was reported for EBM models predicting 90-day readmission, 90-day mortality, and one-year revision for classification; root mean squared error (RMSE) was reported for the regression model predicting LOS. Results for 90-day readmission, 90-day mortality, and one-year revision were validated with five-fold cross-validation stratified by outcome. Results for LOS were validated with five-fold cross-validation.

Feature importance was measured using mean absolute scores of the trained model and used to predict the full data set. This was done by computing the average contribution the feature makes to predicted risk averaged over the patient population. Each feature in the group was passed through the learned shape function from the EBM model. We then used the density function from each feature to generate an absolute value score by row or log odds if it was a classification problem; we then averaged the score to get the feature importance mean absolute score. For the community factors variable, feature importance by group was measured using mean absolute score of the model. Each feature in the group was passed through the learned shape function from the EBM model. We describe the combined weighted-average importance scores as community factors. In a secondary analysis, we used mean imputation to compute missing values for community factors and reran all the models. All analyses were done in python jupyternotebook.[53]

RESULTS

Cohort characteristics. Of the 105,336 patients undergoing THA in the PHC4 who met our inclusion criteria, 8,422 (8%) had readmission within 90 days, 309 (0.3%) died within 90 days after discharge, and 1,617 (1.5%) had revision within one year. The median LOS was 2 days (IQR 1–3). The median age at admission was 65 (IQR 58–73) years, and 46.0% were male. Although 90.3% were White, 6.6% were African American, and 3.0% were other races. Additionally, 78.1% were discharged home or home with health care, and 21.1% were discharged to post–acute care facilities (Table 1).

Patient characteristics for each outcome. Patients who were readmitted within 90 days (n = 8,422) were more likely to be older (median 69.0, IQR 60.0–77.0), female (55.2%), and African American (8.2%) and have a nonhome discharge destination (41.2%) (Supplementary Table S1). Readmitted patients were more likely to live in zip codes with a higher percentage of householders living alone, a lower median household income, and a higher percentage of people outside the labor force with no health insurance (Supplementary Table S2). Surprisingly, readmitted patients were more likely to live in zip codes with a higher National Walkability Index, which may reflect urban populations.

Patients who experienced 90-day mortality (n = 309) were more likely to be older (median 77.0, IQR 67.0–84.0) and have nonhome discharge after THA (52.8%) (Supplementary Table S2). Interestingly, they were more likely to live in zip codes with a lower percentage of residents who speak a language other than English and a lower median household income.

Finally, patients requiring revision within one year (n = 1,617) were more likely to be female (57.6%) and have a nonhome discharge (28.6%). They were more likely to live in zip codes with a higher percentage of householders living alone, a lower median

Table 1. Characteristics of cohort*

Variables	N = 105,336
Age, median (IQR), y	65.0 (58.0–73.0)
Sex, n (%)	
Female	56,977 (54.1)
Race, n (%)	
White	95,158 (90.4)
African American	6,954 (6.61)
Other	3,142 (2.99)
Missing	82 (0.08)
Discharge location, n (%)	
Discharged home/HH	82,262 (78.1)
Post–acute care	22,200 (21.1)
Other	651 (0.62)
Transfer to other acute care hospital	193 (0.18)
Died in hospital	26 (0.02)
Missing	4
Elixhauser comorbidity index, n (%)	
0	17,666 (16.8)
1–4	81,795 (77.7)
≥5	5,875 (5.58)
Community-level variables, median (IQR)	
Householder living alone, %	28.1 (24.0–32.8)
Foreign born, %	4.30 (1.71–8.37)
Missing, no.	9,837
Speaking language other than English, %	6.99 (3.64–12.0)
With computer access, %	85.4 (81.6–89.0)
Missing, no.	9,815
With internet access, %	77.4 (72.4–83.2)
Missing, no.	9,815
Median household income, US$	57,336 (48,078–71,783)
With unpaid family workers, %	0.12 (0.05–0.22)
Missing, no.	20,040
Not in labor force without insurance, %	8.71 (5.93–11.6)
Missing, no.	83
Above high school, %	91.3 (88.2–93.9)
Missing, no.	9,815
Above college, %	26.2 (19.6–38.1)
Missing, no.	9,815
National Walkability Index	8.76 (7.28–11.8)

* Categorical variables are reported as n (%), and continuous variables are reported as median (IQR). HH, home with health care; IQR, interquartile range.

household income, and a higher percentage of people outside the labor force with no health insurance.

Predictive performance and variable importance. Our 90-day readmission classification model had an AUROC of 0.66 on the test sample. Community factors were more important than race in predicting 90-day readmission, and the five-fold cross-validated mean absolute score in log odds units for community factors (0.10 [SD 0.005]) was considerably higher than that of race (0.015 [SD 0.002]). The top three most important factors predicting 90-day readmission were discharge location, age, and Elixhauser comorbidity index (mean absolute scores 0.24, 0.15, and 0.13, respectively) (Figure 1A).

Our 90-day mortality prediction model had an AUROC of 0.76. Community factors were more important than race in predicting 90-day mortality, and the cross-validated mean absolute score for community factors (0.28 [SD 0.01]) was considerably higher than that of race, which was the least important factor in the model (0.02 [SD 0.005]). The top three most important factors predicting 90-day mortality were community factors, discharge location, and age (mean absolute scores 0.31, 0.24, and 0.19, respectively) (Figure 1B).

Our one-year revision prediction model had an AUROC of 0.58. The top three most important factors predicting revision were community factors, discharge location, and Elixhauser comorbidity index (mean absolute score 0.03, 0.01, and 0.005, respectively) (Figure 1C). Race was again the least important factor in the model (mean absolute score 0.002 [SD 0.0002]).

Our LOS prediction model had a RMSE of 0.41, with a coefficient of determination R^2 of 0.2. The top three most important factors predicting LOS were discharge location, community factors, and Elixhauser comorbidity index (mean absolute scores 0.12, 0.056, and 0.030, respectively) (Figure 1D). Race was the least important factor in the model (mean absolute score 0.009 [SD 0.0002]). The predictive performance of all models did not change significantly when we imputed missing values for community factors.

DISCUSSION

In this study of 105,336 patients who underwent elective THA in the state of Pennsylvania in 2012 to 2018, importance of community factors—including household income, neighborhood walkability, living alone, and percentage uninsured—was greater than that of individual race in predicting post-THA 90-day readmission, 90-day mortality, LOS, and one-year revision, even when accounting for patient age, sex, and comorbidities. Community factors were more important than race, age, sex, and comorbidities in predicting 90-day mortality and one-year revision. In our models for 90-day readmission and LOS, the aggregate variable importance score for community factors was considerably more than that of race.

In our statistical analysis comparing cohort characteristics by THA outcome, we found results similar to previous studies. We confirmed that African American race was associated with increased readmission.[6,8,34,36,54] Our results also affirmed previous studies that found that nonhome discharge destination was associated with increased readmission rates, mortality rates, and LOS.[34,54,55] Our claim that nonrace SDOH may be more important in predicting THA outcomes than race is also suggested by other studies. Others showed that socioeconomic disadvantage may mediate the association of race with LOS by 37%, and when controlling for this disadvantaged status, the association of race with 90-day readmission was no longer significant. Still others showed that socioeconomic disadvantaged status but not race was significantly associated with LOS and 90-day complications

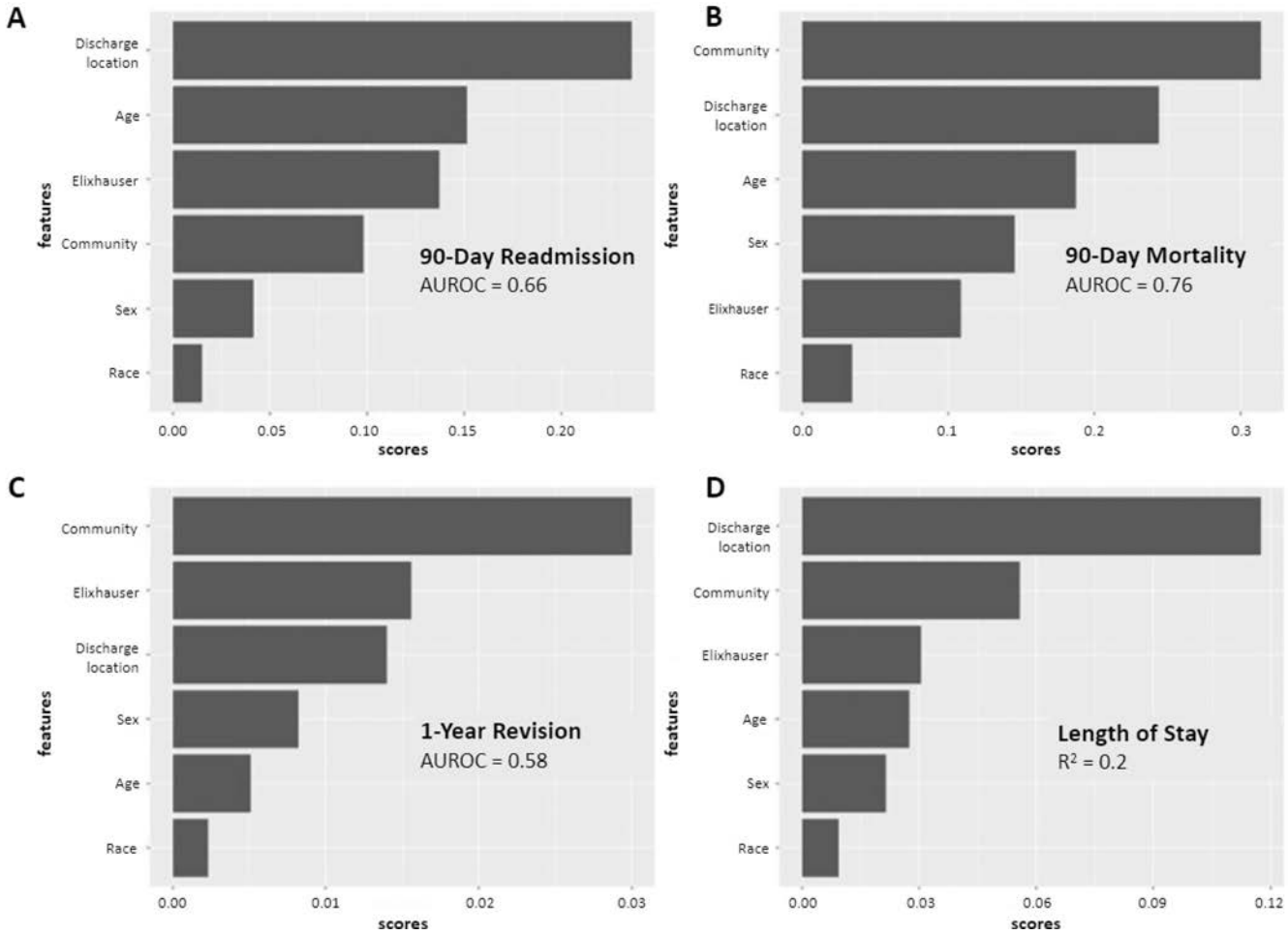

Figure 1. (A) Variable importance scores predicting 90-day readmission based on an EBM (AUROC = 0.6579), (B) 90-day mortality based on an EBM (AUROC = 0.7643), (C) revision based on an EBM (AUROC = 0.5775), and (D) length of stay based on regression (RMSE = 0.41 and R^2 = 0.20). "Community" is the aggregate importance score for the variables percentage of householders living alone, percentage who are foreign born, percentage who speak a language other than English, percentage with computer access, percentage with internet access, National Walkability Index, percentage not in the labor force without insurance, percentage above a high school level, percentage above a college level, percentage with unpaid family workers, and median household income, which were extracted from the American Community Survey or calculated using the National Walkability Index software based on individual patient zip code. AUROC, area under receiver operator characteristic curve; EBM, explainable boosting machine; RMSE, root mean squared error.

in multivariable models.[56,57] These help us affirm that the our model shows similar results in measures previously studied.

Studies have shown that at the patient level, increased social support is associated with better post-THA patient-reported outcome, lower revision rates, and shorter LOS.[24–26] Others have shown that lower income or SES and Medicaid or uninsured status at the patient level are associated with increased post-THA readmission and LOS.[5–7,35,36] Our results suggest that social support, access to care, and income are important at the community level as well.

Although racial disparities on THA use and outcomes are well established—likely also reflecting that race is associated with many other SDOH, such as SES, education, and access to care—it is hard to understand the complexity of how SDOH collectively influence THA outcomes. Along the same lines, prior work has demonstrated that policies increasing access to care potentially lessen racial disparities in THA outcomes and use.[55,58–60] However, further county-level differences in THA outcomes still exist even after accounting for most patient-level characteristics, such as age, sex, body mass index (BMI), Elixhauser comorbidity index, marital status, educational level, and disposable income.[19] Our study expands on this knowledge and demonstrates that community factors were more important than individual race in predicting THA outcomes. Thus, our results indicate that more so than race, unaddressed community factors are often unnoticed and may deserve greater attention to understand and improve THA outcomes.

Our retrospective study has several limitations. First, the PHC4 database lacks granular information about important patient-level variables, such as BMI, ethnicity, and surgical facility.

However, we did have comorbidities included in the models. Although we did not have access to patient-level wealth or income, there is a lot of research showing that neighborhood SES may be a reasonable proxy. Our data set also lacks data on care continuity and patient-reported outcomes, which may help shed insight in our future work. Additionally, our patient population is limited to primary elective THA in Pennsylvania, so our results may not be generalizable to other regions of the country, but we do have a sizable sample in this study. Although our results do not account for clustering correlation by zip code, we built classifiers using EBMs to improve prediction performance using zip code–level SDOH along with other clinical predictors.[61,62] Next, our models have a limited predictive value, with our EBMs for mortality, readmission, and revision having AUROCs of 0.76, 0.66, and 0.58, respectively, and our LOS regression model having an RMSE of 0.41. However, this caliber of performance of our EBMs is comparable to other similar claims-based outcome analyses using various other machine learning techniques as well as logistic regression.[63–67] Even with limited AUROC, GAM is still thought to be potentially useful in developing heuristics for population-level policy.[68] Finally, EBM models are limited by the possibility of overfitting, especially if a vast number of terms are considered. To address this limitation and minimize the possibility of overfitting, we did not use interactions in the final models because they were complex and did not improve performance.

Despite these limitations, our study strengths include a large sample size of 105,336 patients, for whom we had access to detailed patient-level data on age, race, sex, and comorbidity index, as well as a wide variety of census tract community-level variables. Additionally, we use EBMs, which offer higher interpretability, higher accuracy, and lower likelihood of overfitting than other machine learning methods. Feature importance using the EBMs does account for collinearity among variables and can accurately assess the importance of individual variables and groups of variables in the model while accounting for magnitude and direction of the effect sizes. Further, unlike other methods, such as Shapley Additive Explanations (SHAP), the interpretability is exact (ie, there is no loss of fidelity or approximations involved).

A key strength of our study lies in its ability to move beyond the traditional focus of race as a dominant SDOH and highlights the other SDOH offering a more comprehensive perspective of factors influencing outcomes in THA. Although race undoubtedly plays a crucial role, an exclusive focus on this single patient-level SDOH can overlook other significant contributors to outcomes, recognizing that these are shaped by multiple interrelated factors beyond race. By aggregating several community-level SDOH (eg, percentage of householders living alone, percentage of foreign-born individuals, etc), the study highlights that other contextual factors also exert significant influence on health. These factors often reflect the broader social and environmental conditions in which patients live, conditions that can compound or mitigate the effects of individual characteristics such as race.

In summary, in this large-scale study of 105,336 patients undergoing primary THA across 170 Pennsylvania hospitals, we used novel machine learning models to show that community factors such as household income, neighborhood walkability, living alone, and percentage uninsured were more important than individual race in predicting 90-day readmission, 90-day mortality, one-year revision, and LOS after THA. Despite the emphasis on race in THA disparities research, our results indicate that additional focus is needed on the community factors that are associated with poor THA outcomes. These should be included in conversations between physicians and patients while planning treatment and discharge. For example, surgical planning would include a multidisciplinary team involving the surgeon, primary care physicians, physical therapists, occupational therapists, and social work. Additionally, the patient's zip code alone can give the team information that can gauge the overall community-level SDOH, access to health care and internet, and community social support. For patients in zip codes screened with community factors associated with poorer outcomes, the surgeon could center discharge planning discussions to ensure occupational therapy equipment and physical therapy arrangements for better home health services. Social work could provide additional resources such as free or low-cost internet, transportation, or other insurance options to these patients that might not be necessary for those patients from zip codes with community factors associated with better outcomes. These are future implementation studies that need to be done.

Community factors are aspects of systemic inequalities that are difficult to address, but our study suggests that they may need to be considered when enacting policies and resource allocation. Future research is necessary to determine if it is possible to mitigate some of these community-level inequalities.

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 Mehta 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. Wolford ML, Palso K, Bercovitz A. Hospitalization for total hip replacement among inpatients aged 45 and over: United States, 2000-2010. NCHS Data Brief 2015;(186):1–8.
2. Fingar KR, Stocks C, Weiss AJ, et al. Most frequent operating room procedures performed in U.S. hospitals, 2003-2012. Accessed June 24, 2022. https://www.hcup-us.ahrq.gov/reports/statbriefs/sb186-Operating-Room-Procedures-United-States-2012.jsp

3. Kurtz S, Ong K, Lau E, et al. Projections of primary and revision hip and knee arthroplasty in the United States from 2005 to 2030. J Bone Joint Surg Am 2007;89(4):780–785.
4. Johnson MA, Sloan M, Lopez VS, et al. Racial disparities in perioperative complications following primary total hip arthroplasty. J Orthop 2020;21:155–160.
5. Weiner JA, Adhia AH, Feinglass JM, et al. Disparities in hip arthroplasty outcomes: results of a statewide hospital registry from 2016 to 2018. J Arthroplasty 2020;35(7):1776–1783.e1.
6. Maldonado-Rodriguez N, Ekhtiari S, Khan MM, et al. Emergency department presentation after total hip and knee arthroplasty: a systematic review. J Arthroplasty 2020;35(10):3038–3045.e1.
7. Grits D, Emara AK, Klika AK, et al. Neighborhood socioeconomic disadvantage associated with increased healthcare utilization after total hip arthroplasty. J Arthroplasty 2022;37(10):1980–1986.e2.
8. Dharmasukrit C, Chan SYS, Applegate RL II, et al. Frailty, race/ethnicity, functional status, and adverse outcomes after total hip/knee arthroplasty: a moderation analysis. J Arthroplasty 2021;36(6):1895–1903.
9. Jha AK, Fisher ES, Li Z, et al. Racial trends in the use of major procedures among the elderly. N Engl J Med 2005;353(7):683–691.
10. Lawrence RC, Helmick CG, Arnett FC, et al. Estimates of the prevalence of arthritis and selected musculoskeletal disorders in the United States. Arthritis Rheum 1998;41(5):778–799.
11. Jordan JM, Linder GF, Renner JB, et al. The impact of arthritis in rural populations. Arthritis Care Res 1995;8(4):242–250.
12. Nelson A. Unequal treatment: confronting racial and ethnic disparities in health care. J Natl Med Assoc 2002;94(8):666–668.
13. Ahn KO, Shin SD, Hwang SS, et al. Association between deprivation status at community level and outcomes from out-of-hospital cardiac arrest: a nationwide observational study. Resuscitation 2011;82(3):270–276.
14. Miller R, Akateh C, Thompson N, et al. County socioeconomic characteristics and pediatric renal transplantation outcomes. Pediatr Nephrol 2018;33(7):1227–1234.
15. Biswas S, Andrianopoulos N, Duffy SJ, et al. Impact of socioeconomic status on clinical outcomes in patients with ST-segment-elevation myocardial infarction. Circ Cardiovasc Qual Outcomes 2019;12(1):e004979.
16. Brennan-Olsen S, Vogrin S, Holloway KL, et al. Geographic region, socioeconomic position and the utilisation of primary total joint replacement for hip or knee osteoarthritis across western Victoria: a cross-sectional multilevel study of the Australian Orthopaedic Association National Joint Replacement Registry. Arch Osteoporos 2017;12(1):97.
17. Brennan SL, Lane SE, Lorimer M, et al. Associations between socioeconomic status and primary total knee joint replacements performed for osteoarthritis across Australia 2003-10: data from the Australian Orthopaedic Association National Joint Replacement Registry. BMC Musculoskelet Disord 2014;15(1):356.
18. Agabiti N, Picciotto S, Cesaroni G, et al; Italian Study Group on Inequalities in Health Care. The influence of socioeconomic status on utilization and outcomes of elective total hip replacement: a multicity population-based longitudinal study. Int J Qual Health Care 2007;19(1):37–44.
19. Oldsberg L, Garellick G, Osika Friberg I, et al. Geographical variations in patient-reported outcomes after total hip arthroplasty between 2008 - 2012. BMC Health Serv Res 2019;19(1):343.
20. Daneshvar P, Forster AJ, Dervin GF. Accuracy of administrative coding in identifying hip and knee primary replacements and revisions. J Eval Clin Pract 2012;18(3):555–559.
21. Cahue SR, Etkin CD, Stryker LS, et al. Procedure coding in the American Joint Replacement Registry. Arthroplast Today 2019;5(2):251–255.
22. Singh JA, Kwoh CK, Boudreau RM, et al. Hospital volume and surgical outcomes after elective hip/knee arthroplasty: a risk-adjusted analysis of a large regional database. Arthritis Rheum 2011;63(8):2531–2539.
23. Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care 2005;43(11):1130–1139.
24. Simões JL, Soares S, Sa-Couto P, et al. The influence of presurgical factors on the rehabilitation outcome of patients following hip arthroplasty. Rehabil Nurs 2019;44(4):189–202.
25. Young NL, Cheah D, Waddell JP, et al. Patient characteristics that affect the outcome of total hip arthroplasty: a review. Can J Surg 1998;41(3):188–195.
26. Lall AC, Schwarzman GR, Battaglia MR, et al. Effect of marital status on patient-reported outcomes following total hip arthroplasty: a matched analysis with minimum 2-year follow-up. Hip Int 2021;31(3):362–368.
27. Lum TY, Vanderaa JP. Health disparities among immigrant and nonimmigrant elders: the association of acculturation and education. J Immigr Minor Health 2010;12(5):743–753.
28. Gomez SL, France AM, Lee MM. Socioeconomic status, immigration/acculturation, and ethnic variations in breast conserving surgery, San Francisco Bay area. Ethn Dis 2004;14(1):134–140.
29. McEnhill ME, Brennan JL, Winnicki E, et al. Effect of immigration status on outcomes in pediatric kidney transplant recipients. Am J Transplant 2016;16(6):1827–1833.
30. Oshima SM, Tait SD, Thomas SM, et al. Association of smartphone ownership and internet use with markers of health literacy and access: cross-sectional survey study of perspectives from Project PLACE (Population Level Approaches to Cancer Elimination). J Med Internet Res 2021;23(6):e24947.
31. Bailey SC, O'Conor R, Bojarski EA, et al. Literacy disparities in patient access and health-related use of Internet and mobile technologies. Health Expect 2015;18(6):3079–3087.
32. Sarkar U, Karter AJ, Liu JY, et al. The literacy divide: health literacy and the use of an internet-based patient portal in an integrated health system-results from the diabetes study of northern California (DISTANCE). J Health Commun 2010;15(suppl 2):183–196.
33. McCleary-Jones V, Scheideman-Miller C, Rev Dorn JA Jr, et al. Health information technology use and health literacy among community-dwelling African Americans. ABNF J 2013;24(1):10–16.
34. Plate JF, Ryan SP, Bergen MA, et al. Patient risk profile for unplanned 90-day emergency department visits differs between total hip and total knee arthroplasty. Orthopedics 2020;43(5):295–302.
35. Mahomed NN, Barrett JA, Katz JN, et al. Rates and outcomes of primary and revision total hip replacement in the United States Medicare population. J Bone Joint Surg Am 2003;85(1):27–32.
36. Oronce CIA, Shao H, Shi L. Disparities in 30-day readmissions after total hip arthroplasty. Med Care 2015;53(11):924–930.
37. Martin KR, Shreffler J, Schoster B, et al. Associations of perceived neighborhood environment on health status outcomes in persons with arthritis. Arthritis Care Res (Hoboken) 2010;62(11):1602–1611.
38. Mielenz TJ, Callahan LF, Edwards MC. Average vs item response theory scores: an illustration using neighbourhood measures in relation to physical activity in adults with arthritis. Public Health 2017;142:15–21.
39. Gebauer S, Schootman M, Xian H, et al. Neighborhood built and social environment and meeting physical activity recommendations among mid to older adults with joint pain. Prev Med Rep 2020;18:101063.

40. US Census Bureau. Glossary. Census.gov. Accessed June 28, 2022. https://www.census.gov/programs-surveys/geography/about/glossary.html
41. Ewing R, Cervero R. Travel and the built environment. J Am Plann Assoc 2010;76(3):265–294.
42. Mooney SJ, Hurvitz PM, Moudon AV, et al. Residential neighborhood features associated with objectively measured walking near home: revisiting walkability using the Automatic Context Measurement Tool (ACMT). Health Place 2020;63:102332.
43. US Environmental Protection Agency. National walkability index user guide and methodology. May 17, 2021. Accessed June 28, 2022. https://www.epa.gov/smartgrowth/national-walkability-index-user-guide-and-methodology
44. Sasieni P. Generalized additive models. T. J. Hastie and R. J. Tibshirani, Chapman and Hall, London, 1990. No. of Pages: xv + 335. Price: £25. ISBN: 0-412-34390-8. Stat Med 1992;11(7):981–982.
45. Kuk M, Bobek S, Nalepa GJ. Comparing explanations from glass-box and black-box machine-learning models. In: Groen D, de Mulatier C, Paszynski M, et al, eds. Computational Science – ICCS 2022. Springer International Publishing; 2022:668–675. *Lecture Notes in Computer Science*; vol 13352. https://doi.org/10.1007/978-3-031-08757-8_55
46. Lou Y, Caruana R, Gehrke J. Intelligible models for classification and regression. In: *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Association for Computing Machinery; 2012:150–158.
47. Nori H, Jenkins S, Koch P, et al. InterpretML: a unified framework for machine learning interpretability. arXiv Preprint posted online September 19, 2019. https://doi.org/10.48550/ARXIV.1909.09223
48. Lundberg S, Lee SI. A unified approach to interpreting model predictions. arXiv Preprint posted online May 22, 2017. https://doi.org/10.48550/ARXIV.1705.07874
49. Lou Y, Caruana R, Gehrke J, et al. Accurate intelligible models with pairwise interactions. In: *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Association for Computing Machinery; 2013:623–631.
50. Agarwal R, Melnick L, Frosst N, et al. Neural additive models: interpretable machine learning with neural nets. arXiv Preprint posted online April 29, 2020. https://doi.org/10.48550/ARXIV.2004.13912
51. Alahmadi R, Almujibah H, Alotaibi S, et al. Explainable boosting machine: a contemporary glass-box model to analyze work zone-related road traffic crashes. Safety (Basel) 2023;9(4):83.
52. Caruana R, Lou Y, Gehrke J, et al. Intelligible models for healthcare: predicting pneumonia risk and hospital 30-day readmission. In: *Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Association for Computing Machinery; 2015:1721–1730.
53. The Jupyter Notebook. Jupyter Notebook 6.5.4 documentation. Accessed April 21, 2023. https://jupyter-notebook.readthedocs.io/en/stable/
54. Paxton EW, Inacio MCS, Singh JA, et al. Are there modifiable risk factors for hospital readmission after total hip arthroplasty in a US healthcare system? Clin Orthop Relat Res 2015;473(11):3446–3455.
55. Mehta B, Singh JA, Ho K, et al. Race, discharge disposition, and readmissions after elective hip replacement: analysis of a large regional dataset. Health Equity 2019;3(1):628–636.
56. Chandrashekar AS, Hymel AM, Baker CE, et al. socioeconomic indices are associated with increased resource utilizations, but not 90-day complications following total hip and knee arthroplasty. J Arthroplasty 2025;40(2):294–300.e1.
57. Dubin JA, Bains SS, Hameed D, et al. The utility of the area deprivation index in assessing complications after total joint arthroplasty. JBJS Open Access 2024;9(2):e23.00115.
58. Mohamed NS, Remily EA, Wilkie WA, et al. Closing the socioeconomic gap in Massachusetts: trends in total hip arthroplasty from 2013 to 2015. Orthopedics 2021;44(2):e167–e172.
59. Vina ER, Kallan MJ, Collier A, et al. Race and rehabilitation destination after elective total hip arthroplasty: analysis of a large regional data set. Geriatr Orthop Surg Rehabil 2017;8(4):192–201.
60. Kurtz SM, Lau E, Ong KL, et al. Universal health insurance coverage in Massachusetts did not change the trajectory of arthroplasty use or costs. Clin Orthop Relat Res 2016;474(5):1090–1098.
61. Hubbard AE, Ahern J, Fleischer NL, et al. To GEE or not to GEE: comparing population average and mixed models for estimating the associations between neighborhood risk factors and health. Epidemiology 2010;21(4):467–474.
62. Bouwmeester W, Twisk JW, Kappen TH, et al. Prediction models for clustered data: comparison of a random intercept and standard regression model. BMC Med Res Methodol 2013;13(1):19.
63. Baig MM, Hua N, Zhang E, et al. A machine learning model for predicting risk of hospital readmission within 30 days of discharge: validated with LACE index and patient at risk of hospital readmission (PARR) model. Med Biol Eng Comput 2020;58(7):1459–1466.
64. Hur J, Tang S, Gunaseelan V, et al. Predicting postoperative opioid use with machine learning and insurance claims in opioid-naïve patients. Am J Surg 2021;222(3):659–665.
65. Quinn TJ, Singh S, Lees KR, et al; VISTA Collaborators. Validating and comparing stroke prognosis scales. Neurology 2017;89(10):997–1002.
66. Chen KA, Joisa CU, Stitzenberg KB, et al. Development and validation of machine learning models to predict readmission after colorectal surgery. J Gastrointest Surg 2022;26(11):2342–2350.
67. Le ST, Liu VX, Kipnis P, et al. Comparison of electronic frailty metrics for prediction of adverse outcomes of abdominal surgery. JAMA Surg 2022;157(5):e220172.
68. Kunze KN, So MM, Padgett DE, et al. Machine learning on Medicare claims poorly predicts the individual risk of 30-day unplanned readmission after total joint arthroplasty, yet uncovers interesting population-level associations with annual procedure volumes. Clin Orthop Relat Res 2023;481(9):1745–1759.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 77, No. 7, July 2025, pp 900–905
DOI 10.1002/acr.25524

BRIEF REPORT

Association of Pain During Exercise With Exercise-Induced Hypoalgesia in People With Knee Osteoarthritis

Soyoung Lee,[1] Tuhina Neogi,[1] Benjamin M. Senderling,[1] S. Reza Jafarzadeh,[1] Mary Gheller,[1] Pirinka G. Tuttle,[2] Charmaine Demanuele,[2] Lars Viktrup,[3] Paul Wacnik,[2] and Deepak Kumar[1]

Objective. A paradoxical relationship between pain during exercise and the hypoalgesic effect of exercise has not been studied well in the knee osteoarthritis (OA) population. We sought to investigate the relation of pain evoked during exercise to exercise-induced hypoalgesia (EIH) and to determine if the efficiency of conditioned pain modulation (CPM), a proxy of the descending pain inhibitory system, mediates this relationship in people with knee OA.

Methods. We used baseline data from two clinical trials for people with symptomatic knee OA (n = 68). The maximum pain rating (0–10) during a series of knee exercises was defined as the outcome. EIH was assessed as an increase (ie, improvement) in the pressure pain threshold (PPT) after a bout of exercises. Efficient CPM was defined as an increase (ie, improvement) in PPT after a painful conditioning stimulus (forearm ischemia). We performed a causal mediation analysis to examine the association between pain during exercise and EIH as well as the mediating role of CPM efficiency on the relation of pain during exercise with EIH.

Results. People with knee OA who had at least a one-unit increase in pain with exercise were 43% more likely (odds ratio [OR] 1.43, 95% confidence interval [CI] 1.05–1.94) to experience subsequent EIH than those without pain increase. The efficiency of CPM did not mediate the relationship between pain during exercise and EIH (OR 1.00, 95% CI 0.96–1.04).

Conclusion. Our finding suggests that some amount of discomfort or pain during exercise may have beneficial analgesic effects; however, this is not likely via activation of the descending pain inhibitory system.

INTRODUCTION

Knee osteoarthritis (OA) is a leading cause of chronic pain and disability in middle-aged and older adults worldwide.[1] Although exercise is recommended as the first-line intervention for managing pain in people with knee OA,[2] fear of exercise-induced pain flares can interfere with exercise engagement and adherence.[3] In healthy individuals, painful exercise has been reported to induce a larger improvement in pain or pain sensitivity immediately after the bout of exercise, known as exercise-induced hypoalgesia (EIH).[4] In people with chronic pain, including those with knee OA, the EIH response can often be absent, with individuals experiencing no change or an increase in pain after a bout of exercise.[5] Stronger EIH may be related to greater improvements in pain and function after an exercise intervention in people with knee OA.[6] Whether pain evoked during exercise is related to EIH response in people with knee OA is unknown. Understanding the relevance of pain during exercise can help clinicians educate patients about this common clinical complaint and may facilitate improved exercise adherence.

With painful exercise, descending neural pathways can be activated and could inhibit nociceptive signaling via the "pain inhibits pain" paradigm.[7] This phenomenon is one potential mechanism underlying EIH; others are speculated to involve the

ClinicalTrials.gov identifiers: NCT03064139 and NCT04243096.

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

Supported by Boston University. NCT03064139 was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS), NIH (award K01-AR-069720), and by the Pfizer Inc ADVANCE grant. NCT04243096 was supported by funding from Pfizer Inc and Eli Lilly and Company. Dr Neogi's work was supported by the NIAMS, NIH (grants K24-AR-070892 and P30-AR-072571). Dr Jafarzadeh's work was supported by the NIAMS, NIH (grant P30-AR-072571).

[1]Soyoung Lee, PhD, Tuhina Neogi, MD, PhD, Benjamin M. Senderling, MS, S. Reza Jafarzadeh, PhD, Mary Gheller, MS, Deepak Kumar, PT, PhD: Boston University, Boston, Massachusetts; [2]Pirinka G. Tuttle, MS, Charmaine Demanuele, PhD, Paul Wacnik, PhD: Pfizer Inc, Cambridge, Massachusetts; [3]Lars Viktrup, MD, PhD: Eli Lilly and Company, Indianapolis, Indiana.

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.25524).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25524.

Address correspondence via email to Deepak Kumar, PT, PhD, at kumard@bu.edu.

Submitted for publication August 7, 2024; accepted in revised form March 6, 2025.

SIGNIFICANCE & INNOVATIONS

- Mild to moderate pain during a bout of knee-strengthening exercise was associated with greater likelihood of experiencing subsequent exercise-induced hypoalgesia (EIH) in people with knee osteoarthritis (OA).
- Efficiency of conditioned pain modulation did not mediate the association between pain increase during exercise and EIH.
- Some pain with exercise may be beneficial for people with knee OA.

endocannabinoid, serotonergic, opioid, immune, and autonomic systems.[5] Conditioned pain modulation (CPM) is an indirect measure commonly used for estimating the efficiency of the descending inhibitory system. Whereas some studies have reported an association between CPM and EIH response,[4,7] others have not.[8] Whether the efficiency of CPM mediates the relation between pain evoked during exercise and EIH in people with knee OA merits further study. Our objectives were (1) to examine the relation of pain evoked during exercise to the presence of EIH in people with knee OA and (2) to determine if the efficiency of CPM mediates the relationship between pain during exercise and EIH in people with knee OA.

PATIENTS AND METHODS

Participants. This study reports secondary analyses from baseline data from two clinical trials (ClinicalTrials.gov identifiers: NCT03064139 and NCT04243096). Participants for both trials were recruited from the general community using advertisements that included flyers posted in the community, online and social media advertising, and print advertising in local newspapers and newsletters relevant to the target population. Key inclusion criteria were age over 50 years, body mass index (BMI) ≤40, a confirmed clinical diagnosis of knee OA based on American College of Rheumatology criteria,[9] and ability to walk unassisted for 20 minutes. Key exclusion criteria were contraindications to exercise, partial or total joint replacements in any lower extremity joint, knee osteotomy, glucocorticoid or hyaluronic acid injections in either knee within the past three months, and other health conditions that could impact motor function. More information on the inclusion and exclusion criteria of each clinical trial is provided in Supplementary Table 1. All study procedures received approval from the Boston University Institutional Review Board, and participants provided informed consent before any study procedures. For each participant, an "index knee" was identified as the knee that had been diagnosed with knee OA by a physician, or if both knees were diagnosed with OA, the knee experiencing more pain was chosen. In cases in which pain scores were equal, the index knee was selected randomly.

Experimental protocol. The experimental protocol and the order of the assessments are shown in Figure 1. Additionally, participants completed the Knee Injury and Osteoarthritis Outcome Score (KOOS) during the visit.

Pain during exercise (exposure). All participants completed a single exercise bout only for the index knee on an isokinetic dynamometer (System3, Biodex). Exercises included three maximum voluntary isometric trials each of ankle plantar flexion, knee extension, and knee flexion and two isokinetic flexion–extension trials at 60°/s and 120°/s each with five repetitions (Figure 1). We collected pain scores for the index knee on a 0 to 10 numeric rating scale (NRS) after each trial. The exposure was the maximum pain rating across all exercise trials in the exercise bout. We used maximal pain during the exercise because prior studies suggested that this measure has greater clinical validity than a measure of change in pain from before the exercise to maximal pain during the exercise in people with chronic musculoskeletal pain.[10]

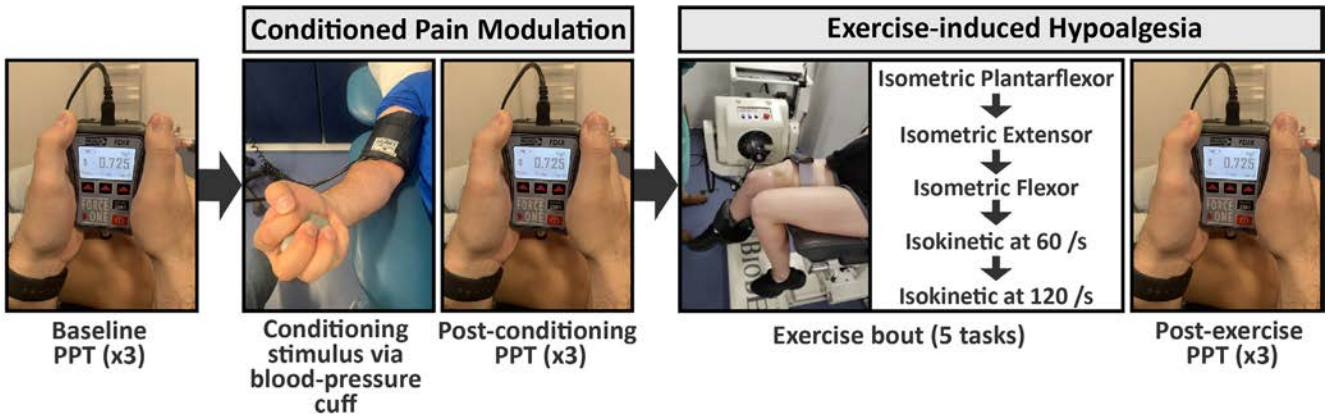

Figure 1. Experimental protocol of conditioned pain modulation and exercise-induced hypoalgesia. During the exercise session, there were multiple trials of each exercise task (five tasks in total). Pain score for the index knee was collected on a 0 to 10 numeric rating scale at every trial. PPT, pressure pain threshold.

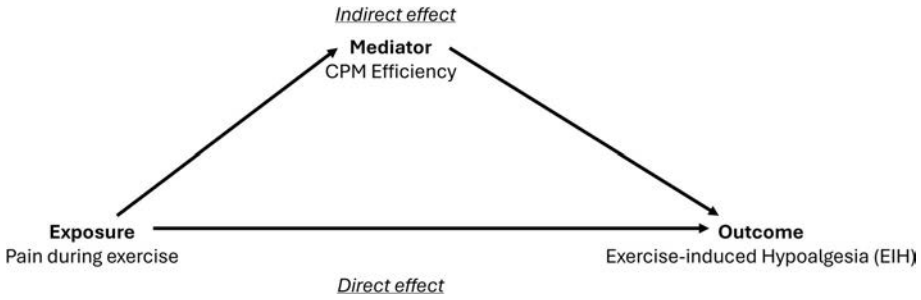

Figure 2. Directed acyclic graph indicating indirect and direct effects of the mediation analysis model. CPM, conditioned pain modulation.

EIH (outcome). We assessed EIH as the change in the pressure pain threshold (PPT) from before to immediately after the exercise bout. We assessed PPT using a handheld algometer (Force One; WAGNER) applied at a constant rate (0.5 kgf/s) at the center of the patella of the index knee while patients were in the supine position with a bolster placed under the knee[7] (Figure 1). PPT measurements were collected by three investigators who were certified on previously published protocols in which we have reported intraclass coefficients[11] between 0.85 and 0.90. The pressure at which the participant first felt the pressure change to slight pain was recorded as the PPT (in kgf). Three trials of PPTs were obtained before and after exercise in the same position. We used the average of the last two trials for the analysis. An EIH response was defined to be present if the ratio of the postexercise PPT to pre-exercise PPT was >1 (ie, improvement in PPT after exercise).[7]

CPM (mediator). CPM is known to assess the effectiveness of the descending pain modulation pathway, employing a pain inhibits pain paradigm.[4,7,12] CPM was assessed before the EIH session (Figure 1). We used PPT at the patella of the index knee as the test stimulus (averaging results from the last two trials) both before and after inducing forearm ischemia pain as the conditioning stimulus. We inflated a blood pressure cuff on the right upper arm to 10 mm Hg above systolic pressure and instructed participants to perform hand-gripping exercises using a sponge ball until they reported forearm pain rated as 4 of 10 or until two minutes elapsed if the pain did not reach 4 of 10. Subsequently, we re-evaluated PPT at the patella of the study knee (averaging results from the last two trials) before releasing the cuff pressure. We defined efficient CPM as a post-conditioning PPT/preconditioning PPT ratio >1.

Statistical analyses. We implemented a causal mediation analysis via natural-effects modeling to investigate the association between pain during exercise and the EIH response (present or absent) (ie, total effects), as well as the mediating role of CPM (efficient vs inefficient) (ie, indirect effects) on the relationship between pain during exercise and EIH response, while adjusting for age, sex, and BMI (Figure 2). Age, sex, and BMI were included as confounders because they are known to be associated with the pain threshold, which in turn is linked to both the exposure and the outcome in this analysis.[13,14] The linearity assumption was evaluated by examining the relationship between continuous confounders and the log-odds of the EIH response. We estimated the parameters of natural-effects models corresponding to direct, indirect, and total effects using the Vansteelandt et al imputation-based method.[15] Analyses were performed using R (version 4.2.2) (Supplementary File 1). Data from all available participants were used, and no a priori sample size estimations were performed.

RESULTS

Of the 104 participants enrolled in the two studies, data were available from 68 participants for this analysis (Table 1). Participants were mainly women, with ages ranging from 51 to 79 years, and many were overweight. On average, the KOOS scores during the study visit indicated moderate to severe pain and limitations of physical function. Overall, participants experienced mild to moderate pain during exercise (mean 3.4 of

Table 1. Participant characteristics*

	Total (n = 68)[a]
Age, mean (SD), y	65.0 (7.0)
Female, n (%)	47 (69.1)
BMI, mean (SD)	29.0 (4.6)
BMI, n (%)	
<25	14 (21)
≥25 and ≤30	26 (38)
>30	28 (41)
KOOS pain score (0-100), mean (SD)	61.9 (12.8)
KOOS ADL (0-100), mean (SD)	72.3 (16.2)
Pain during exercise (0-10), mean (SD)	3.4 (2.3)
Pain during exercise, n (%)	
<1	6 (9)
≥1 and ≤3	35 (51)
>3	27 (40)
EIH present, n/N (%)	36/68 (52.9)
Had efficient CPM, n/N (%)	40/68 (58.8)

* ADL, activity of daily living; BMI, body mass index; CPM, conditioned pain modulation; EIH, exercise-induced hypoalgesia; KOOS, Knee Injury and Osteoarthritis Outcome Score.
[a] Of the 104 participants enrolled in the two studies, 16 were excluded for missing the EIH assessment, 14 were excluded for missing the CPM assessment, and 6 were excluded for missing both the EIH and CPM assessments.

10 [95% confidence interval (CI) 2.9–3.9]). On a 0 to 10 pain rating scale, 9% reported pain levels lower than 1, 51% reported pain levels between 1 and 3, and the remaining 40% reported pain levels greater than 3. Among 68 participants, 53% exhibited an EIH response and 59% exhibited an efficient CPM response (Table 1). In the subsample of people who reported knee pain ≥5 with the exercise (n = 20), 70% experienced EIH.

In our causal mediation analysis, for each one-unit increase in pain during exercise, the odds of experiencing EIH in persons with knee OA increased by 43% (odds ratio [OR] 1.43, CI 1.05–1.94). This effect, adjusted for age, sex, and BMI, was primarily driven by the natural direct effect (OR 1.43, 95% CI 1.05–1.95), representing the effect of exercise pain on EIH not mediated by CPM efficiency. The natural indirect effect through CPM efficiency was negligible (OR 1.00, 95% CI 0.96–1.04), indicating that CPM efficiency did not mediate the effect of exercise-induced pain on EIH response.

DISCUSSION

We observed that mild to moderate pain during knee exercise was associated with greater odds of having EIH response in people with knee OA. However, CPM efficiency did not appear to mediate this relationship. Our results suggest that mild to moderate pain with exercise may have beneficial analgesic effects immediately following the exercise. It has been suggested that mechanisms other than descending pain inhibition that involve endocannabinoid, serotonergic, opioid, immune, and autonomic systems and/or distraction due to pain may underlie the relationship between pain increase with exercise and the presence of EIH.[12]

All major guidelines recommend exercise as a core intervention for managing pain in people with knee OA.[2] However, pain evoked by exercise remains an important barrier to exercise adherence because patients with knee OA often believe that activities that cause pain must be harmful.[3] Clinicians currently do not have any evidence-based recommendations to educate patients about dealing with pain evoked by exercise. Our findings suggest that mild to moderate pain (3–4 on a 0–10 scale) during exercise may increase the likelihood of experiencing analgesic effects immediately after exercise. Prior studies on this topic have been few and have had conflicting findings. In healthy women, the EIH response was greater after exercise when pain was induced using blood flow restriction versus pain-free exercise.[12] In contrast, another study reported that muscle pain ratings during exercise after experimentally induced pain via injection with saline were not related to subsequent EIH response in healthy adults.[16] In people with chronic pain due to shoulder myalgia or knee OA, prior studies suggested greater EIH response when exercising nonpainful muscles rather than painful muscles.[17,18] However, our study differs from these prior studies in people with chronic pain in one key aspect. The prior studies compared exercising painful muscles (ie, shoulder muscles in people with shoulder myalgia and knee muscles in people with knee OA) to exercising nonpainful muscles (ie, leg muscles in people with shoulder myalgia and upper extremity muscles in people with knee OA).[17,18] However, in our study, all participants completed the same set of knee exercises, and we examined the association of pain during this exercise with EIH. It is possible that there was variability in the EIH response after exercising painful muscles in the cohorts of prior studies, but this was not reported. In a recent study, the authors examined EIH after knee exercise with varying levels of blood flow restriction in people with end-stage knee OA and reported an increase in pain with blood flow restriction and an increase in PPT, somewhat consistent with our study.[19] Our study provides important new information about variability in EIH response as being related to pain during knee exercises that are commonly prescribed to people with knee OA. It is important to note that the magnitude of the pain increase during exercise in our cohort was limited. It is plausible that larger increases in pain may lead to different outcomes. For instance, in people with low back pain, those who experienced a pain increase of <2 of 10 during walking had EIH, whereas those who experienced a larger pain increase of ≥2 of 10 did not.[20] However, in our small sample (n = 20) of people with knee pain ≥5 during the exercise, 70% showed an EIH response, suggesting that the moderating effect of pain intensity may require further study.

Our findings may have implications for responses to exercise interventions in people with knee OA. In a study of an eight-week neuromuscular exercise intervention for people with knee OA, pain was accepted during exercise, with pain (as measured by a 0–10 NRS) between 0 and 2 being considered safe, pain between 3 and 5 being considered acceptable, and pain >5 being considered high risk.[21] The results showed that this approach resulted in a reduction in exercise-evoked pain flares (ie, change in pain from before to after exercise) over the course of the intervention. Interestingly, the participants who were deemed noncompliant (ie, attended <75% of the sessions) had a larger exercise-evoked pain flare (0.79 NRS) versus the compliant group (0.43 NRS). The authors speculated that this difference may have influenced compliance. A systematic review reported that significant benefits were seen in the short term with protocols that included painful exercises versus protocols that included pain-free exercises in people with chronic musculoskeletal pain conditions, with no clear difference over the medium term and long term.[22] Our findings provide further evidence that may help clinicians educate patients about the potential safety of mild to moderate pain with exercise and boost compliance.

CPM efficiency may not mediate the relation between pain evoked by exercise and EIH response. CPM is interpreted as a proxy for the descending pain inhibitory pathway, involving endogenous opioids and serotonin—substances well-known for their release after exercise as well.[4,7] Consequently, the activation of opioidergic descending inhibitory pathways has been speculated as a shared mechanism of CPM and EIH.[7] This can be supported

by several previous studies showing a significant association between the efficiency of CPM and EIH, which have shown that CPM explains 23% of EIH responses after adjusting for demographic confounders.[4] However, to our knowledge, our study is the first to investigate whether CPM mediates the association between pain during exercise and EIH. Our finding that CPM did not mediate the relation of pain during exercise with EIH could be due to several reasons. First, the pain increase (95% CI 2.9–3.9) during exercise in our study may be insufficient to activate the descending inhibitory system, considering the pain level we set for the conditioning stimulus was ≥4 during the CPM. It is also possible that the descending inhibitory system was already fully engaged owing to ongoing chronic pain in these individuals and there was no remaining reserve for further activation. Second, we regarded pain evoked at the affected knee joint during exercise as a conditioning stimulus for the EIH response, whereas we used forearm pain due to cuff-induced ischemia at a nonpainful location as a conditioning stimulus for the CPM. Lastly, the overall age-related, as well as the OA-derived, attenuation in the efficiency of the descending pain inhibitory system can be one factor leading to a nonsignificant mediation effect.[23] Our findings are aligned with those in the study by Ellingson et al, in which EIH was observed after both painful and pain-free exercises, leading the authors to conclude that CPM is likely not the primary mechanism of EIH.[12] However, given that our analysis was limited to a single session of exercise, the potential mediating role of CPM over a longer exercise intervention may be examined in future studies.

Some limitations in this study should be acknowledged. We only examined EIH after exercise involving the painful knee. It is plausible that the findings may differ with exercises involving other joints or other types of exercise (eg, aerobic exercise, upper extremity exercise). Also, when determining the EIH response, we did not account for the SE of PPT measurement, which could have resulted in categorizing two individuals with only small differences in PPT into different groups. Similarly, we did not consider different thresholds of increase in PPT to define EIH or CPM, which may alter participant classification. We cannot rule out interactions due to centrally acting medications, such as serotonin and norepinephrine reuptake inhibitors, and/or central effects due to cancer, which was only excluded in one of the cohorts. Finally, we only examined the relation of pain with exercise and an immediate EIH response after a single bout of exercise. Whether pain with exercise is related to better outcomes after a long-term exercise intervention needs further study.

In conclusion, approximately half of the participants with knee OA did not experience an EIH response after a single knee exercise session. Mild to moderate increases in pain during knee exercise were associated with a greater likelihood of EIH in people with knee OA. CPM efficiency did not mediate the association of pain increase during exercise with EIH. Clinicians may educate patients that some pain with exercise is likely safe and may help with treatment benefits.

ACKNOWLEDGMENTS

We would like to acknowledge Brian Friscia and Michael Rose for assisting with data collection and data processing. We also want to thank the participants for their contribution.

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 Kumar 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

NCT04243096 was conducted as a collaboration between Boston University, Pfizer, and Eli Lilly. Investigators from Pfizer and Eli Lilly were involved in the study design, data analysis and interpretation, and drafting of the manuscript. Publication of this article was not contingent upon approval by Pfizer and Eli Lilly.

ADDITIONAL DISCLOSURES

Authors Tuttle, Demanuele, and Wacnik are employees of Pfizer Inc. Author Viktrup is an employee of Eli Lilly and Company.

REFERENCES

1. Cui A, Li H, Wang D, et al. Global, regional prevalence, incidence and risk factors of knee osteoarthritis in population-based studies. EClinicalMedicine 2020;29–30:100587.
2. Conley B, Bunzli S, Bullen J, et al. Core recommendations for osteoarthritis care: a systematic review of clinical practice guidelines. Arthritis Care Res (Hoboken) 2023;75(9):1897–1907.
3. Dobson F, Bennell KL, French SD, et al. Barriers and facilitators to exercise participation in people with hip and/or knee osteoarthritis: synthesis of the literature using behavior change theory. Am J Phys Med Rehabil 2016;95(5):372–389.
4. Lemley KJ, Hunter SK, Bement MK. Conditioned pain modulation predicts exercise-induced hypoalgesia in healthy adults. Med Sci Sports Exerc 2015;47(1):176–184.
5. Rice D, Nijs J, Kosek E, et al. Exercise-induced hypoalgesia in pain-free and chronic pain populations: state of the art and future directions. J Pain 2019;20(11):1249–1266.
6. Hansen S, Vaegter HB, Petersen KK. Pretreatment exercise-induced hypoalgesia is associated with change in pain and function after standardized exercise therapy in painful knee osteoarthritis. Clin J Pain 2020;36(1):16–24.
7. Fingleton C, Smart KM, Doody CM. Exercise-induced hypoalgesia in people with knee osteoarthritis with normal and abnormal conditioned pain modulation. Clin J Pain 2017;33(5):395–404.
8. Balasch-Bernat M, Lluch E, Vaegter HB, et al. Should exercises be painful or not? Effects on clinical and experimental pain in individuals with shoulder pain. J Pain 2021;22(10):1246–1255.
9. Altman R, Asch E, Boch D, et al. Development of criteria for the classification and reporting of osteoarthritis. Classification of Osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the

American Rheumatism Association. Arthritis Rheum 1986; 29(8): 1039–1049.

10. Knox PJ, Simon CB, Pohlig RT, et al. a standardized assessment of movement-evoked pain ratings is associated with functional outcomes in older adults with chronic low back pain. Clin J Pain 2021; 38(4):241–249.
11. Neogi T, Guermazi A, Roemer F, et al. Association of joint inflammation with pain sensitization in knee osteoarthritis: the Multicenter Osteoarthritis Study. Arthritis Rheumatol 2016;68(3):654–661.
12. Ellingson LD, Koltyn KF, Kim JS, et al. Does exercise induce hypoalgesia through conditioned pain modulation? Psychophysiology 2014;51(3):267–276.
13. Gløersen M, Steen Pettersen P, Neogi T, et al. Associations of body mass index with pain and the mediating role of inflammatory biomarkers in people with hand osteoarthritis. Arthritis Rheumatol 2022; 74(5):810–817.
14. Ostrom C, Bair E, Maixner W, et al. Demographic predictors of pain sensitivity: results from the OPPERA study. J Pain 2017;18(3):295–307.
15. Vansteelandt S, Bekaert M, Lange T. Imputation strategies for the estimation of natural direct and indirect effects. Epidemiol Methods 2012;1(1):131–158.
16. Hansen S, Petersen KK, Sloth E, et al. Hypoalgesia after exercises with painful vs. non-painful muscles in healthy subjects - a randomized cross-over study. *Scand* J Pain 2021;22(3):614–621.
17. Lannersten L, Kosek E. Dysfunction of endogenous pain inhibition during exercise with painful muscles in patients with shoulder myalgia and fibromyalgia. Pain 2010;151(1):77–86.
18. Burrows NJ, Booth J, Sturnieks DL, et al. Acute resistance exercise and pressure pain sensitivity in knee osteoarthritis: a randomised crossover trial. Osteoarthritis Cartilage 2014;22(3):407–414.
19. Ogrezeanu DC, López-Bueno L, Sanchís-Sánchez E, et al. Exercise-induced hypoalgesia with end-stage knee osteoarthritis during different blood flow restriction levels: sham-controlled crossover study. PM R 2023;15(12):1565–1573.
20. Vaegter HB, Petersen KK, Sjodsholm LV, et al. Impaired exercise-induced hypoalgesia in individuals reporting an increase in low back pain during acute exercise. Eur J Pain 2021;25(5):1053–1063.
21. Sandal LF, Roos EM, Bøgesvang SJ, et al. Pain trajectory and exercise-induced pain flares during 8 weeks of neuromuscular exercise in individuals with knee and hip pain. Osteoarthritis Cartilage 2016;24(4):589–592.
22. Smith BE, Hendrick P, Smith TO, et al. Should exercises be painful in the management of chronic musculoskeletal pain? A systematic review and meta-analysis. Br J Sports Med 2017;51(23): 1679–1687.
23. Edwards RR, Fillingim RB, Ness TJ. Age-related differences in endogenous pain modulation: a comparison of diffuse noxious inhibitory controls in healthy older and younger adults. Pain 2003;101(1–2): 155–165.

Arthritis Care & Research
Vol. 77, No. 7, July 2025, pp 906–915
DOI 10.1002/acr.25514

American College of Rheumatology
Empowering Rheumatology Professionals

Sequence Analysis to Phenotype Health Care Patterns in Adults With Musculoskeletal Conditions Using Primary Care Electronic Health Records

Smitha Mathew,[1] George Peat,[1,2] Emma Parry,[1] Ross Wilkie,[1] Kelvin P. Jordan,[1] Jonathan C. Hill,[1] and Dahai Yu[1]

Objective. The aim of this study was to apply sequence analysis (SA) to phenotype health care patterns of adult patients with musculoskeletal (MSK) conditions using primary care electronic health records and to investigate the association between these health care patterns and patients' self-reported outcomes after consultation.

Methods. Data from the Multilevel Integrated Data for musculoskeletal health intelligence and Actions program conducted in North Staffordshire and Stoke-on-Trent, United Kingdom, was used. The study included patients aged ≥18 years who consulted primary care for MSK conditions between September 2021 and July 2022. SA was employed to categorize patients with similar health care patterns in primary care in the five years before their index consultation in respect to consultations, analgesic prescriptions, imaging, physiotherapy, and secondary care referrals. Associations of sociodemographic characteristics and self-reported outcome with clusters were determined.

Results. In total, 1,875 patients consulting primary care for MSK conditions were available for analysis. SA identified five clusters of previous health care patterns among patients with MSK conditions, including "increasing consultation and analgesia" (5.60%), "low consultation and health care use" (57.39%), "high consultation and health care use" (8.32%), "low consultation but high analgesia" (13.01%), and "low consultation but moderate health care use" (15.68%). Patients in the "high consultation and health care use" group were predominantly female, were older, had obesity, had more comorbidities, and lived in the most deprived areas compared to those in the "low consultation and health care use" group. Additionally, self-reported outcomes varied significantly among clusters, with patients in the "high consultation and health care use" group reporting worse self-reported outcomes.

Conclusion. This analysis identified five distinct clusters of health care patterns for patients with MSK conditions in primary care and observed substantial variations in patients' self-reported outcomes and sociodemographic profiles across these different groups of patients.

INTRODUCTION

Musculoskeletal (MSK) conditions are a major cause of pain and disability worldwide. In the United Kingdom, more than 20 million people live with an MSK condition.[1] MSK conditions are primarily assessed and managed in primary care. It accounts for 12% to 14% of primary care consultations in adults, and a substantial portion of health care expenditure is allocated to managing these conditions.[2] A range of different interventions are recommended for the management of MSK conditions, including providing advice on self-management and exercise, referring patients for nonpharmacologic treatments such as

ISRCTN Registry identifier: ISRCTN18132064.

The views expressed herein are those of the authors and do not necessarily represent those of the NHS, the National Institute for Health and Care Research, or the Department of Health and Social Care.

The MIDAS-GP study has been funded by the Nuffield Foundation's Oliver Bird Fund and Versus Arthritis (OBF/43390; visit www.nuffieldfoundation.org and www.versusarthritis.org) as well as by Keele Clinical Trials Unit. Ms Mathew's work was supported by a PhD studentship from the Faculty of Medicine and Health Sciences at Keele University. Dr Parry's work was supported by a National Institute for Health and Care Research (NIHR) Academic Clinical Lectureship (CL-2020-10-001). Dr Yu and Dr Jordan's work was partly suported by the NIHR Applied Research Collaboration West Midlands.

[1]Smitha Mathew, MSc, George Peat, MSc, PhD, Emma Parry, MRCGP, PhD, Ross Wilkie, BSc, PhD, Kelvin P. Jordan, MSc, PhD, Jonathan C. Hill, MSc, PhD, Dahai Yu, MSc, PhD: Keele University, Keele, Staffordshire, United Kingdom; [2]George Peat, MSc, PhD: Sheffield Hallam University, Sheffield, United Kingdom.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25514).

Author disclosures and graphical abstract are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25514.

Address correspondence via email to Dahai Yu, MSc, PhD, at d.yu@keele.ac.uk.

Submitted for publication October 7, 2024; accepted in revised form January 29, 2025.

SIGNIFICANCE & INNOVATIONS

- Our study identified five distinct patterns of health care utilization in primary care among adult patients with musculoskeletal conditions using sequence analysis.
- We observed inequalities in health care utilization patterns based on patients' characteristics and significant variations in patients' self-reported outcomes across different clusters of health care utilization patterns. Specifically, patients from socioeconomically deprived areas who were predominantly older and female and who had obesity and multiple comorbidities showed higher consultation rates, higher health care use, and poorer short-term outcome.
- These findings highlighted the importance of addressing disparities in health care access and the need for targeted interventions for patients at risk of poorer health outcomes.

physiotherapy, and prescribing analgesics to alleviate pain and symptoms.[3]

Pain associated with MSK conditions leads to high health care use, and patients seeking health care may find themselves consulting a diverse range of health care professionals and receiving a mix of analgesic prescriptions, imaging, physiotherapy, and secondary care referrals.[4,5] Understanding patterns within these interactions can provide insights into how different patient subgroups use health care services. By analyzing these patterns, health care providers can identify the specific needs of patient subgroups. It enables health care planners to allocate resources more strategically, ensuring that they are directed to where they are most needed. Moreover, a comprehensive understanding of care patterns helps identify service gaps and areas for improvement. This knowledge allows for the optimization of health care delivery by addressing disparities in service utilization. Overall, it supports identifying specific health care needs, informs strategic resource allocation, and contributes to improving health care delivery and patient outcomes.[6,7]

Patients' self-reported outcome measures are valuable for evaluating perceptions of health, symptoms, and the effectiveness of MSK management.[8] These measures capture information primarily focusing on pain levels, activity limitations, and overall quality of life rather than clinical measures.[9] Several studies have highlighted an association between chronic pain and increased health care utilization.[10–12] Additionally, a correlation has been observed between low health-related quality of life and high health care utilization.[12] Evidence from a primary care prospective observational cohort study further indicates that subgroups of individuals with different levels of risk for poor MSK pain outcomes exhibit different levels of health care utilization.[13] Relating health care utilization patterns to patients' self-reported outcomes might direct attention toward potentially poorly targeted or ineffective patterns of care.

In recent years, sequence analysis (SA) has emerged as a promising analytical approach in health care research due to its ability to uncover valuable insights and patterns from real-world data.[14] SA is used to analyze ordered sets of data, often referred to as sequences. This method is commonly used in social science to identify patterns in life course trajectories and to study transitions into adulthood[15,16] or career patterns[17] by examining longitudinal data representing events experienced by individuals over time. In health care, SA allows researchers to analyze sequences of medical events, such as diagnoses, treatments, and procedures, to understand disease progression and care pathways.[6,18,19] SA enables the exploration of health care utilization patterns, including patient journeys through the health care system, patterns of service utilization, and transitions between different levels of care.[20–22]

A conventional SA involves three steps: defining events as sequences of successive categorical states, calculating dissimilarities between pairs of sequences, and building a typology of the sequences.[23] The states in the sequence should be clinically meaningful and relevant to the research objectives. Dissimilarity is a quantitative measure indicating the degree to which two individuals followed distinct sequences. There are different dissimilarity measures based on alignment and nonalignment techniques. The choice of dissimilarity measure may affect the results of SA; therefore, researchers select an appropriate measure aligned with their research objective.[24] Finally, a cluster analysis is performed to classify individuals with similar sequences.

In this study, we focus on the identification of different health care patterns among adult patients with MSK conditions in primary care more than five years before their index consultation, as well as examining the effect of these patterns on patients' self-reported outcomes. By examining historic care patterns, we can comprehensively understand the various treatment strategies patients have experienced, which might influence their current health status and outcome.

Therefore, the primary objective of this study was to apply SA to phenotype health care patterns of patients with MSK conditions from routinely collected primary care electronic health records (EHRs). The secondary objective was to investigate the association between the identified health care patterns and patients' self-reported outcomes after consultation.

MATERIALS AND METHODS

Data source and population. The Multilevel Integrated Data for musculoskeletal health intelligence and Actions (MIDAS) program, funded by the Nuffield Foundation and Versus Arthritis, aims to develop a comprehensive, place-based system for MSK health data in North Staffordshire and Stoke-on-Trent, United Kingdom. MIDAS-GP is one observational cohort study within the overall MIDAS project, and is designed to collect, link,

and explore data from patient-report, electronic health records, and other sources for adults presenting with common, painful MSK conditions presenting in general practiced (GP). The study focuses on integrating data from various clinical settings to enhance MSK care pathways. The prespecified MIDAS-GP study protocol is available at Open Science Framework (https://osf.io/e542w/). The study received ethical approval from Yorkshire & The Humber-Leeds West Research Ethics Committee (Reference: 21/YH/0178).

The eligible participants for this study included patients aged 18 years and older, registered with 30 participating general practices and who consulted any primary care health care professional within the practice for a painful, noninflammatory MSK condition. Recruitment was conducted from September 2021 to July 2022, staggered across different practices, with recruitment periods lasting from three to six months. Relevant MSK pain-related consultations were identified using a pre-specified Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) code list (Supplementary Table S1). Eligible participants were invited to complete a baseline questionnaire on MSK health and care and were asked for their consent to link the questionnaire to EHRs. The consenting participants were further asked to complete the follow-up questionnaires at three and six months.

The information on patient's demographic, socioeconomic, comorbidities, and MSK management strategies were derived from the primary care EHR in the five years before index MSK consultation. The list of comorbidities used was produced after cross-mapping morbidities in National Institute for Health and Care Excellence (NICE) multimorbidity indicator for general practice,[25] Charlson,[26] and Elixhauser[27] comorbidity indices, and potentially relevant case-mix adjustment methods.[28] Comorbidity code lists are available at Open Science Framework (https://osf.io/e542w/). The MSK management information included MSK-related primary care consultations, relevant prescriptions for medications, referrals for imaging (eg, radiographs, magnetic resonance imaging, or CT scans), referrals for physiotherapy, and referrals for secondary care (MSK triage, rheumatology, trauma, and orthopedic departments). Patients' neighborhood deprivation was also considered. We used the English index of multiple deprivation (IMD) 2019 rank as a composite measure of neighborhood deprivation, which covers seven domains of material deprivation including income, employment, education and skills training, health deprivation and disability, barriers to housing and services, crime, and living environment.[29] The IMD classifies the areas into five quintiles based on relative disadvantage, with quintile 1 being the most deprived and quintile 5 being the least deprived. Additionally, patients' MSK Health Questionnaire (HQ) scores at baseline, three months, and six months after index consultation were considered. The MSK-HQ is a 14-item questionnaire that captures key outcomes that patients with MSK conditions have prioritized as important for use across clinical pathways.[30] Scores range from 0 to 56, with higher scores indicating better MSK health over the past two weeks.[30] The data of this study are available upon request.

Table 1. Patients' baseline characteristics*

Variables	Patients (N = 1,875)
Sex, n (%)	
Female	1,233 (65.76)
Male	642 (34.24)
Age, mean (SD)	57.74 (15.50)
Age group, n (%)	
18–34 y	157 (8.37)
35–44 y	234 (12.48)
45–54 y	362 (19.31)
55–64 y	430 (22.93)
65–74 y	407 (21.71)
75–84 y	241 (12.85)
85+ y	44 (2.35)
BMI, mean (SD)	29.18 (6.91)
BMI, n (%)	
Underweight (<18.5)	26 (1.39)
Normal (18.5–24.9)	399 (21.28)
Overweight (25–29.9)	562 (29.97)
Obese (≥30)	608 (32.43)
Missing	280 (14.93)
Comorbidity count, n (%)	
0	829 (44.21)
1	577 (30.77)
2	329 (17.55)
3+	140 (7.47)
Index of multiple deprivation, n (%)	
Quintile 1 (most deprived)	530 (28.27)
Quintile 2	383 (20.43)
Quintile 3	398 (21.23)
Quintile 4	320 (17.07)
Quintile 5 (least deprived)	244 (13.01)
Race and ethnicity, n (%)	
White	1,788 (95.36)
Asian	31 (1.65)
Mixed	11 (0.59)
Black	28 (1.49)
Other	17 (0.91)

* BMI, body mass index.

Statistical analysis. To explore the patterns of utilization of key MSK management strategies in primary care, we employed a multichannel SA involving five domains: MSK-related consultations, analgesic prescriptions, imaging referrals, physiotherapy referrals, and secondary care referrals. The primary step in SA was defining the states within the sequence, the observation period, and the time unit. The health care patterns of patients with MSK conditions were observed for five years before their index consultation. The MSK management information was retrieved as annual count data. So, we defined three categorical states for consultations and analgesic prescriptions: "none," "low," and "high," representing zero, one to three, and four or more instances, respectively, and two categorical states for imaging, physiotherapy, and secondary care referrals: "no" and "yes" occurrence during the year (detailed in the Supplementary Material S1). If the care event is not recorded in the system, it is considered to have not occurred. We defined care sequences for each domain for each patient, with each sequence consisting of five states (one for each year).

For the analysis of sequences, we chose optimal matching (OM) edit distance, the most often used approach to measure the dissimilarity between pairs of sequences.[14] OM measures the dissimilarity between two sequences by determining the minimum cost required to transform one sequence into another by edit operations such as insertion, deletion, or substitution of states. We opted for a data-driven cost for insertion/deletion, and substitutions based on the frequency of the states in the sequences, referred to as INDELSLOG. In this approach, insertion/deletion costs were calculated initially as the logarithm of the inverse of the relative frequency of the states, as log(2/[1 + f]), in which "f" is the relative frequency of the states. Then, the substitution costs between the two states are computed by summing their insertion/deletion costs.[31] The rationale behind this approach is that inserting or deleting rare states is more costly than inserting or deleting frequent states, and substituting rarely observed states costs more than substituting common states.[31] The multidomain dissimilarity matrix was computed by adding the domain-specific dissimilarity matrixes. Based on the computed dissimilarity matrix, we performed an agglomerative hierarchical cluster analysis with Ward's linkage to classify patients with similar care patterns. The optimum number of clusters was determined based on the dendrogram, inertia jump curve, cluster quality indices, and clinical relevance and interpretability (explanation of the selection criteria is given in the Supplementary Material S2). To visualize the care patterns, we used sequence index plots and state distribution plots provided by the SA. State distribution plot shows the distribution of states for each time unit, whereas each line in the sequence index plot represents an individual sequence.[24]

We compared patients' demographic and health characteristics among the derived clusters using the chi-square test, *t*-test, and analysis of variance. A multinomial logistic regression model was used to assess the association between patients' profiles and cluster membership. A linear mixed model was used to test

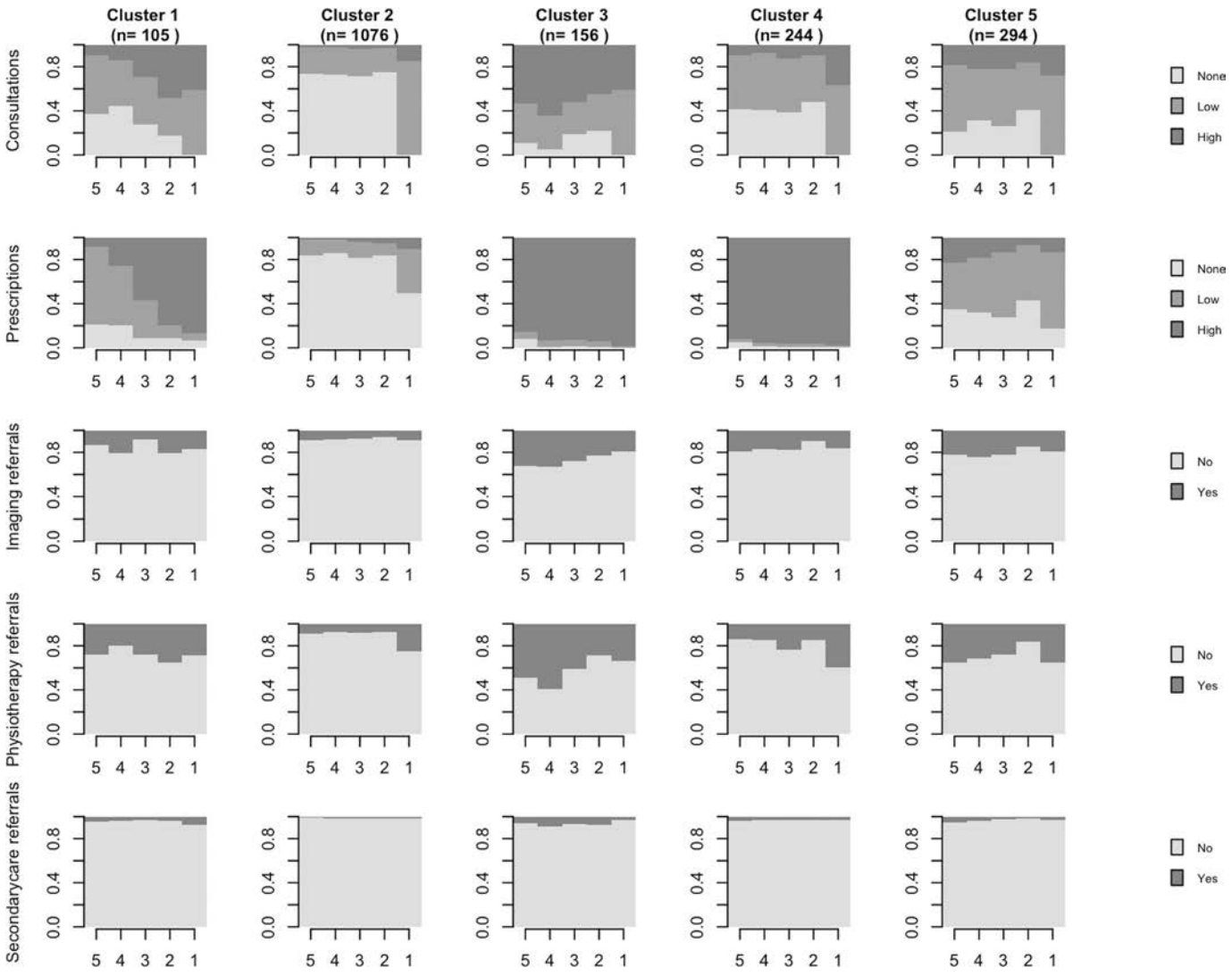

Figure 1. State distribution plot of care sequence typology by domain (consultations, prescriptions, imaging, physiotherapy, and secondary care referrals).

the difference in patient-reported MSK-HQ scores between clusters. The model included the fixed, categorical effects of cluster, time, cluster-by-time interaction, sex, comorbidities, and IMD, alongside continuous, fixed covariates for age and body mass index (BMI). To account for within-participant variability, an unstructured covariance structure was applied to model the within-participant errors. The missing data in BMI (n = 280) were imputed by multiple imputation using chained equations.[32] Sensitivity analyses were conducted to ensure the reproducibility of the results. For this, patients with MSK conditions were subgrouped into those with osteoarthritis (OA) and those with low back pain (LBP), and the SA was repeated within these subgroups. The SA was conducted using the TraMineR and WeightedCluster packages in R, and all other analyses were performed using Stata version 18 (StataCorp).

RESULTS

Participants. A total of 2,008 patients (14.9%) responded at baseline, of whom 1,875 patients consented and were successfully linked to their EHRs and hence form the primary population for analysis (detailed in Supplementary Figure S1). Among these patients, the mean ± SD age was 57.74 ± 15.50 years, and the mean ± SD BMI was 29.18 ± 6.91. Female participants accounted for 65.76% of the patients, 32.43% were classified as obese, and 28.27% were from the most deprived areas (Table 1). Patients' care sequences of each domain were presented in sequence index plots (Supplementary Figure S2).

SA. By the multichannel SA of the domains—MSK consultations, analgesic prescriptions, imaging referrals, physiotherapy referrals, and secondary care referrals—patients with similar care sequences were classified into five distinct clusters (Figure 1) based on the dendrogram, inertia jump curve, and cluster quality indices (Supplementary Table S2 and Supplementary Figures S3 and S4). The characteristics of the identified clusters are as follows:

- Cluster 1 (n = 105, 5.60%) patients were characterized by a marked increase in high-level (ie, four or more) consultations and analgesic prescriptions over the five years, accompanied by moderate imaging and physiotherapy, and minimum secondary care referrals. This cluster can be labeled as "increasing consultation and analgesia."
- Cluster 2 (n = 1,076, 57.39%) consisted of patients with low-level (one to three) consultations and analgesic

Table 2. Multinomial logistic regression model for association between patients' characteristics and different clusters*

Patient characteristic	Clusters of similar care sequences			
	Increasing consultation and analgesia, OR (95% CI)	High consultation and health care use, OR (95% CI)	Low consultation but high analgesia, OR (95% CI)	Low consultation but moderate health care use, OR (95% CI)[b]
Sex				
Male	1	1	1	1
Female	1.51 (0.97–2.35)	2.55 (1.69–3.88)[a]	1.79 (1.30–2.46)[a]	1.87 (1.40–2.51)[a]
Age group				
18–34 y	1	1	1	1
35–44 y	0.62 (0.23–1.69)	1.24 (0.43–3.57)	2.98 (0.96–9.21)	1.41 (0.75–2.62)
45–54 y	1.71 (0.76–3.84)	2.82 (1.10–7.21)[a]	5.73 (1.97–16.66)[a]	1.80 (0.99–3.24)
55–64 y	1.16 (0.50–2.69)	3.50 (1.39–8.82)[a]	7.04 (2.45–20.21)[a]	1.87 (1.04–3.33)[a]
65–74 y	1.96 (0.85–4.52)	5.31 (2.09–13.49)[a]	11.66 (4.06–33.53)[a]	2.79 (1.55–5.01)[a]
75–84 y	2.73 (1.10–6.77)[a]	9.05 (3.41–24.03)[a]	18.99 (6.46–55.84)[a]	3.20 (1.68–6.11)[a]
85+ y	1.69 (0.32–8.89)	8.42 (2.23–31.80)[a]	12.51 (3.26–47.94)[a]	4.31 (1.65–11.26)[a]
BMI				
Underweight/normal (<25)	1	1	1	1
Overweight (25–29.9)	0.98 (0.52–1.84)	1.04 (0.59–1.84)	1.18 (0.77–1.80)	1.22 (0.84–1.76)
Obese (≥30)	2.03 (1.15–3.59)[a]	2.54 (1.52–4.25)[a]	1.79 (1.18–2.71)[a]	1.80 (1.25–2.59)[a]
Comorbidity count				
0	1	1	1	1
1	1.74 (1.07–2.83)[a]	1.58 (0.97–2.56)	2.19 (1.51–3.18)[a]	1.34 (0.99–1.82)
2	1.93 (1.07–3.47)[a]	4.38 (2.72–7.05)[a]	3.90 (2.60–5.85)[a]	1.28 (0.86–1.91)
3+	3.49 (1.65–7.38)[a]	6.65 (3.60–12.28)[a]	5.55 (3.21–9.61)[a]	2.07 (1.19–3.61)[a]
Index of multiple deprivation				
Quintile 1 (most deprived)	1.42 (0.71–2.82)	2.65 (1.34–5.23)[a]	2.09 (1.21–3.62)[a]	1.22 (0.78–1.90)
Quintile 2	1.24 (0.61–2.52)	2.33 (1.16–4.65)[a]	1.84 (1.05–3.22)[a]	1.11 (0.70–1.76)
Quintile 3	0.70 (0.33–1.49)	0.78 (0.36–1.66)	1.30 (0.75–2.26)	0.85 (0.55–1.35)
Quintile 4	0.86 (0.40–1.83)	1.46 (0.70–3.02)	1.34 (0.76–2.38)	0.86 (0.53–1.39)
Quintile 5 (least deprived)	1	1	1	1

* BMI, body mass index; CI, confidence interval; OR, odds ratio.
[a] These are significant results.
[b] The reference cluster is low consultation and healthcare use.

prescriptions mainly in the index year, and minimal imaging, physiotherapy, and secondary care referrals. This cluster can be labeled as "low consultation and health care use."

- Cluster 3 (n = 156, 8.32%) was made up of patients with consistently higher levels of consultation, analgesic prescriptions, imaging, physiotherapy, and secondary care referrals. This cluster can be labeled as "high consultation and health care use."
- Cluster 4 (n = 244, 13.01%) included patients with low-level (one to three) consultations, low imaging, physiotherapy, and secondary care referrals, but having higher levels (four or more) of analgesic prescriptions over the five years. This cluster can be labeled as "low consultation but high analgesia."
- Cluster 5 (n = 294, 15.68%) consisted of patients with low-levels (one to three) of consultations, analgesic prescriptions, and secondary care referrals, but moderate levels of imaging and physiotherapy referrals. This cluster can be labeled as "low consultation but moderate health care use."

Potential predictors of cluster membership. Patients' characteristics by clusters of similar care patterns were presented in Supplementary Table S3. Table 2 shows the findings of the multinomial logistic regression model computed to examine potential predictors of the identified clusters. Odds ratios were calculated to indicate the likelihood of being in a particular cluster compared to the reference cluster. The reference cluster used in the analysis was "low consultation and health care use." Female patients were significantly more likely to be in "high consultation and health care use," "low consultation but high analgesia," and "low consultation but moderate health care use" clusters, as compared to being in the "low consultation and health care use" cluster. Additionally, older age, obesity, a higher comorbidity index, and socioeconomic deprivation (most deprived) were identified as significant predictors for membership in the "increasing consultation and analgesia," "high consultation and health care use," "low consultation but high analgesia," and "low consultation but moderate health care use" clusters.

The effect of health care patterns and patients' MSK-HQ scores. Table 3 presents the adjusted estimates for the association between health care patterns and MSK-HQ score. Figure 2 illustrates the predicted mean MSK-HQ score values among different clusters at index consultation (baseline), and at three months and six months following the index consultation. The mean patient-reported MSK-HQ score was significantly lower (worse MSK health) in the "increasing consultation and analgesia," "high consultation and health care use," "low consultation but high analgesia," and "low consultation but moderate

Table 3. Longitudinal linear mixed model to assess association between clusters of similar care sequence and MSK-HQ score*

	MSK-HQ Score
Fixed effects, coefficient (95% CI)	
Intercept	26.19 (23.04 to 29.33)[a]
Cluster of similar care sequences, coefficient (95% CI)	
Increasing consultation and analgesia	−5.90 (−7.90 to −3.89)[a]
High consultation and healthcare use	−7.26 (−9.01 to −5.51)[a]
Low consultation but high analgesia	−5.79 (−7.24 to −4.35)[a]
Low consultation but moderate healthcare use	−2.73 (−4.03 to −1.43)[a]
Time, coefficient (95% CI)	
3 mo	5.41 (4.78 to 6.04)[a]
6 mo	6.42 (5.69 to 7.16)[a]
Interaction terms cluster of similar care sequence x time, coefficient (95% CI)	
Increasing consultation and analgesia x 3 mo	−2.05 (−4.12 to 0.02)
Increasing consultation and analgesia x 6 mo	−2.82 (−3.30 to −0.57)[a]
High consultation and healthcare use x 3 mo	−5.18 (−6.92 to −3.43)[a]
High consultation and healthcare use x 6 mo	−4.57 (−6.55 to −2.58)[a]
Low consultation but high analgesia x 3 mo	−2.88 (−4.28 to −1.49)[a]
Low consultation but high analgesia x 6 mo	−2.98 (−4.61 to −1.35)[a]
Low consultation but moderate healthcare use x 3 mo	−1.93 (−3.30 to −0.57)[a]
Low consultation but moderate healthcare us x 6 mo	−1.10 (−2.65 to 0.45)
Random effects, SD	
Intercept	7.39
Time	2.76

* The reference cluster is low consultation and health care use. Model was controlled for sex, age, body mass index, comorbidity count, and index of multiple deprivation. CI, confidence interval; MSK-HQ, Musculoskeletal Health Questionnaire.
[a] These are significant results.

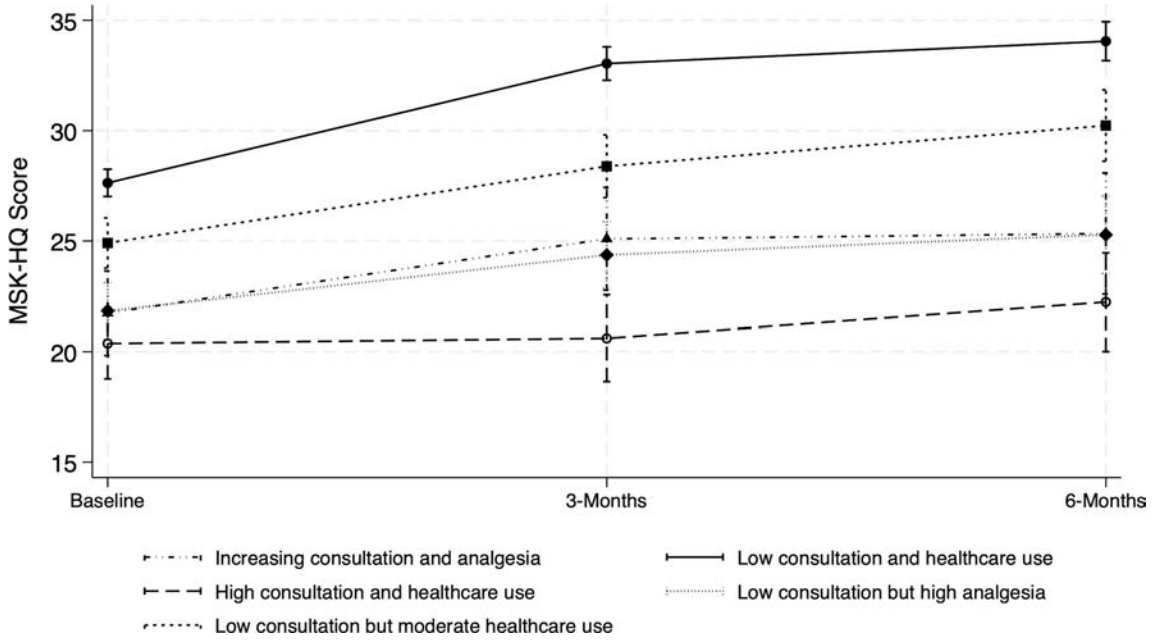

Figure 2. Predicted values of MSK-HQ score among the distinct clusters. Predicted values were controlled for sex, age, body mass index, comorbidity count, and index of multiple deprivation. MSK-HQ, Musculoskeletal Health Questionnaire.

health care use" clusters compared to the "low consultation and health care use" cluster at baseline, three months, and six months; the estimated differences in mean score are presented in Table 4. Additionally, the MSK-HQ score over time, as indicated by the interaction terms of clusters with similar care sequences and time, showed significantly less improvement at month 3 in the "high consultation and health care use" (coefficient –5.18 [95% confidence interval (CI) –6.92 to –3.43]), "low consultation but high analgesia" (coefficient –2.88 [95% CI –4.28 to –1.49]), and "low consultation but moderate health care use" (coefficient –1.93 [95% CI –3.30 to –0.57]) clusters compared to the improvement in the "low consultation and health care use" cluster. Similarly, less improvement was observed at month 6 in the "increasing consultation and analgesia" (coefficient –2.82 [95% CI –5.27 to –0.37]), "high consultation and health care use" (coefficient –4.57 [95% CI –6.55 to –2.58]), and "low consultation but high analgesia" (coefficient –2.98 [95% CI –4.61 to –1.35]) clusters (Table 3).

Sensitivity analysis. To test the reproducibility of the results, two additional SAs were performed by subgrouping the patients with conditions into those with OA and those with LBP. Agglomerative hierarchical cluster analysis with OM and INDELSLOG cost produced five clusters for patients with OA (Supplementary Figure S5), which were similar to the results obtained for patients with MSK conditions. Similarly, the analysis of patients with LBP also resulted in five clusters (Supplementary Figure S6). Another sensitivity analysis with similar SA methods was conducted excluding patients who have less than five years of continuous retrospective record, and it yielded similar clusters of the main analysis (Supplementary Figure S7).

DISCUSSION

Our study examined health care patterns among 1,875 adult patients who sought consultation for MSK conditions in primary care settings and investigated the relationship between these

Table 4. Difference in MSK-HQ score from low consultation and health care use at baseline, three months, and six months*

	MSK-HQ score		
	Baseline, difference (95% CI)	3 mo, difference (95% CI)	6 mo, difference (95% CI)
Increasing consultation and analgesia	−5.90 (−7.91 to −3.89)[a]	−7.95 (−10.40 to −5.49)[a]	−8.72 (−11.61 to −5.83)[a]
Low consultation and health care use	0	0	0
High consultation and health care use	−7.26 (−9.01 to −5.51)[a]	−12.43 (−14.56 to −10.31)[a]	−11.82 (−14.24 to −9.40)[a]
Low consultation but high analgesia	−5.79 (−7.24 to −4.35)[a]	−8.68 (−10.39 to −6.97)[a]	−8.78 (−10.76 to 6.79)[a]
Low consultation but moderate health care use	−2.73 (−4.03 to −1.43)[a]	−4.66 (−6.28 to −3.04)[a]	−3.82 (−5.67 to −1.98)[a]

* Model was controlled for sex, age, body mass index, comorbidity count, and index of multiple deprivation. CI, confidence interval; MSK-HQ, Musculoskeletal Health Questionnaire.
[a] These are significant results.

health care patterns and the patients' self-reported MSK-HQ outcomes. Using SA, we identified five distinct clusters that differed in terms of MSK-related pain consultations, analgesic prescriptions, imaging, physiotherapy, and secondary care referrals. The data tell us that the low consultation and health care use group has the best MSK health. Factors associated with being in the other clusters and poorer health are sex, age, BMI, comorbidities, and neighborhood deprivation.

To our knowledge, this is the first study to use SA methodology to uncover health care patterns of MSK conditions in primary care using routinely collected EHR data. A Canadian study by Nguena Nguefack et al[6] used SA to identify five two-year care trajectories among patients living with arthritic conditions. However, their focus was on patterns of health care visits across different health care services (eg, emergency department visits, hospitalizations, and pain clinics) without considering multiple treatment strategies. This may be due to variations in the health care systems, which may influence the applicability of different primary care approaches. Similarly, the study by Mose et al[5] employed latent class growth analysis to identify five 10-year patterns of MSK health care utilization among adult Danes who reported chronic MSK pain. Although they modeled the number of health care contacts, they did not analyze the sequence of services used. Our findings have similarities with trajectories from studies analyzing single components of health care. However, in contrast, our study examined jointly all the main components of MSK management in primary care settings. Additionally, Meisingset et al[33] identified five distinct MSK phenotypes using latent class analysis, but their focus was on key prognostic factors over the biopsychosocial domains across common MSK pain. Although these phenotypes may support the development of targeted interventions, our study, which integrates different care strategies for MSK pain in primary care, offers practical insights that may enhance clinical practice and inform decision-making in primary care settings.

This study demonstrated that patients in the "high consultation and health care use" group experienced the worst outcome in terms of MSK-HQ score. This finding aligns with the results of the study by Nguena Nguefack et al,[6] which indicated that belonging to a high health care utilization group was associated with a higher likelihood of perceiving a poor or fair quality of life.[6] This high-utilization group in our study represented 8.32% of consulters with MSK pain and predominantly consisted of female participants, older patients, individuals with obesity, and those coming from the most deprived areas. Additionally, this group had the highest proportion of patients with a comorbidity count of three or above, suggesting a significant burden of comorbidities.[6]

In contrast, patients in the "low consultation and health care use" group exhibited the best MSK health (highest MSK-HQ score). This was the largest group, comprising 57.39% of consulters with MSK pain, and included a higher proportion of male participants and younger patients, fewer individuals with obesity, and a greater proportion of patients with no comorbidities. Notably, 389 patients (36.15%) in this group had consultations only in year 1, suggesting they might be incident consulters. Furthermore, individuals from the least deprived areas typically use health care services less frequently than those from the most deprived areas, a finding consistent with other studies reporting socioeconomic differences in the prevalence and management of chronic pain.[34] These results indicate that more sophisticated SA nevertheless confirms the general observation made in previous studies of a subset of patients with high levels of pain and disability and high health care use, in which issues of quality and effectiveness of care may be more important than simple lack of access to primary care.

By evaluating data from the five years before the index consultation, we gained insights into the longitudinal treatment strategies experienced by patients. This helps health care providers learn from previous cases, refining treatment guidelines and care strategies based on actual outcomes. Furthermore, our approach helps identify patient groups that require more intensive and tailored care, allowing for a more effective allocation of resources to where they are needed most. Our findings reveal that nearly half of the patients consulting for MSK conditions have a long history of health care interactions, which is associated with poorer short-term outcome. These patients typically come from socioeconomically deprived areas, are predominantly older and female, and have obesity and multiple comorbidities. Our assessment of patients' profiles and outcome variations among health care utilization patterns can be used to improve care pathways and highlights areas in which policy interventions could substantially enhance health equity.

The strength of this study lies in its innovative multidimensional approach to SA, enabling a comprehensive exploration of the most shared health care utilization patterns for MSK conditions in primary care, considering patterns of consultations, analgesic prescriptions, imaging, physiotherapy, and secondary care referrals. There are potential limitations in this study. The inclusion of only those patients who consented to participate might have introduced a selection bias, as evidenced by the poor response rate. Additionally, the IMD data suggest that the sample was less deprived compared to the general population. Consequently, the patterns of health care identified here, and their relative frequency, may not reflect those in the target population of all adult consulters with MSK pain. In particular, the frequency of low consultation and health care use may be overestimated in our sample, given indirect evidence of lower study participation among more deprived patients. Moreover, our analysis was based on continuous retrospective records of five years before the index consultation. The registration period of the patients was not available in the data, so we were not sure whether the patients with missing health care events had no recorded events or were not registered during that period. We checked whether the patients had five years of continuous records by computing the difference

between the index date and the date of the first recorded event. We found 738 patients had less than five years of continuous retrospective record. Excluding these patients does reduce the sample size. Therefore, we conducted a sensitivity analysis excluding these patients, and the full results are provided in the supplementary file (Supplementary Figures S8–S11 and Supplementary Tables S4–S9).

Optimizing primary care and linkage to effective approaches is crucial for reducing the effect of MSK conditions. Understanding the patterns of patients' journeys through various health care services contributes to the achievement of this goal. SA could serve as a feasible method for identifying patient interactions with the health care system by delineating sequences of care events and identifying distinct health care utilization patterns. This study offers initial insights into patterns of health care by consulters with MSK pain to primary care, which have been directed by clinicians. Further investigations are warranted to gain a deeper understanding of care patterns for MSK conditions in primary and secondary care settings and focus on specific MSK subpopulations such as OA and LBP.

In conclusion, this study identified five distinct health care patterns among adult patients with MSK conditions using SA. Patients' self-reported outcomes and sociodemographic profiles varied across the five clusters. Patients with high health care utilization reported poorer outcomes, whereas those with lower utilization had better outcomes. These findings underscore the association among socioeconomic status, extensive health care utilization, and poorer health outcome, emphasizing the need for targeted policy interventions to improve health equity and quality of care.

ACKNOWLEDGMENTS

This study is based on data from MIDAS-GP study. We are extremely grateful for the involvement and contributions of our Patient Advisory Group to this study. We thank James Bailey for his assistance in retrieving data from EHRs. We wish to acknowledge the contributions of Simon Wathall, Gerri Mulcahy, and members of the National Institute for Health and Care Research Clinical Research Network: West Midlands. Practice managers and staff at participating practices; staff at MJog by Livi; and Sarah Lawton, Steff Garvin, Clare Thompson, Jo Smith, Sarah Lewis, Rachael Heath, Jacqui Carter, and the administration support staff in Keele Clinical Trials Unit contributed to the design and implementation of practice-based patient recruitment methods for MIDAS-GP.

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 Yu 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. Versus Arthritis. The State of Musculoskeletal Health 2024. 2024. Accessed August 2, 2024. https://www.versusarthritis.org/media/25649/versus-arthritis-state-msk-musculoskeletal-health-2023.pdf
2. Yu D, Missen M, Jordan KP, et al. Trends in the annual consultation incidence and prevalence of low back pain and osteoarthritis in England from 2000 to 2019: comparative estimates from two clinical practice databases. Clin Epidemiol 2022;14:179–189.
3. Welsh VK, Mason KJ, Bailey J, et al. Trends in consultations and prescribing for rheumatic and musculoskeletal diseases: an electronic primary care records study. Br J Gen Pract 2023;73(736):e858–e866.
4. Babatunde OO, Bishop A, Cottrell E, et al. A systematic review and evidence synthesis of non-medical triage, self-referral and direct access services for patients with musculoskeletal pain. PLoS One 2020;15(7):e0235364.
5. Mose S, Kent P, Smith A, et al. Trajectories of musculoskeletal healthcare utilization of people with chronic musculoskeletal pain – a population-based cohort study. Clin Epidemiol 2021;13:825–843.
6. Nguena Nguefack HL, Pagé MG, Choinière M, et al. Distinct care trajectories among persons living with arthritic conditions: a two-year state sequence analysis. Front Pain Res (Lausanne) 2022;3:1014793.
7. Flothow A, Novelli A, Sundmacher L. Analytical methods for identifying sequences of utilization in health data: a scoping review. BMC Med Res Methodol 2023;23(1):212.
8. Husselbee R, Price J. Implementing and evaluating patient reported outcome measures (MSK-HQ) using electronic patient records in musculoskeletal practice: analysis of over 11,000 records. Physiotherapy 2022;114:e94–e95.
9. Hill JC, Thomas E, Hill S, et al. Development and validation of the Keele musculoskeletal patient reported outcome measure (MSK-PROM). PLoS One 2015;10(4):e0124557.
10. Kinge JM, Knudsen AK, Skirbekk V, et al. Musculoskeletal disorders in Norway: prevalence of chronicity and use of primary and specialist health care services. BMC Musculoskelet Disord 2015;16(1):75.
11. Häuser W, Wolfe F, Henningsen P, et al. Untying chronic pain: prevalence and societal burden of chronic pain stages in the general population - a cross-sectional survey. BMC Public Health 2014;14(1):352.
12. Emilson C, Åsenlöf P, Demmelmaier I, et al. Association between health care utilization and musculoskeletal pain. A 21-year follow-up of a population cohort. *Scand* J Pain 2020;20(3):533–543.
13. Oppong R, Lewis M, Campbell P, et al. Comparison of health-care utilization, costs and health-related quality of life across the subgroups defined by the Keele STarT MSK Tool. Rheumatology (Oxford) 2023;62(6):2076–2082.
14. Mathew S, Peat G, Parry E, et al. Applying sequence analysis to uncover 'real-world' clinical pathways from routinely collected data: a systematic review. J Clin Epidemiol 2024;166:111226.
15. Schwanitz K. The transition to adulthood and pathways out of the parental home: a cross-national analysis. Adv Life Course Res 2017;32:21–34.
16. Lorentzen T, Bäckman O, Ilmakunnas I, et al. Pathways to adulthood: sequences in the school-to-work transition in Finland, Norway and Sweden. Soc Indic Res 2019;141(3):1285–1305.
17. Zhou Y. Work trajectories and status attainment process: a study using sequence analysis. J Chin Sociol 2023;10(1):1.
18. Brodeur S, Vanasse A, Courteau J, et al. Antipsychotic utilization trajectories three years after initiating or reinitiating treatment of schizophrenia: a state sequence analysis approach. Acta Psychiatr Scand 2022;145(5):469–480.

19. Vanasse A, Courteau J, Courteau M, et al. Multidimensional analysis of adult patients' care trajectories before a first diagnosis of schizophrenia. Schizophrenia (Heidelb) 2022;8(1):52.
20. Vanasse A, Courteau J, Courteau M, et al. Healthcare utilization after a first hospitalization for COPD: a new approach of State Sequence Analysis based on the "6W" multidimensional model of care trajectories. BMC Health Serv Res 2020;20(1):177.
21. Henri S, Herrera R, Vanasse A, et al. Trajectories of care in patients with chronic obstructive pulmonary disease: a sequence analysis. Can J Respir Crit Care Sleep Med 2022;6(4):237–247.
22. Le Meur N, Vigneau C, Lefort M, et al. Categorical state sequence analysis and regression tree to identify determinants of care trajectory in chronic disease: example of end-stage renal disease. Stat Methods Med Res 2019;28(6):1731–1740.
23. Abbott A, Tsay A. Sequence analysis and optimal matching methods in sociology. Sociol Methods Res 2000;29(1):3–33.
24. Liao TF, Bolano D, Brzinsky-Fay C, et al. Sequence analysis: its past, present, and future. Soc Sci Res 2022;107:102772.
25. National Institute for Health and Care Excellence. Multiple long-term conditions: multimorbidity register. July 31, 2019. Accessed December 11, 2024. https://www.nice.org.uk/indicators/ind205-multiple-long-term-conditions-multimorbidity-register
26. Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40(5):373–383.
27. Elixhauser A, Steiner C, Harris DR, et al. Comorbidity measures for use with administrative data. Med Care 1998;36(1):8–27.
28. Wray NP, Hollingsworth JC, Peterson NJ, et al. Case-mix adjustment using administrative databases: a paradigm to guide future research. Med Care Res Rev 1997;54(3):326–356.
29. Ministry of Housing, Communities & Local Government. The English Indices of Deprivation 2019 (IoD2019). September 26, 2019. Accessed July 12, 2024. https://assets.publishing.service.gov.uk/media/5d8e26f6ed915d5570c6cc55/IoD2019_Statistical_Release.pdf
30. Hill JC, Kang S, Benedetto E, et al. Development and initial cohort validation of the arthritis research UK Musculoskeletal Health Questionnaire (MSK-HQ) for use across musculoskeletal care pathways. BMJ Open 2016;6(8):e012331.
31. Ritschard G, Liao TF, Struffolino E. Strategies for multidomain sequence analysis in social research. Sociol Methodol. 2023;53(2):288–322.
32. Azur MJ, Stuart EA, Frangakis C, et al. Multiple imputation by chained equations: what is it and how does it work? Int J Methods Psychiatr Res 2011;20(1):40–49.
33. Meisingset I, Vasseljen O, Vøllestad NK, et al. Novel approach towards musculoskeletal phenotypes. Eur J Pain 2020;24(5):921–932.
34. Lynch M, Peat G, Jordan K, et al. Where does it hurt? Small area estimates and inequality in the prevalence of chronic pain. Eur J Pain 2023;27(10):1177–1186.

Arthritis Care & Research
Vol. 77, No. 7, July 2025, pp 916–927
DOI 10.1002/acr.25502

American College of Rheumatology
Empowering Rheumatology Professionals

Pregnancy Outcomes of Targeted Synthetic Disease-Modifying Antirheumatic Drugs Among Patients With Autoimmune Diseases: A Scoping Review

Vienna Cheng,[1] Neda Amiri,[2] Vicki Cheng,[3] Ursula Ellis,[4] Jacquelyn J. Cragg,[5] Mark Harrison,[6] Laurie Proulx,[7] and Mary A. De Vera[6]

Objective. Targeted synthetic disease-modifying antirheumatic drugs (tsDMARDs) have expanded the management of autoimmune diseases, including rheumatic diseases. As the use of these drugs grows, it is important to understand their effects on pregnancy. We conducted a scoping review to synthesize the current evidence on the impacts of tsDMARDs on pregnancy outcomes.

Methods. We searched the Embase, MEDLINE, and CENTRAL databases in November 2023. We included studies that examined tsDMARD exposure for chronic autoimmune disease(s), particularly in mothers during pregnancy, fathers before conception, and/or fetuses/neonates in utero. We extracted data on sample size, study design, tsDMARD exposure (dose and duration), and reproductive health outcomes.

Results. Of 6,712 studies screened, eight were included, namely nine case reports, one case series, four cross-sectional studies, and one cohort study among patients with ulcerative colitis, rheumatoid arthritis, and psoriasis. Sample sizes ranged from 1 to 116 pregnancies or offspring, with six studies on tofacitinib, one on baricitinib, one on upadacitinib, and no studies on apremilast. Overall, 19 fetal/neonatal outcomes, six fetal/neonatal-maternal outcomes, and three maternal outcomes were extracted. The most frequently reported fetal/neonatal outcomes were congenital anomaly (n = 4), preterm birth (n = 4), and the fetal/neonatal-maternal outcome of spontaneous abortion (n = 4). Only one study reported on the maternal outcome of delivery via Cesarean section.

Conclusion. Our scoping review of evidence to date on the perinatal use of tsDMARDs reveal small sample sizes and a limited number of studies, all largely descriptive in nature. Findings highlight evidence gaps that preclude providers and patients from making informed decisions when considering the perinatal use of tsDMARDs.

INTRODUCTION

Chronic autoimmune diseases, such as rheumatic disease and inflammatory bowel disease, are associated with significant morbidity.[1,2] Within these, rheumatoid arthritis and spondyloarthritis affect 0.5% to 1%[1] and 0.5% to 2%[3] of the global population, respectively. These conditions often strike during the childbearing years, with rheumatoid arthritis disproportionately impacting more women than men.[2,4,5] Uncontrolled autoimmune disease during pregnancy is not only associated with reduced fertility[6,7] but also adverse maternal outcomes (eg, pre-eclampsia[8] and gestational diabetes[9,10]) and neonatal outcomes (eg, preterm birth[11–13] and congenital anomalies[12,14]). Treatment with disease-modifying antirheumatic drugs

Supported by The Arthritis Society Canada to the Following Maternal, Neonatal, and Childhood Outcomes Associated With Use of Arthritis Medications Perinatally (FAMILY) study (grant SOG-22-192). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

[1]Vienna Cheng, PharmD: University of British Columbia, Collaboration for Outcomes Research and Evaluation, and Arthritis Research Canada, Vancouver, British Columbia, Canada; [2]Neda Amiri, MD: University of British Columbia, Arthritis Research Canada, and Mary Pack Arthritis Centre, Vancouver, British Columbia, Canada; [3]Vicki Cheng, PharmD, MSc: University of British Columbia and Collaboration for Outcomes Research and Evaluation, Vancouver, British Columbia, Canada; [4]Ursula Ellis, MLIS: University of British Columbia, Vancouver, British Columbia, Canada; [5]Jacquelyn J. Cragg, PhD: University of British Columbia, Collaboration for Outcomes Research and Evaluation, and International Collaboration on Repair Discoveries, Vancouver, British Columbia, Canada; [6]Mark Harrison, PhD, Mary A. De Vera, PhD: University of British Columbia, Collaboration for Outcomes Research and Evaluation, Arthritis Research Canada, and Centre for Health Evaluation and Outcome Sciences, Vancouver, British Columbia, Canada; [7]Laurie Proulx, BCom(Hons): Canadian Arthritis Patient Alliance, Ottawa, Ontario, Canada.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25502).

Author disclosures and graphical abstract are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25502.

Address correspondence via email to Mary A. De Vera, PhD, at mdevera@arthritisresearch.ca.

Submitted for publication October 11, 2024; accepted in revised form January 14, 2025.

SIGNIFICANCE & INNOVATIONS

- To our knowledge, this is the first synthesis and characterization of current evidence on targeted synthetic disease-modifying antirheumatic drugs (tsDMARDs) and pregnancy outcomes. Overall, we identified a limited number of studies on each tsDMARD, with a total of eight included publications (tofacitinib, n = 6; baricitinib, n = 1; and upadacitinib, n = 1). Notably, no studies on apremilast were found, and two included publications were abstracts. This speaks to the limited published information on these emerging drugs to date.
- Included studies were largely descriptive in nature, highlighting a notable lack of analytical evidence on the effects of tsDMARDs on human pregnancy (two case reports, one case series, and four cross-sectional studies, as opposed to one cohort study).
- Overall, we found inconsistent reporting of sample size units, exposure definitions, and pregnancy outcomes. Although most reported outcomes were related to the fetus/neonate, only two studies reported on maternal outcomes (eg, gestational diabetes), and no studies reported on paternal outcomes (eg, fertility or disease activity).
- Our study highlights the need for more analytical research on tsDMARD exposure and pregnancy outcomes. Findings from this synthesis have implications for informing our understanding of the existing perinatal research to date on tsDMARDs and highlighting gaps to be addressed in future research.

(DMARDs) is often necessary throughout the perinatal period to maintain low disease activity and minimize maternal-fetal risk.[15]

Despite the proven efficacy of conventional synthetic DMARDs for decades, many patients with moderate to severe rheumatic disease remain suboptimally controlled unless treated with a biologic or targeted synthetic (ts)DMARD.[15,16] Specifically, owing to their unique oral formulation and efficacy in targeting specific intracellular signaling pathways of the immune system,[17–19] the advent of tsDMARDs in the recent decade has significantly expanded the treatment options of autoimmune diseases.[20–23] There are four tsDMARDs approved by the US Food and Drug Administration for various rheumatic and inflammatory bowel diseases: tofacitinib, baricitinib, upadacitinib, and apremilast[24,25] (Supplementary Table 1). The Janus kinase (JAK) inhibitors (tofacitinib, baricitinib, and upadacitinib) are indicated for rheumatoid arthritis, ankylosing spondylitis, ulcerative colitis, and psoriatic arthritis.[26] Apremilast is a phosphodiesterase-4 inhibitor indicated for adult psoriatic arthritis and plaque psoriasis.[27] Although evidence has emerged on the perinatal impacts of conventional synthetic DMARDs and biologic originator DMARDs over the past decade,[28,29] data on newer therapies, particularly tsDMARDs, are scarce.

In 2016, the European Alliance of Associations for Rheumatology (EULAR) defined points to consider for the use of antirheumatic drugs before and during pregnancy.[30] Regarding tsDMARDs, EULAR indicates that tofacitinib should be discontinued 2 months before conception because of insufficient data concerning safety during pregnancy. There were no points for baricitinib, upadacitinib, or apremilast until the 2024 EULAR update—to which insufficient data was also stated.[31] The American College of Rheumatology (ACR) published guidelines in 2020 indicating they were unable to assess the pregnancy compatibility of tofacitinib, baricitinib, and apremilast due to insufficient data.[32] ACR also states that no recommendations can be made for men planning to father a child while taking tofacitinib, baricitinib, and apremilast due to insufficient data.[32] Indeed, preclinical animal studies have demonstrated teratogenicity, and given their small molecular size (unlike monoclonal antibodies), which allows transplacental passage,[17,22,23,33–35] there are rising calls for research evaluating tsDMARD use during human pregnancy and associated pregnancy outcomes. The lack of data may negatively impact patients and babies, as treatment plans are often modified before conception to medications compatible with pregnancy. With this, it is important to synthesize the research to date to identify knowledge gaps and inform much-needed subsequent research. Therefore, we conducted a scoping review as this is an appropriate approach to achieve our objective of mapping and synthesizing the current evidence on the reproductive health impact of tsDMARDs on fetal/neonatal, maternal, and paternal outcomes in individuals living with chronic autoimmune diseases (eg, rheumatoid arthritis, psoriasis, or inflammatory bowel disease).

METHODS

Search strategy. We conducted a scoping review following the Arksey and O'Malley framework: (1) developed the research question, (2) identified relevant studies (search electronic databases, check reference lists of relevant studies and systematic reviews, hand search, and citation mining), (3) included studies matching our inclusion criteria, (4) extracted and charted key data points, and (5) organized, mapped, and reported our results.[36] We also followed the Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews reporting checklist.[37,38] Our rationale for choosing to conduct a scoping review is to map the evidence, identify gaps in literature, and gain a broad understanding of evidence on the reproductive health impact of tsDMARDs.

We developed our search strategy with a research librarian (UE) who applied it to Embase (Ovid), Cochrane CENTRAL (Ovid), and MEDLINE (Ovid) from inception until November 30, 2023 (Supplementary Table 2). Because multiple DMARDs are likely to be examined within the same study, our search strategy remained broad to encompass all DMARDs to ensure we

captured all possible publications with tsDMARD exposure. The search was limited to publications in English, French, German, and Korean due to the time and resource limitations of the team. Covidence software was used to deduplicate search results.[39]

Study screening and inclusion criteria. The search strategy was designed based on our inclusion criteria, which were categorized into the Population-Concept-Context outline described by the Joanna Briggs Institute for scoping reviews.[40] The inclusion criteria included original, peer-reviewed observational studies (Context) that examined exposure to tsDMARDs (Supplementary Table 1) (Concept) in (1) mothers during pregnancy, (2) fathers before conception, and/or (3) fetuses/neonates in utero among parents with chronic autoimmune disease(s) (eg, rheumatoid arthritis, psoriasis, inflammatory bowel disease) (Population). Reviews, treatment guidelines, commentaries, and opinion pieces were excluded. If newer publication(s) for the same study were identified, the earlier publication(s) were excluded unless they presented information that was not reported in the most recent publication. All studies identified were screened for eligibility initially by title and abstract, then by full-text review to determine final inclusion (Vienna Cheng, Vicki Cheng, and MADV). Any uncertainties were resolved through discussion to achieve consensus (Vienna Cheng, Vicki Cheng, and MADV).

Data extraction and patient and public involvement. We extracted available data on publication year, country, study sample autoimmune condition(s) (eg, rheumatoid arthritis, psoriasis, inflammatory bowel disease), tsDMARD exposure, sample size, study design (eg, case reports, cohort studies, cross-sectional studies), publication type (eg, abstract or full manuscript), and data source (eg, medical records, administrative data, pregnancy registry, manufacturer safety database). Of interest, we extracted information on tsDMARD exposure in terms of medicinal ingredient, timing (ie, timing of last dose in gestational weeks), duration (ie, total length of exposure during pregnancy), and maternal and/or paternal exposure. Finally, outcomes were extracted and characterized according to the person(s) experiencing the outcome (fetus/neonate, mother, or father) and the timing of the outcomes (before pregnancy, during pregnancy, intrauterine, at delivery, or after delivery). All data were entered into a data extraction form developed on Microsoft Excel. Our patient research partner (LP), who has lived experience with rheumatoid arthritis and multiple pregnancies, contributed patient perspectives that highlighted the priorities and experiences of patients during the review of the manuscript, discussion and interpretation of results, and suggestions for future research.

RESULTS

Search results. Of 6,712 studies screened, 8 studies (reported in 6 full manuscripts and 2 abstracts) were eligible for inclusion (Figure 1). Characteristics of the included studies are summarized in Table 1. One case report reported on baricitinib, and six studies, including one case report, one case series, three cross-sectional studies, and one cohort study, reported on tofacitinib. One cross-sectional study reported on upadacitinib, and no studies reported on apremilast. Overall, the cross-sectional studies reported a total of 456 pregnancies from both maternal and paternal tsDMARD exposures. The cohort study published as an abstract reported on four offspring of mothers exposed to tsDMARDs during pregnancy.

Reproductive health outcomes reporting framework. Altogether, we extracted 28 reproductive health outcomes reported in the included studies. Informed by our prior work on synthesizing reproductive health outcomes,[41] we organized outcomes reported in the included studies according to who experienced the outcome (fetus/neonate, fetus/neonate and mother, mother, or father), when the outcome was assessed or occurred (before pregnancy, during pregnancy, intrauterine phase, at delivery, and after delivery), and what the specific outcome was. This resulted in a reproductive health outcomes reporting framework as follows: (1) fetal/neonatal outcomes assessed intrauterine or after delivery, (2) fetal/neonatal-maternal outcomes occurring during pregnancy or at delivery, (3) maternal outcomes occurring before or during pregnancy, and (4) paternal outcomes occurring before pregnancy (eg, fertility[42], disease activity[43]). Figure 2 depicts this framework, which includes an additional feature of flexibility with gray boxes indicating outcomes not reported in our included studies, but could be incorporated in future perinatal studies. These gray boxes serve as placeholders to encourage and guide future expansions to the framework as additional research becomes available. Colors in the figure were used to distinguish different elements of the framework, guided by the principles of accessible visual design.

Synthesis of included studies. *Case reports and case series.* Two case reports described maternal exposure during pregnancy to baricitinib and tofacitinib, respectively. One case series reported on six cases of maternal exposure to tofacitinib (Tables 2 and 3).

Costanzo et al reported on a 43-year-old woman in Italy with rheumatoid arthritis who took baricitinib from conception until 17 weeks' gestation, with normal fetal growth and no abnormalities detected. A live baby at normal weight (3,200 g) and length (50 cm) was delivered at 38 weeks' gestation with no perinatal infections.[44] The authors did not report on the presence of congenital anomaly. The mother was observed to be in clinical remission during pregnancy.

Fernández-Sánchez et al reported on a 40-year-old woman in Spain with psoriatic arthritis who took tofacitinib from conception until discontinuation at 6 weeks' gestation, with normal fetal

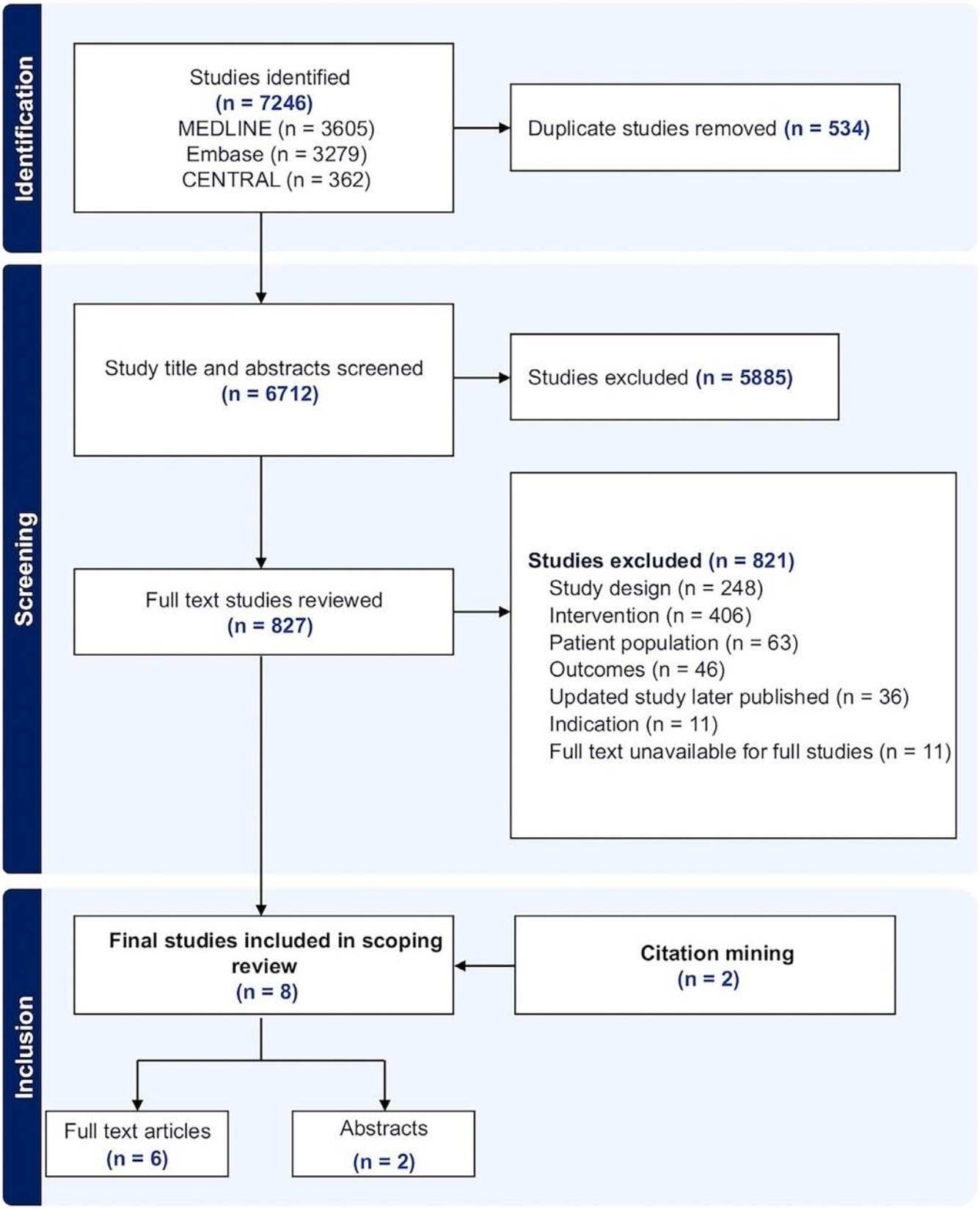

Figure 1. Preferred Reporting Items for Systematic reviews and Meta-Analyses flow diagram. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25502/abstract.

Table 1. Characteristics of all included studies

Study	Country	Condition(s)	Exposure	Sample size	Study design	Publication type	Data source(s)
Baricitinib (n = 1 study)							
Costanzo et al, 2020[44]	Italy	Rheumatoid arthritis	Maternal	1	Case report	Full manuscript	Medical records
Tofacitinib (n = 6 studies)							
Clowse et al, 2016[35]	USA	Psoriasis and rheumatoid arthritis	Maternal Paternal	13[a] 34[b] 44[c,d]	Cross-sectional	Full manuscript	Pregnancy registry and manufacturer safety database
Mahadevan et al, 2018[33]	USA	Psoriasis, psoriatic arthritis, rheumatoid arthritis, ulcerative colitis, and unspecified[e]	Maternal Paternal	116[c] 87[c,d]	Cross-sectional	Full manuscript	Pregnancy registry and manufacturer safety database
Vinet et al, 2019[48]	Canada	Unspecified[e]	Maternal	4[f]	Cohort study	Abstract	Administrative data
Mahadevan et al, 2020[47]	USA	Ulcerative colitis	Maternal Paternal	15[c] 19[c,d]	Cross-sectional	Abstract	Manufacturer safety database
Fernández-Sánchez et al, 2021[45]	Spain	Psoriatic arthritis	Maternal	1	Case report	Full manuscript	Medical records
Mitrova et al, 2025[46]	Czech Republic	Ulcerative colitis	Maternal	6	Case series	Full manuscript	Medical records
Upadacitinib (n = 1 study)							
Mahadevan et al, 2024[34]	USA	Atopic dermatitis, Crohn disease, psoriatic arthritis, rheumatoid arthritis, nonradiographic axial spondyloarthritis, ulcerative colitis, and unspecified[e]	Maternal	21[g] 59[h] 48[d,i]	Cross-sectional	Full manuscript	Manufacturer safety database

[a] Number of pregnancies that received tofacitinib plus methotrexate combination therapy.
[b] Number of pregnancies that received tofacitinib monotherapy.
[c] Number of exposed pregnancies.
[d] Drug exposure occurred around the time of conception, in the first trimester, and/or undetermined in relation to the date of conception.
[e] Authors did not specify which condition the mother or father exposed to targeted synthetic disease-modifying antirheumatic drugs were being treated for.
[f] Number of exposed offspring.
[g] Number of pregnancies that received upadacitinib plus methotrexate combination therapy.
[h] Number of pregnancies that received upadacitinib monotherapy.
[i] Number of exposed 'cases'.

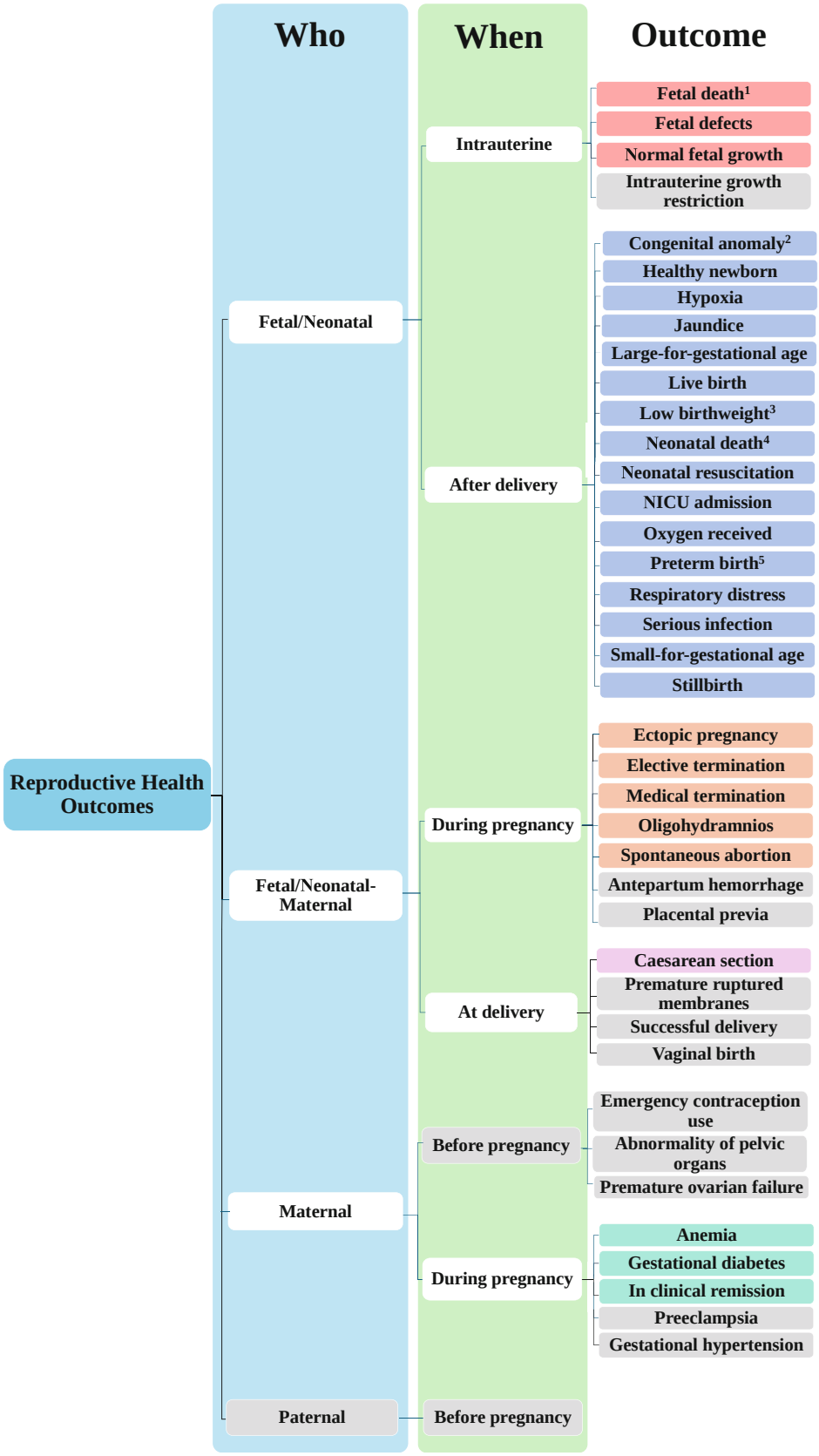

Figure 2. Reproductive health outcomes reporting framework.[1] Death past 20 weeks' gestation.[2] Includes major and minor malformations and congenital disorders.[3] Birth weight <2,500 g (5.5 lbs).[4] Death among live births within the first 28 days of life.[5] Live birth ≤37 weeks' gestation. NICU, neonatal intensive care unit.

Table 2. Exposure data of case reports and case series

Study	Mother/ baby, n	Exposure				
		Drug	Exposure	Total length of exposure, weeks' gestation	Timing of last dose, weeks' gestation	Timing of delivery, weeks' gestation
Costanzo et al, 2020[44]	1	Baricitinib	Maternal	17	17	38
Fernández-Sánchez et al, 2021[45]	1	Tofacitinib	Maternal	6	6	36
Mitrova et al, 2025[46]	1[a]	Tofacitinib	Maternal	9	9	–
	2			7	7	37
	3			7	7	37
	4			40	–	40
	5			38	–	38
	6			38	–	38

[a] Elective termination of pregnancy during the first trimester.

growth and no fetal defects detected. A live preterm baby (2,515 g weight and 48 cm length) was delivered at 36 weeks' gestation via an emergent Cesarean section owing to a preterm membrane rupture, with no signs of congenital malformations or dysfunctions.[45] The study did not report whether further neonatal outcomes were observed, such as infection. The mother was observed to be in clinical remission during pregnancy.

Lastly, Mitrova et al presented a case series on six pregnancies from six women with ulcerative colitis in the United States, three of whom were exposed to tofacitinib throughout pregnancy and three during the first trimester.[46] Of the six mothers, three were not in clinical remission during pregnancy. No intrauterine fetal outcomes were reported. One pregnancy was electively terminated at 9 weeks' gestation owing to safety concerns from the patient. The remaining five pregnancies all resulted in live, healthy, full-term newborns with no congenital anomalies or identification of small for gestational age (median birth weight 2,970 g [range 2,270–3,200 g]). One infant was born with low birth weight (2,270 g) via planned Cesarean section to a mother who was not in disease remission throughout pregnancy. The other infant delivered via planned Cesarean section was born to a mother in remission throughout pregnancy; however, this infant was diagnosed with jaundice and hypoxia requiring oxygen after birth. In terms of perinatal complications, one other infant was diagnosed with jaundice and hypoxia requiring oxygen after birth; this infant was also born to a mother in remission throughout pregnancy. Of note, maternal outcomes such as presence of gestational diabetes, preeclampsia, or gestational hypertension were not reported in any of the case reports or case series.

Cross-sectional studies and cohort study. Reported reproductive health outcomes from four included cross-sectional studies and one cohort study are illustrated in Figure 3. Altogether, there were 14 fetal/neonatal outcomes (n = 5 studies), six fetal/neonatal-maternal outcomes (n = 4 studies), three maternal outcomes (n = 2 studies[33,34]), and no paternal outcomes reported (eg, fertility[42] or disease activity[43]). Overall, the most frequently reported fetal/neonatal outcomes were congenital anomaly (n = 4 studies) and preterm birth (n = 4 studies). The main fetal/

Table 3. Observed outcomes of case reports and case series*

Study	Mother/ baby, n	Observed outcomes							
		Fetal/neonatal					Fetal/neonatal-maternal		Maternal during pregnancy
		Intrauterine		After delivery			At delivery		
		Fetal defect	Normal fetal growth	Congenital anomaly	Live birth	Preterm birth	Cesarean section	Preterm membrane rupture	In clinical remission
Costanzo et al, 2020[44]	1	0	1	–	1	0	1	–	1
Fernández-Sánchez et al, 2021[45]	1	0	1	0	1	1	1	1	1
Mitrova et al, 2025[46]	1[a]	–	–	–	0	–	–	–	1
	2	–	–	0	1	0	1	–	0
	3	–	–	0	1	0	0	–	1
	4	–	–	0	1	0	0	–	0
	5	–	–	0	1	0	0	–	0
	6	–	–	0	1	0	1	–	1

* A dash indicates that this study did not report this outcome or it was not applicable. 0, outcome not observed; 1, outcome observed.
[a] Elective termination of pregnancy during the first trimester.

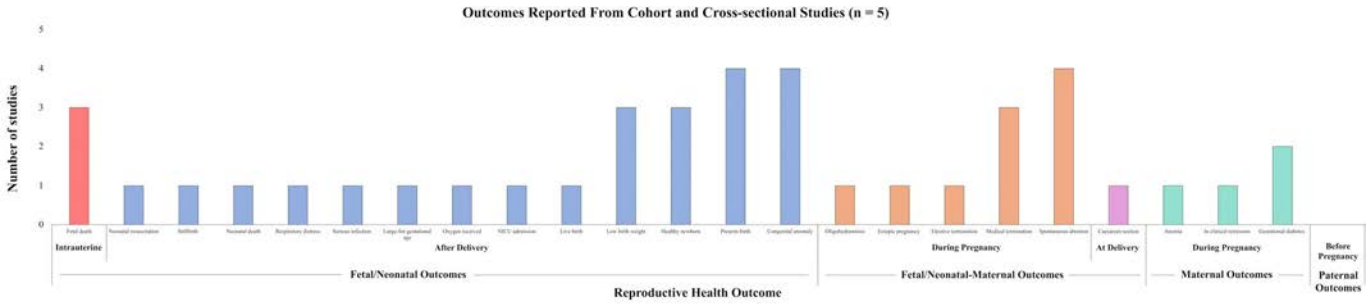

Figure 3. Reproductive health outcomes reported from cohort and cross-sectional studies. NICU, neonatal intensive care unit.

neonatal-maternal outcomes reported were spontaneous abortion (n = 4 studies) and medical termination (n = 3 studies). All cross-sectional studies included cases pending or lost to follow-up.

The 2016 cross-sectional study by Clowse et al was based on a pregnancy registry and manufacturer safety databases studying maternal (n = 13 pregnancies exposed to tofacitinib and methotrexate therapy, and n = 34 pregnancies exposed to tofacitinib monotherapy) and paternal (n = 80 pregnancies) tofacitinib exposure during pregnancy.[35] Among women with rheumatoid arthritis receiving tofacitinib monotherapy (n = 18 pregnancies), authors reported on the following fetal/neonatal outcomes: healthy newborn (n = 11, 61.1%), preterm or low birth weight newborn (n = 2, 11.1%), congenital malformation (pulmonary valve stenosis, approximately 35 days of in utero tofacitinib exposure, and delivered at 38 weeks' gestation) (n = 1, 5.6%), and fetal death (n = 0). Fetal/neonatal-maternal outcomes reported include spontaneous abortion (n = 3, 16.7%) and medical termination (n = 1, 5.6%). Among women with rheumatoid arthritis receiving tofacitinib and methotrexate combination therapy (n = 13 pregnancies), reported fetal/neonatal outcomes included healthy newborn (n = 5, 38.5%), no congenital malformations, and no fetal deaths.[35] Lastly, in terms of paternal drug exposure (n = 44 pregnancies) that occurred approximately at conception or the first trimester, outcomes reported included healthy newborns (n = 23, 52.2%) and spontaneous abortions (n = 5, 11.4%).

The 2018 cross-sectional study by Mahadevan et al used data from a pregnancy registry and manufacturer safety database on maternal (n = 116 pregnancies) and paternal (n = 87 pregnancies) tofacitinib exposure during pregnancy.[33] Among 116 maternal exposures, six fetal/neonatal outcomes were observed, including delivering a healthy newborn (n = 43, 37.1%), preterm birth (n = 7, 6.0%), congenital malformations (n = 2, 1.7%), low birth weight (n = 1, 0.1%), and no neonatal or fetal deaths.[33] There were two fetal/neonatal-maternal outcomes reported: spontaneous abortion (n = 15, 12.9%) and medical termination (n = 14, 12.1%). This is one of the two included studies that reported the maternal outcome of gestational diabetes (n = 1, 0.9%). Of 87 paternal exposures (within the first trimester or undetermined), five fetal/neonatal outcomes were reported, including healthy newborn (n = 56, 64.4%), preterm birth (n = 2, 2.3%), neonatal death (n = 1, 1.2%) and no congenital malformations or fetal deaths. Fetal/neonatal-maternal outcomes included spontaneous abortion (n = 7, 8.1%) and no medical terminations.

The subsequent 2020 cross-sectional study by Mahadevan et al, published as an abstract, used manufacturer safety databases and reported on maternal (n = 15 pregnancies) and paternal (n = 19 pregnancies) tofacitinib exposure during pregnancy.[47] All who received tofacitinib presented with ulcerative colitis. Among 15 maternal exposures (during the first trimester), the authors reported four fetal/neonatal outcomes: healthy newborn (n = 9, 60.0%), preterm birth (n = 1, 6.7%) and no congenital malformations or fetal deaths (after 20 weeks' gestation). Two fetal/neonatal-maternal outcomes were reported: medical terminations (n = 2, 13.3%) and spontaneous abortions (n = 2, 13.3%). Among 19 paternal exposures (timing of exposure unspecified), the authors reported two fetal/neonatal outcomes: healthy newborn (n = 15, 78.9%) and preterm birth (n = 1, 5.3%). Two fetal/neonatal-maternal outcomes were reported: spontaneous abortions (n = 2, 10.5%) and no medical terminations.

In 2024, Mahadevan et al published another cross-sectional study using manufacturer safety databases on maternal upadacitinib exposure during pregnancy (n = 128 pregnancies with known outcomes).[34] Cases reported were from clinical trial data (n = 80) and postmarket surveillance data (n = 48). Patients presented with atopic dermatitis, Crohn disease, psoriatic arthritis, RA, nonradiographic axial spondyloarthritis, ulcerative colitis, and unspecified. Among 80 pregnancies reported from clinical trial data (n = 59 receiving upadacitinib monotherapy and n = 21 receiving concomitant upadacitinib and methotrexate therapy), fetal/neonatal outcomes reported include live birth (n = 43, 53.8%), stillbirth (n = 0), preterm birth (n = 3, 3.8%), congenital anomaly (n = 1, 1.3%), and low birth weight (n = 1, 1.3%).[34] The newborn with congenital anomaly was born prematurely (34 weeks' gestation) with an atrial septal defect and low birthweight (2,030 g), having experienced in utero upadacitinib monotherapy exposure until 4 weeks and 4 days' gestation. The baby's mother experienced anemia and

oligohydramnios during pregnancy. Fetal/neonatal-maternal outcomes during pregnancy were reported on spontaneous abortion (n = 19, 23.4%), elective termination (n = 17, 21.3%), and ectopic pregnancy (n = 1, 1.3%). This was the only included study that reported on the fetal/neonatal outcome of large for gestational age (n = 1, 2.1%), collected through postmarket surveillance data.

Lastly, the 2019 cohort study by Vinet et al, published as an abstract, used administrative data to report maternal exposure during pregnancy, consisted of four tofacitinib-exposed offspring.[48] Although mothers with inflammatory conditions were studied, the abstract did not report specific conditions for the mothers who received tofacitinib. There was one reported neonatal outcome of serious infection, defined as at least one hospitalization because of infection in the first year of life. No other outcomes were reported. Data extracted from the abstract revealed that among four newborns of mothers with inflammatory disease exposed to tofacitinib, one case of serious infection was observed (25%). Our ability to distinguish and extract further study results specific to tofacitinib was limited mainly due to the abstract form of the publication.

DISCUSSION

Our scoping review synthesizing the impact of tsDMARDs on pregnancy outcomes captured eight studies on tofacitinib (n = 6), baricitinib (n = 1), and upadacitinib (n = 1), with no studies on apremilast. There was substantial variability in the reporting of tsDMARD-exposed sample size unit but, nonetheless, we synthesized a total of 468 exposed pregnancies, mothers, and/or newborns. Studies were largely descriptive in nature and mainly reported on fetal/neonatal outcomes, particularly after delivery, with only one included analytic study published as an abstract. Aside from synthesizing the evidence to date on the reproductive health impacts of tsDMARDs, we also established a flexible framework for reporting reproductive health outcomes to guide much-needed future research and postmarket surveillance reporting.

To our knowledge, this is the first synthesis and characterization of the current evidence on tsDMARDs and pregnancy outcomes, which provides an opportunity to identify gaps in the literature and future directions. One gap we identified is the inconsistent and incomplete reporting of study samples and sample sizes. For example, all cross-sectional studies reported the number of exposed pregnancies, but the number of exposed mothers and fathers were unclear or not reported,[33–35,47] and the cohort study only reported the number of exposed offspring.[48] This is particularly problematic when evaluating reproductive health impacts, because it is important to distinguish exposures among mothers (or fathers, where relevant), which occur as a *single unit*, from pregnancies/deliveries/offspring, which occur as *multiple units*. This variability in reporting of study sample units and sample sizes is reflected in our inability to synthesize information across included studies. To facilitate comparisons and synthesis, including potential meta-analyses, we recommend that future studies ensure comprehensive reporting of study samples, sample sizes, and analytic units (ie, number of mothers, fathers, pregnancies, deliveries, and offspring), particularly for drug-exposed patients.

Throughout our synthesis of the included studies, we also found inconsistent reporting of tsDMARD exposure, particularly in cross-sectional studies. Details regarding the timing of exposure to tsDMARDs (ie, preconception, trimesters during pregnancy, or both) and the duration (eg, days or weeks) were either reported inconsistently or not specified. Additionally, the included cohort study was only available in the form of an abstract, which further limited our ability to extract comprehensive data on exposure. If known, it is important for future studies to specify the timing of tsDMARD exposure relative to before or after conception, as well as the time periods during pregnancy, as this can enable the reporting of outcomes specific to certain exposure timings.[49] For example, the embryonic period, between 14 days to approximately 60 days after conception, is where the highest chance of malformation may occur owing to teratogen exposure.[50] Therefore, in studies of pregnancy outcomes, it is important to ascertain the timing of potential insults between conception and birth. We also identified very few studies on each tsDMARD, with a total of eight included publications (tofacitinib, n = 6; baricitinib, n = 1; and upadacitinib, n = 1). Notably, no studies on apremilast were found, and two of the included publications were abstracts. This speaks to the limited published information on each of these emerging drugs to date. However, we recognize that these drugs were introduced in recent years (the earliest being tofacitinib in 2012[26,51]); therefore, their use during pregnancy continues to be cautioned because of limited evidence. This highlights the value of syntheses such as our present study, because they provide a clearer understanding of where there is limited progress made to date on perinatal tsDMARD research.

With respect to outcomes, many extracted from the included studies were not originally investigated as primary endpoints (ie, maternal anemia, large for gestational age, or gestational diabetes[34]). Critically, we identified inconsistent approaches to the reporting of outcomes. To address this, we developed a reporting framework for reproductive health outcomes (Figure 2). This framework not only supported our synthesis but may also serve as a tool to guide future reporting, such as postmarket surveillance and upcoming research on reproductive health outcomes associated with tsDMARD exposure and other disease-modifying drugs. Of note, there was a particular lack of reporting on outcomes that may be considered important by patients and families themselves, such as maternal outcomes, pregnancy monitoring, and medication changes during pregnancy, which can have impacts on the health of female patients with rheumatic diseases (eg, joint damage or increased pain). These invaluable outcomes can be integrated into future studies through active,

iterative engagement with patient research partners throughout the research process to enrich the reporting of DMARD reproductive health outcomes.

We also synthesized study limitations reported by the authors of the included studies. These limitations included challenges in determining the exact date of conception and gestational age, significant missing data (eg, uncertain tsDMARD exposure timing, pending or lost to follow-up, and unknown pregnancy outcomes), a lack of prospective data, and small sample sizes.[33–35] Authors from the included studies also acknowledged limitations in establishing causal relationships between tsDMARDs and perinatal outcomes because of potential confounders unaccounted for, such as concomitant medications and comorbidities.[35] Synthesizing these limitations of current perinatal studies provides valuable considerations for authors planning future study designs on perinatal research.

Arguably, the most important gap we identified in our scoping review is the persistent knowledge gap on the safety of tsDMARDs during pregnancy, particularly given the very limited analytic studies to date. Currently, tsDMARDs are indicated for various rheumatic and inflammatory bowel diseases, such as rheumatoid arthritis and ulcerative colitis, all of which require continuous pharmacological treatment to delay disease progression, alleviate symptoms, and maintain low disease activity.[1] However, these conditions often present during the childbearing years.[2,4] Uncontrolled rheumatic disease during pregnancy is associated with adverse maternal (eg, pre-eclampsia[8,52]) and neonatal outcomes (eg, prematurity[11,12]). Therefore, ongoing treatment throughout the perinatal period is crucial to maintain low disease activity and reduce maternal-fetal risk.[15]

Understanding the effects of arthritis medications on pregnancy outcomes is essential to ensuring maternal-fetal safety while enabling providers and patients to make informed treatment decisions. Of particular interest to our scoping review was the class of tsDMARDs, because their small molecular size allows them to potentially cross the placenta.[35, 53–55] In preclinical animal studies, tofacitinib demonstrated teratogenic and feticidal effects at significantly greater doses than approved human doses, whereas upadacitinib demonstrated teratogenic effects at doses equivalent to human therapeutic and subtherapeutic doses.[33,46] Although animal studies have revealed adverse outcomes, such as skeletal malformations and reduced fetal birth weight, it is difficult to determine the impact of tsDMARDs on human pregnancy based on animal data alone.[17,22,23,33–35] Given the clear lack of well-designed, analytic evidence on the effect of tsDMARDs during human pregnancy, it is important to consider how clinicians and patients can navigate conversations around taking tsDMARDs during pregnancy in clinical settings. Engaging in early, individualized prepregnancy discussions between clinicians and patients to establish a safe and effective treatment plan has been associated with fewer flares, fewer medication changes during pregnancy, and lower disease activity in the first trimester.[56]

It is important to consider the strengths and limitations of our scoping review. Our search strategy was codesigned with experts in perinatal research and a research librarian. To ensure we captured all potential publications, we included a comprehensive list of search terms relating to fetal/neonatal, fetal/neonatal-maternal, maternal, and paternal outcomes (eg, fertility[42] and disease activity[43]) as well as all possible indications of tsDMARDs and other autoimmune diseases, informed by our previous work.[28] Uncertainties during the study screening process were discussed among the team to achieve consensus. The limitations of our search include the fact that our search found studies in French, German, and Korean but was ultimately limited to including studies with texts available in English.

Since the introduction of tsDMARDs and their increasingly widespread use, it is important to address questions around their impacts on pregnancy outcomes. Findings from this synthesis inform our understanding of the existing research to date and highlight gaps to be addressed in future research.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr De Vera 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. Uhlig T, Moe RH, Kvien TK. The burden of disease in rheumatoid arthritis. PharmacoEconomics 2014;32(9):841–851.
2. Chakravarty EF, Nelson L, Krishnan E. Obstetric hospitalizations in the United States for women with systemic lupus erythematosus and rheumatoid arthritis. Arthritis Rheum 2006;54(3):899–907.
3. Dean LE, Jones GT, MacDonald AG, et al. Global prevalence of ankylosing spondylitis. Rheumatology (Oxford) 2014;53(4):650–657.
4. Kvien TK, Uhlig T, Ødegård S, et al. Epidemiological aspects of rheumatoid arthritis: the sex ratio. Ann N Y Acad Sci 2006;1069(1): 212–222.
5. Boel A, López-Medina C, van der Heijde DMFM, et al. Age at onset in axial spondyloarthritis around the world: data from the Assessment in SpondyloArthritis International Society Peripheral Involvement in Spondyloarthritis study. Rheumatology (Oxford) 2022;61(4):1468–1475.
6. Brouwer J, Fleurbaaij R, Hazes JMW, et al. Subfertility in women with rheumatoid arthritis and the outcome of fertility assessments. Arthritis Care Res (Hoboken) 2017;69(8):1142–1149.
7. Hashash JG, Kane S. Pregnancy and inflammatory bowel disease. Gastroenterol Hepatol (N Y) 2015;11(2):96–102.
8. Secher AEP, Granath F, Glintborg B, et al. Risk of pre-eclampsia and impact of disease activity and antirheumatic treatment in women with rheumatoid arthritis, axial spondylarthritis, and psoriatic arthritis: a collaborative matched cohort study from Sweden and Denmark. RMD Open 2022;8(2):e002445.

9. Smith CJF, Förger F, Bandoli G, et al. Factors associated with preterm delivery among women with rheumatoid arthritis and women with juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2019;71(8):1019–1027.
10. Tarar ZI, Farooq U, Zafar MU, et al. A national study of pregnancy-related maternal and fetal outcomes in women with inflammatory bowel disease. Int J Colorectal Dis 2022;37(7):1535–1543.
11. Reed SD, Vollan TA, Svec MA. Pregnancy outcomes in women with rheumatoid arthritis in Washington state. Matern Child Health J 2006;10(4):361–366.
12. Aljary H, Czuzoj-Shulman N, Spence AR, et al. Pregnancy outcomes in women with rheumatoid arthritis: a retrospective population-based cohort study. J Matern Fetal Neonatal Med 2020;33(4):618–624.
13. Baird DD, Narendranathan M, Sandler RS. Increased risk of preterm birth for women with inflammatory bowel disease. Gastroenterology 1990;99(4):987–994.
14. Magro F, Estevinho MM. Newborn congenital abnormalities and inflammatory bowel disease: unveiling an unexplored relationship. J Crohns Colitis 2020;14(8):1033–1034.
15. Vinet E, Pineau C, Gordon C, et al. Biologic therapy and pregnancy outcomes in women with rheumatic diseases. Arthritis Rheum 2009;61(5):587–592.
16. Conran CA, Moreland LW. A review of biosimilars for rheumatoid arthritis. Curr Opin Pharmacol 2022;64:102234.
17. Otezla. Product monograph. Amgen Canada Inc; 2020.
18. Xeljanz. Product monograph. Pfizer Canada ULC; 2014.
19. Olumiant. Product monograph. Eli Lilly Canada Inc; 2024.
20. Massalska M, Maslinski W, Ciechomska M. Small molecule inhibitors in the treatment of rheumatoid arthritis and beyond: latest updates and potential strategy for fighting COVID-19. Cells 2020;9(8):1876.
21. Aletaha D, Smolen JS. Diagnosis and management of rheumatoid arthritis: a review. JAMA 2018;320(13):1360–1372.
22. Benjamin O, Goyal A, Lappin SL. Disease-modifying antirheumatic drugs (DMARD). In: StatPearls. StatPearls Publishing; 2024.
23. Mother to Baby. Apremilast (Otezla). Organization of Teratology Information Specialists. Published February 1, 2024. Accessed August 14, 2024. https://mothertobaby.org/fact-sheets/apremilast/
24. Coricello A, Mesiti F, Lupia A, et al. Inside perspective of the synthetic and computational toolbox of JAK inhibitors: recent updates. Molecules 2020;25(15):3321.
25. US Food and Drug Administration. Drugs@FDA: FDA-approved drugs. Accessed August 14, 2024. https://www.accessdata.fda.gov/scripts/cder/daf/index.cfm?event=overview.process&ApplNo=205437
26. Winthrop KL, Cohen SB. Oral surveillance and JAK inhibitor safety: the theory of relativity. Nat Rev Rheumatol 2022;18(5):301–304.
27. OTEZLA (apremilast). Prescribing information. Amgen Inc; 2021.
28. Tsao NW, Rebic N, Lynd LD, et al. Maternal and neonatal outcomes associated with biologic exposure before and during pregnancy in women with inflammatory systemic diseases: a systematic review and meta-analysis of observational studies. Rheumatology (Oxford) 2020;59(8):1808–1817.
29. Tsao NW, Lynd LD, Sayre EC, et al. Use of biologics during pregnancy and risk of serious infections in the mother and baby: a Canadian population-based cohort study. BMJ Open 2019;9(2):e023714.
30. Götestam Skorpen C, Hoeltzenbein M, Tincani A, et al. The EULAR points to consider for use of antirheumatic drugs before pregnancy, and during pregnancy and lactation. Ann Rheum Dis 2016;75(5):795–810.
31. Förger F, Pluma Sanjurjo A, Rüegg L, et al. AB1439 update of the EULAR points to consider for use of antirheumatic drugs in reproduction, pregnancy and lactation. Ann Rheum Dis 2024;83(suppl 1):2075–2076.
32. Sammaritano LR, Bermas BL, Chakravarty EE, et al. 2020 American College of Rheumatology guideline for the management of reproductive health in rheumatic and musculoskeletal diseases. Arthritis Care Res (Hoboken) 2020;72(4):461–488.
33. Mahadevan U, Dubinsky MC, Su C, et al. Outcomes of pregnancies with maternal/paternal exposure in the tofacitinib safety databases for ulcerative colitis. Inflamm Bowel Dis 2018;24(12):2494–2500.
34. Mahadevan U, Levy G, Gensler L, et al. Pregnancy outcomes in patients treated with upadacitinib: analysis of data from clinical trials and postmarketing reports. Drug Saf 2024;47(10):1039–1049.
35. Clowse ME, Feldman SR, Isaacs JD, et al. Pregnancy outcomes in the tofacitinib safety databases for rheumatoid arthritis and psoriasis. Drug Saf 2016;39(8):755–762.
36. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005;8(1):19–32.
37. Peters MDJ, Marnie C, Colquhoun H, et al. Scoping reviews: reinforcing and advancing the methodology and application. Syst Rev 2021;10(1):263.
38. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018;169(7):467–473.
39. Covidence Systematic Review Software. Accessed August 14, 2024. www.covidence.org
40. The Joanna Briggs Institute. The Joanna Briggs Institute Reviewers' Manual 2015: Methodology for JBI Scoping Reviews. The Joanna Briggs Institute; 2015.
41. Oveisi N, Cheng V, Ellis U, et al. Reproductive health outcomes among adolescent and young adult cancer patients: a systematic review and meta-analysis. Cancers (Basel) 2023;15(6):1707.
42. Larsen MD, Friedman S, Magnussen B, et al. Birth outcomes in children fathered by men treated with anti-TNF-α agents before conception. Am J Gastroenterol 2016;111(11):1608–1613.
43. Perez-Garcia LF, Röder E, Smeele HTW, et al. Paternal inflammatory arthritis is associated with a higher risk of miscarriage: results of a large multicentre study (iFAME-Fertility). Rheumatology (Oxford) 2022;61(8):3390–3395.
44. Costanzo G, Firinu D, Losa F, et al. Baricitinib exposure during pregnancy in rheumatoid arthritis. Ther Adv Musculoskelet Dis 2020;12:1759720X19899296.
45. Fernández-Sánchez M, Ribes-Artero H, Romá-Sánchez E, et al. Fetal exposure to tofacitinib during the first trimester: a healthy newborn case report. Birth Defects Res 2021;113(17):1275–1279.
46. Mitrova K, Julsgaard M, Augustijns P, et al. Tofacitinib in pregnancy: assessing pregnancy and infant outcomes, cord blood, and breast milk concentrations. Clin Gastroenterol Hepatol 2025;23(1):163–165.e3.
47. Mahadevan U, Baumgart DC, Dubinsky MC, et al. S0847 pregnancy outcomes in the tofacitinib ulcerative colitis OCTAVE studies: an update as of February 2020. Am J Gastroenterol 2020;115:S437–S438.
48. Vinet E, St-Pierre Y, Moura C, et al. OP0225 serious infections in offspring exposed in utero to non-TNFi biologics and tofacitinib. Ann Rheum Dis 2019;78(suppl 2):189.
49. Savitz DA, Dole N, Herring AH. Methodologic issues in the design and analysis of epidemiologic studies of pregnancy outcome. Stat Methods Med Res 2006;15(2):93–102.
50. Alwan S, Chambers CD. Identifying human teratogens: an update. J Pediatr Genet 2015;4(2):39–41.
51. US Food and Drug Administration. Drug approval package: Xeljanz (tofacitinib) tablets. Published December 28, 2012. Accessed August 14, 2024. https://www.accessdata.fda.gov/drugsatfda_docs/nda/2012/203214orig1s000toc.cfm

52. Lin HC, Chen SF, Lin HC, et al. Increased risk of adverse pregnancy outcomes in women with rheumatoid arthritis: a nationwide population-based study. Ann Rheum Dis 2010;69(4):715–717.
53. Saavedra MA, Romo-Rodríguez R, Gutiérrez-Ureña SR, et al. Targeted drugs in spondyloarthritis during pregnancy and lactation. Pharmacol Res 2018;136:21–28.
54. Jorgensen SC, Tabbara N, Burry L. A review of COVID-19 therapeutics in pregnancy and lactation. Obstet Med 2022;15(4):225–232.
55. Eliesen GAM, Fransen M, van Hove H, et al. Placental transfer of tofacitinib in the ex vivo dual-side human placenta perfusion model. Curr Res Toxicol 2024;6:100149.
56. Gupta A, De Vera MA, Rebić N, et al. Pre-pregnancy planning for female patients seen at a pregnancy and rheumatic diseases clinic: a retrospective analysis of patients with rheumatic diseases seeking pregnancy-related care. Rheumatol Int 2024;44(2): 283–289.

Arthritis Care & Research
Vol. 77, No. 7, July 2025, pp 928–938
DOI 10.1002/acr.25515

American College of Rheumatology
Empowering Rheumatology Professionals

Effectiveness of a Telephone-Delivered Walk With Ease Program on Arthritis-Related Symptoms, Function, and Activity: A Randomized Trial

Christine A. Pellegrini, Sara Wilcox, Yesil Kim, Scott Jamieson, Katherine DeVivo, and Daniel Heidtke

Objective. Walk With Ease (WWE) is a six-week arthritis-appropriate evidence-based physical activity program traditionally offered in a face-to-face format. Because many populations encounter participation barriers to in-person programs, WWE was modified for telephone delivery (WWE-T). The short- and long-term effects of this program on physical activity and arthritis-related outcomes were examined.

Methods. Participants (n = 267) with arthritis were randomized to WWE-T or a wait list control. WWE-T participants received two telephone calls per week (one group and one individual call) for six weeks. Group calls focused on arthritis education and social support. Individual calls focused on problem-solving and goal setting. Physical function tests, patient-reported outcomes, and physical activity were assessed at baseline, 6 weeks, 6 months, and 12 months.

Results. Participants were 92% female and 60% Black and had a mean ± SD age of 64.1 ± 9.4 years and a body mass index of 34.2 ± 7.7. Retention ranged from 93.6% at 6 weeks to 83.8% at 12 months. Participants attended a mean ± SD of 9.8 ± 2.6 calls. At six weeks, WWE-T participants had greater improvements in physical function ($P = 0.03$), fatigue ($P = 0.03$), self-efficacy ($P \leq 0.0001$), and activity impairment due to health ($P = 0.01$) as compared to the control group. By 12 months, WWE-T participants had better physical function ($P = 0.02$), higher arthritis self-efficacy ($P \leq 0.0001$), lower depression symptoms ($P = 0.02$), and lower impairment of daily activities ($P = 0.02$) than at baseline.

Conclusion. A WWE-T program led to improvements in physical function, self-efficacy, and impairment related to daily activities in adults with arthritis. Although changes were not seen in all outcomes, this remotely delivered program may be an effective alternative for adults with arthritis who face barriers to in-person programs.

INTRODUCTION

More than 50 million adults in the United States have arthritis.[1] Nearly half of those with arthritis experience arthritis-attributable activity limitations.[2] Arthritis places substantial burden on the economy as a result of health care use and productivity or work loss.[3] As a result, there are several objectives within Healthy People 2030 that focus on improving the quality of life in those with arthritis.[4]

Physical activity is recommended as a nonpharmacological treatment approach for people with arthritis.[5,6] Among adults with arthritis, participation in regular aerobic physical activity is an effective strategy for improving pain, physical function, psychological wellness, and fatigue.[7–9] Despite the recommendations, almost half of adults with arthritis are insufficiently active or do not participate in any leisure time physical activity.[2]

Walk with Ease (WWE) is a six-week evidence-based physical activity program for people with arthritis.[10,11] WWE has led to improvements in arthritis symptoms, self-reported physical activity, mental health, and function.[10–12] Although the program is effective, many populations face barriers to attending in-person group WWE sessions, and regular attendance to three group sessions each week remains a challenge.[12] Alternative formats of WWE that have the capability to reach those who are most affected by arthritis and likely experience barriers to participating in in-person programs are needed.

The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the Centers for Disease Control and Prevention or the US Department of Health and Human Services.

Supported by Cooperative Agreement Number U48DP006401 from the Centers for Disease Control and Prevention.

Christine A. Pellegrini, PhD, Sara Wilcox, PhD, Yesil Kim, PhD, Scott Jamieson, MS, Katherine DeVivo, MPH, Daniel Heidtke, MPH: University of South Carolina, Columbia.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25515.

Address correspondence to Christine Pellegrini, PhD, at cpellegrini@sc.edu.

Submitted for publication September 23, 2024; accepted in revised form February 18, 2025.

SIGNIFICANCE AND INNOVATIONS

- A six-week telephone-delivered Walk With Ease (WWE) program is effective at improving physical function, self-efficacy, and impairment of daily activities in adults with arthritis more than one year.
- The telephone-delivered WWE program can be an effective short-term option for those adults with arthritis who may face challenges attending in-person sessions yet would benefit from supportive accountability to aid with behavior change and improve arthritis-attributable symptoms.

Telephone-based programs have been effective for changing numerous health behaviors, including physical activity.[13] Telephone-based contacts are highly acceptable to older adults,[14] are less burdensome and costly, and show greater potential for scalability. Thus, the WWE program was adapted for telephone delivery (WWE-T), and the purpose of this study was to evaluate the short- (6 weeks) and long-term (6 and 12 months) effects of WWE-T in adults with arthritis on pain and physical function (primary outcomes) as well as physical activity, self-efficacy, disability, depressive symptoms, weight, blood pressure, work loss, and health care use (secondary outcomes).

MATERIALS AND METHODS

Study design. This study was a randomized controlled trial examining the effects of a six-week WWE-T program as compared to a wait list control. Participants were recruited in 11 cohorts (10–33 participants per cohort) between March 2022 and August 2023 in the greater Columbia, South Carolina, area. Assessments were completed in-person. All study procedures were approved by the University of South Carolina Institutional Review Board, and participants provided informed written consent before participation.

Participants. Participants were adults ≥18 years old who meet criteria for the Centers for Disease Control and Prevention's definition of self-reported arthritis.[15] Specifically, participants who responded yes to the following question were eligible: "Have you ever been told by a doctor or other health professional that you have some form of arthritis, rheumatoid arthritis, gout, lupus, or fibromyalgia?" Additionally, participants had to be able to read and write in English; plan to live in the Columbia, South Carolina, region for the next year; and be willing to be randomized to either study condition. Participants were excluded if they had any contraindications to exercise (besides arthritis), were pregnant, or had a serious cognitive impairment.

The enrollment goal was 300 participants, with the aim to have 50% of participants between 18 and 64 years old or with low socioeconomic status. Assuming 25% attrition, a sample of 120 participants per group would provide 80% or greater power to detect a clinically meaningful difference in each primary outcome, with a minimum effect size of $d = 0.37$. Participants were recruited by using a variety of strategies (DeVivo K, et al: unpublished observations). Strategies included posting flyers throughout the community, attending community and local senior events, and having paid advertisements on local radio stations. Across the university community, emails were sent to university employees and flyers were posted around campus. Participants were also recruited from boosted Facebook posts, newspaper advertisements, and local health departments. Enrollment was stopped early because of lower attrition than anticipated.

A methods-motivational interviewing approach was used during the consent process to aid with making an informed decision to participate and to assist with retention.[16] During this process, study staff discussed with potential participants all study details and the pros and cons of participating in clinical trials and specifically the current study. Following the discussion, interested participants provided written informed consent.

Procedures. Following the completion of the baseline assessments, participants were randomized to start the WWE-T program either immediately or after a year (wait list control). Randomization was stratified by age (<65 years and ≥65 years) and sex, using randomly permuted blocks. A randomization list was generated by a biostatistician.

WWE-T program. Participants randomized to start the WWE-T program immediately were given a brief overview of the program over the phone and were mailed an Arthritis Foundation *Walk With Ease Guidebook* and supplemental handouts for the phone-delivered program. The WWE-T program was led by a certified WWE-T leader who completed the Arthritis Foundation's WWE Program Leader Training as well as supplemental telephone-delivery training. WWE-T leaders followed session-by-session instructions in a WWE-T leaders guide, which included the leader's notes, lecturettes, and participant handouts. All WWE-T leaders had backgrounds in exercise science, nutrition, public health, or a related field and were either health department employees or graduate students. For each cohort, the same WWE-T leader facilitated the six group calls. This leader also completed one-on-one calls with the help of one to two additional WWE-T leaders. WWE-T leaders always stayed with the same participants over the six-week program.

Participants were given the goal after randomization to work toward walking at least 30 min/day for three to five days per week. The program uses motivational strategies, which includes action planning, goal setting, and social support. Additionally, WWE-T provides participants with appropriate health education necessary to safely increase walking and exercise into their daily lives and assists in tailoring the program to fit individuals' unique needs and goals. Examples of topics discussed include exercise and arthritis, preparing to walk (eg, identifying appropriate footwear,

location, and walking pace), stretching, and overcoming barriers, including pain and discomfort.

During the six-week program, participants received two phone calls each week. The first call each week was a group call led by a WWE-T certified trained leader and included other WWE-T participants. During group phone sessions, the WWE-T leader covered topics relating to exercise and arthritis. There were 5 to 17 participants per group, and group calls were designed to last approximately 45 to 60 minutes.

The second call each week was a one-on-one call with a WWE-T certified trained leader. During individual phone sessions, the WWE leader focused on tailoring the WWE program for the individual based on their walking behaviors and barriers encountered the previous week, setting weekly walking goals, and assisting with developing a walking plan for the upcoming week. Individual calls were designed to last approximately 10 to 15 minutes.

Wait list control. Participants randomized to the wait list control group did not receive any intervention between baseline and 12 months. After participants in this condition completed the 12-month assessment, they were provided with an Arthritis Foundation *Walk With Ease Guidebook* and were offered the full six-week WWE-T program.

Assessments. In-person study assessments were completed at all time points by assessors masked to randomization. Participants received $25 for completing the baseline and 6-week assessment, $40 for completing the 6-month assessment, and $50 for completing the 12-month assessment. Additionally, participants were given study-branded incentives, including WWE-T pens, reusable bags, t-shirts, and water bottles. Surveys were completed on paper, and participants were given a padded envelope with prepaid postage to mail the activity monitor back after the seven-day wear period.

Demographic and health variables. At baseline, participants completed general demographic and health history surveys. The demographic variables assessed included age, sex, race, ethnicity, education, income, and marital status. For race and ethnicity, participants self-reported these variables from a set of fixed categories. The health history survey included an assessment of arthritis type, years since arthritis diagnosis, and presence of comorbid chronic health conditions (eg, hypertension, hyperlipidemia).

Primary outcome measures. *Self-reported pain.* A visual analog scale (VAS) was used to assess pain. Participants were asked to mark their experience with symptoms over the past seven days on a 100-mm line. Participants' response for pain was measured in millimeters from the left anchor (no pain, 0) to their mark. Higher scores were indicative of higher levels of pain.

Physical function. The 30-Second Chair Stand Test and the 6-Minute Walk Test assessed physical function, following standard procedures.[17] For the 30-Second Chair Stand Test, participants completed as many chair stand repetitions as possible during a 30-second period. For the 6-Minute Walk Test, participants walked between two cones 50 feet apart for six minutes. The maximal distance in feet a participant could walk during the six-minute period was measured. Higher scores on both tests indicate better physical function.

Secondary outcome measures. *Physical activity.* An ActiGraph GT9X Link accelerometer assessed minutes of moderate- to vigorous-intensity physical activity (MVPA). At each assessment, participants wore the device around the waist during waking hours for seven days. Participants also completed a daily log indicating the times the device was put on and taken off. Nonwear time was defined as ≥90 minutes with zero activity counts, allowing for up to 2 minutes of <100 counts/min.[18] Total minutes of MVPA per week (≥2,020 counts/min)[19] and average steps per day were calculated. Only those with four or more valid monitoring days (≥10 wear hours/day) were used in the analyses.[19]

Fatigue and stiffness. Fatigue and stiffness over the past seven days were assessed using a VAS on a 100-mm line. Responses were measured in millimeters from the left anchor (no pain, 0) to their mark, with higher scores indicating higher levels of fatigue or stiffness.

Arthritis management self-efficacy. An eight-item Arthritis Self-Efficacy Scale assessed participants' confidence to manage symptoms of arthritis.[20] Each item was answered on a 1- (very uncertain) to 10-point (very certain) scale, and the items were averaged to calculate a score. Higher scores indicate higher levels of self-efficacy for managing arthritis.

Depression symptoms. The Center for Epidemiologic Studies Depression Scale assessed depressive symptoms.[21] Participants rated the frequency of symptoms on 10 items using a 4-point Likert scale (0, rarely or none of the time, to 3, most or all of the time). Items 5 and 8 were reverse scored, and responses were summed to yield a total score ranging from 0 to 30, with a higher score indicative of higher levels of depressive symptoms.

Body weight. Height and weight were measured by trained research staff. Participants removed any extra clothing, shoes, and belts and emptied pockets before weighing. Participants' weight to the nearest 0.1 kg was recorded using an electronic scale (seca). Height was measured using a stadiometer to the nearest 0.1 cm. Height and weight were used to calculate body mass index (BMI) as weight in kg divided by height in m^2.

Activities of daily living. The Health Assessment Questionnaire (HAQ)[22] assessed eight categories of activities of daily living (ie, dressing, arising, eating, walking, hygiene, reach, grip, and common activities). There were 20 questions using a scale of 0 (without difficulty) to 3 (unable to do). These items were

averaged for a total score of 0 to 3, with a higher score indicating more impairment or disability.

Blood pressure. Systolic and diastolic blood pressure were measured according to the American Heart Association guidelines in triplicate each visit using an oscillometric blood pressure cuff (Omron) on the left arm at each visit in a seated position after a five-minute rest in a quiet room. The average of the three measures in mm Hg was used in analyses. If blood pressure was greater than 160/100, the session was stopped and the rest of the visit was rescheduled.

Work loss. The Work Productivity and Activity Impairment (WPAI) Questionnaire[23,24] includes six items that result in the calculation of four outcome percentages: absenteeism, presenteeism, overall work impairment due to health, and activity impairment due to health. Higher scores (range 0%–100%) indicate negative impact on health.

Health care use. The University of California at San Diego Healthcare Utilization Questionnaire[25] estimated health care use. For this study, the numbers of physician visits, telephone calls with health care providers, and medications (prescription and nonprescription) used in the last three months were examined.

Process measures and program evaluation. WWE-T leaders recorded session attendance and duration for both individual and group calls. The number of calls completed was summed for each participant, with 12 total calls possible (6 group and 6 individual). Participants in the WWE-T completed the WWE Post-Program Evaluation.[26] This evaluation asks participants to rate the extent to which specific program components benefited them over the past six weeks. Additional questions assess satisfaction with group and individual calls, length of group and individual calls, and rapport with the WWE-T leader. Participants responded to 13 items using a 4-point Likert scale from 0 (not at all) to 3 (very well).

Statistical analysis. Descriptive statistics were performed for baseline characteristics, process measures, and program evaluation. Baseline characteristics of the groups were compared using *t*-tests (continuous variables) or chi-square tests (categorical variables). The distributional assumption for outcomes was assessed and, where necessary, addressed by log transformation. As a result, total MVPA (minutes per week) scores were normally distributed after transformation (skewness ranged from –0.65 to 0.03, and kurtosis ranged from –0.86 to 0.04 at each time point).

An intent-to-treat analysis was conducted on the full sample, and the primary and secondary outcomes were addressed using repeated-measures mixed models with maximum likelihood estimation. A difference-in-difference analysis was used to examine the relationship between group assignment and changes in outcomes. Within-group changes in primary and secondary outcomes were contrasted from baseline to 6 weeks, 6 months, and 12 months. The magnitudes of within-group changes were also compared between groups, computed by subtracting change scores in the wait list control group from those in WWE-T participants. The mixed model estimated adjusted group means. Time and group were main effects variables, and the time by group variable was an interaction effect specified in the model. We included age, sex, race, income, the three most prevalent types of arthritis (osteoarthritis, rheumatoid arthritis, and unsure), and the four most common comorbidities (diabetes, hypertension, anxiety, and sleep apnea) in the model as covariates. For the WPAI Questionnaire, analyses were only run among those who reported full-time or part-time employment (n = 106). All analyses were conducted using SAS 9.4 (SAS Institute, Inc). To test whether the least squares mean differences were significantly different from zero or not, we examined the *t*-test *P* value associated with the least squares mean comparison. The statistically significant level for a two-sided *P* value was set to 0.05.

RESULTS

A total of 595 participants were screened to participate, and 291 participants were consented and completed the baseline assessment (Figure 1). Of those who were consented, 91.7% of participants (n = 267) remained eligible and were randomized. See Table 1 for baseline demographic and sociodemographic characteristics. Participants were primarily female (92%) and Black (60%), and 52% of participants had an income <$73,801. Participants' average ± SD age was 64.1 ± 9.4 years. The most commonly reported forms of arthritis were osteoarthritis (61%) and rheumatoid arthritis (24%). Hypertension (56%), diabetes (23%), sleep apnea (22%), and anxiety (20%) were the most common comorbid conditions reported.

Baseline health and behavior characteristics by randomized group are shown in Table 2. There were no differences between groups besides in the number of participants with sleep apnea (*P* = 0.0476). Participants had an average ± SD BMI of 34.2 ± 7.7. Participants had an average ± SD of 57.0 ± 75.4 min/wk of MVPA and took 4,171.8 ± 2,185.1 steps/day. Participants wore the activity monitor for a mean ± SD of 857.4 ± 99.7 min/day at baseline for 6.5 ± 0.9 days/wk. There were no differences in wear time or days worn between groups at any time point.

Retention was 93.6%, 87.6%, and 83.8% at 6 weeks, 6 months, and 12 months, respectively (Figure 1). There were no differences between groups at any time point. Over the course of the study, there were five adverse events in the WWE-T condition that were deemed possibly related to the study and walking. The events included swollen knees (n = 3), knee injury after fall (n = 1), and a stroke (n = 1).

WWE-T participants completed an average ± SD of 9.8 ± 2.6 calls (4.5.3 ± 1.7 group calls and 5.3 ± 1.3 individual calls) of the 12 possible calls. Group calls were an average ± SD of 52.0 ± 5.6 minutes in duration, and individual calls were 12.0 ± 4.0 minutes. Ninety-three percent of WWE-T participants who

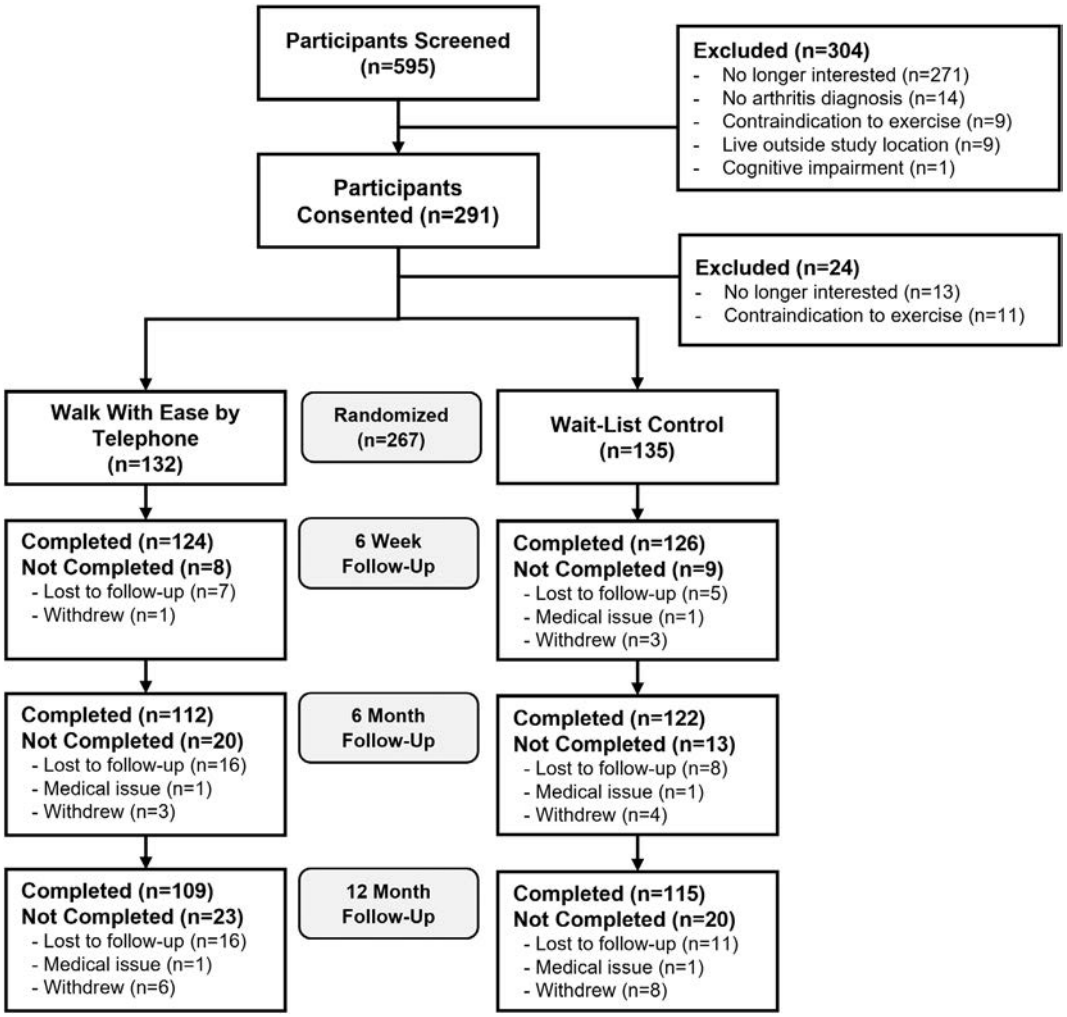

Figure 1. Walk With Ease by telephone flowchart.

completed the program evaluation reported the extent they were satisfied with the program as "fairly well" or "very well" (Table 3).

Changes in all outcomes across the three time points by group and between groups are presented in Table 4. The WWE-T group had greater improvements over the six weeks in chair stand repetitions ($P = 0.03$), fatigue ($P = 0.03$), arthritis management self-efficacy ($P \leq 0.0001$), and activity impairment due to health ($P = 0.01$) as compared to the control group. HAQ disability ($P = 0.0012$) and depressive symptoms ($P = 0.01$) significantly improved in the WWE-T group, but the group × time interactions were not significant. In the wait list control group, physician visits decreased over time ($P = 0.01$).

At six months, arthritis management self-efficacy ($P = 0.01$) was significantly greater in the WWE-T group as compared to the control group. Additionally, at six months, the WWE-T group had more 30-second chair stand repetitions ($P = 0.0035$), greater reductions in depressive symptoms ($P = 0.01$), and lower activity impairment due to health ($P = 0.02$) than the control group. Both groups had an increased number of physician visits at six months.

By 12 months, the WWE-T group had significantly greater improvements in arthritis self-efficacy ($P = 0.03$) as compared to the control group. From baseline to 12 months, the WWE-T group had more chair stand repetitions ($P = 0.02$), lower depression symptoms ($P = 0.02$), and lower activity impairment due to health ($P = 0.02$). Log-transformed MVPA was lower in WWE-T group at 12 months than in the control group ($P = 0.04$).

DISCUSSION

In the short term, a WWE-T program was effective at improving physical function, reducing fatigue and impairment of daily activities, and improving arthritis management self-efficacy in adults with arthritis. By 12 months, participants in the WWE-T condition had better physical function, higher arthritis self-efficacy, lower depression symptoms, and lower impairment of daily activities than at baseline. Overall satisfaction with the telephone-delivered program and adherence to the telephone

Table 1. Baseline demographic characteristics and health history

Characteristics	Walk With Ease by telephone (n = 132)	Wait list control (n = 135)	*P* value
Age, mean ± SD, y	64.06 ± 9.38	64.13 ± 9.36	0.95
Female, n (%)	122 (92)	123 (91)	0.70
Race, n (%)			0.82
Black	80 (61)	80 (59)	
Non-Black	52 (39)	55 (41)	
Ethnicity, n (%)			0.57
Not Hispanic or Latino	131 (99)	132 (99)	
Income, n (%)			0.91
<$73,801	69 (52)	69 (51)	
≥$73,801	40 (30)	44 (33)	
Not reported	23 (17)	22 (16)	
Time since arthritis diagnosis, mean ± SD, mo	144.9 ± 116.60	148.50 ± 131.70	0.81
Type of arthritis, n (%)			
Osteoarthritis	80 (61)	82 (61)	0.98
Rheumatoid arthritis	32 (24)	31 (23)	0.81
Fibromyalgia	12 (9)	12 (9)	0.95
Gout	9 (7)	14 (10)	0.30
Psoriatic arthritis	2 (2)	7 (5)	0.10
Lupus	4 (3)	8 (6)	0.25
Sjögren disease	2 (2)	2 (1)	0.98
Unsure of arthritis diagnosis	25 (19)	19 (14)	0.28
Education, n (%)			0.66
High school/some college	42 (32)	37 (27)	
College graduate	41 (31)	48 (36)	
Advanced degree	49 (37)	50 (37)	
Employment, n (%)			0.21
Full-time employment	30 (23)	41 (30)	
Part-time employment	21 (16)	14 (10)	
Unemployed or retired	81 (61)	80 (59)	
Comorbidities, n (%)			
Diabetes	30 (23)	32 (24)	0.85
Hypertension	75 (57)	75 (56)	0.84
Anxiety	32 (24)	21 (16)	0.08
Depression	25 (19)	21 (16)	0.46
Sleep apnea	22 (17)	36 (27)	0.0476
Asthma	21 (16)	13 (10)	0.12
Thyroid problems	23 (17)	18 (13)	0.35

calls were high. Additionally, WWE-T was safe and provided a feasible and acceptable alternative to the in-person or self-guided formats.

Among primary outcomes, significant differences were only seen in the 30-Second Chair Stand Test. The improvement in chair stands is similar to previous examinations of the group or self-directed versions of WWE.[10] The 30-Second Chair Stand Test is a measure of lower body strength and dynamic balance, whereas the 6-Minute Walk Test examines aerobic capacity, endurance, and dynamic balance. It is possible that participants may have increased walking but did not walk at a high enough intensity to improve aerobic capacity. Even if aerobic capacity did not improve, an increase in walking of any intensity among a highly inactive population may have helped to improve lower body strength; thus, only changes were seen in the 30-Second Chair Stand Test. Previous studies examining the effectiveness of the WWE program have relied on performance-based testing[10] or self-reported physical activity[11,12,27]; thus, this study provides an objective indicator of intensity of physical activity levels in adults with arthritis. Although activity levels still remained below federal physical activity guidelines, any walking can still result in substantial health benefits.[28]

In addition to not seeing changes in the 6-Minute Walk Test, the current study did not see changes in pain, which contradicts previous examinations of WWE.[10,12] It is unclear why pain levels did not change in the current study. Fortunately, there were numerous improvements in secondary outcomes. The telephone-delivered version of WWE led to short-term improvements in fatigue, arthritis management self-efficacy, and impairment to daily activities. The magnitude of effect sizes of these changes was small to medium (0.30–0.58) and higher than that of previous examinations of either the self-guided or group versions of WWE (effect sizes 0.09–0.22).[10,12] With the exception of fatigue, improvements in these outcomes continued through 12 months, suggesting long-term benefits of the program.

Table 2. Baseline health and behavior characteristics by randomized condition*

Variable	Walk With Ease by telephone (n = 132), mean ± SE	Wait list control (n = 135), mean ± SE
Pain, mm	42.93 ± 3.75	40.16 ± 3.75
Fatigue, mm	47.62 ± 4.47	44.13 ± 4.44
Stiffness, mm	54.49 ± 3.99	50.08 ± 4.08
Function or physical limitations		
30-s chair stand, repetitions	8.62 ± 0.42	8.43 ± 0.45
6-min walk test, ft	1,292.20 ± 42.00	1,282.32 ± 44.66
Health Assessment Questionnaire	0.47 ± 0.05	0.47 ± 0.06
Actigraph-assessed activity		
Total MVPA, min/wk	51.32 ± 10.36	54.83 ± 11.24
Steps, steps/day	3,690.30 ± 326.89	3,883.17 ± 356.07
Arthritis management self-efficacy	6.31 ± 0.29	6.30 ± 0.30
Depression (CES-D score)	8.46 ± 0.76	8.40 ± 0.82
Weight, kg	101.08 ± 2.85	101.67 ± 2.86
Body mass index	34.59 ± 1.05	34.68 ± 1.08
Blood pressure, mm Hg		
Systolic	129.24 ± 2.09	128.45 ± 2.25
Diastolic	71.79 ± 1.46	72.19 ± 1.57
Work productivity and activity impairment, %		
Work time missed	1.72 ± 1.54	1.07 ± 1.28
Impairment while working	24.88 ± 5.62	26.69 ± 5.07
Overall work impairment	25.74 ± 5.79	27.43 ± 5.15
Activity impairment	44.01 ± 4.04	40.11 ± 3.83
Health care use		
Physician visits	2.68 ± 0.35	3.06 ± 0.37
Telephone calls made to doctors or staff	1.37 ± 0.24	1.81 ± 0.30
Prescription medicines taken regularly	4.99 ± 0.38	5.21 ± 0.38
Nonprescription medicines taken regularly	3.32 ± 0.42	3.39 ± 0.47

* CES-D, Center for Epidemiologic Studies Depression Scale; MVPA, moderate- to vigorous-intensity physical activity.

One reason the WWE-T program was successful may have been due to the high adherence to the program. In the current study, participants completed an average of 9.8 of the possible 12 phone sessions (82%). This adherence is high, particularly in comparison to a previous scaling-up of WWE in which participants attended approximately 8.5 sessions of the possible

Table 3. Walk With Ease by telephone postprogram evaluation (n = 107)

To what extent	Not at all, n (%)	A little, n (%)	Fairly well, n (%)	Very well, n (%)
Did you learn basic information about arthritis?	0 (0)	5 (4.7)	29 (27.1)	72 (67.3)
Did you increase your understanding of the rationale and principles of exercise for people with arthritis?	2 (1.9)	2 (1.9)	21 (19.6)	82 (76.6)
Did you increase your knowledge about walking in a safe and comfortable matter?	2 (1.9)	8 (7.5)	15 (14.0)	82 (76.6)
Do you feel knowledgeable about how to do warm-up and cool-down exercises before and after walking?	1 (0.9)	3 (2.8)	17 (15.9)	86 (80.4)
Were the problem-solving strategies useful to you?	2 (1.9)	4 (3.7)	31 (29.0)	70 (65.4)
Were the self-test tools useful to you?	1 (0.9)	6 (5.6)	37 (34.6)	63 (58.9)
Were the contract and walking diary tools useful to you?	3 (2.8)	8 (7.5)	34 (31.8)	62 (57.9)
Were the group calls helpful for you?	3 (2.8)	10 (9.3)	26 (24.3)	68 (63.6)
Were the one-on-one calls with the leader helpful for you?	2 (1.9)	2 (1.9)	12 (11.2)	90 (84.1)
Are you happy with the length of the program?	1 (0.9)	9 (8.4)	26 (24.3)	71 (66.4)
Did Walk With Ease fulfill your expectations?	2 (1.9)	6 (5.6)	23 (21.5)	76 (71.0)
Are you satisfied with the program?	2 (1.9)	4 (3.7)	15 (14.0)	85 (79.4)

Table 4. Changes in arthritis-related symptoms, function, and physical activity by group across time*

Variable	Walk With Ease by telephone		Wait list control		Difference between groups (95% CI)	Effect size, *d*	*P* value
	Change from baseline (95% CI)	*P* value	Change from baseline (95% CI)	*P* value			
Pain, mm							
To 6 wk	−4.00 (−11.30 to 3.30)	0.28	−4.49 (−11.59 to 2.60)	0.21	0.49 (−5.52 to 6.50)	0.02	0.87
To 6 mo	−1.04 (−8.99 to 6.92)	0.80	0.35 (−7.05 to 7.76)	0.93	−1.39 (−8.59 to 5.80)	−0.06	0.70
To 12 mo	−5.02 (−13.87 to 3.83)	0.27	−4.89 (−12.71 to 2.93)	0.22	−0.13 (−9.07 to 8.81)	−0.01	0.98
Fatigue, mm							
To 6 wk	−7.60 (−15.22 to 0.02)	0.05	−0.44 (−7.29 to 6.41)	0.90	**−7.16 (−13.69 to −0.64)**	−0.30	**0.03**
To 6 mo	−0.23 (−8.09 to 7.64)	0.95	0.19 (−6.61 to 6.99)	0.96	−0.42 (−7.63 to 6.80)	−0.02	0.91
To 12 mo	−5.93 (−16.22 to 4.36)	0.26	2.89 (−6.10 to 11.88)	0.53	−8.82 (−19.70 to 2.06)	−0.31	0.11
Stiffness, mm							
To 6 wk	−5.55 (−12.08 to 0.97)	0.10	−4.47 (−11.21 to 2.27)	0.19	−1.08 (−7.08 to 4.91)	−0.05	0.72
To 6 mo	−0.99 (−8.32 to 6.34)	0.79	−0.64 (−8.40 to 7.13)	0.87	−0.36 (−8.04 to 7.33)	−0.01	0.93
To 12 mo	−5.49 (−14.79 to 3.80)	0.25	−1.80 (−10.21 to 6.61)	0.67	−3.70 (−13.78 to 6.39)	−0.14	0.47
30-Second Chair Stand Test, repetitions							
To 6 wk	**0.98 (0.43 to 1.52)**	**0.0005**	0.48 (−0.08 to 1.04)	0.09	**0.50 (0.05 to 0.95)**	0.26	**0.03**
To 6 mo	**1.00 (0.33 to 1.67)**	**0.0035**	**0.69 (0.12 to 1.25)**	**0.02**	0.31 (−0.20 to 0.83)	0.18	0.23
To 12 mo	**0.79 (0.13 to 1.45)**	**0.02**	0.45 (−0.16 to 1.06)	0.15	0.33 (−0.32 to 0.99)	0.18	0.32
6-Minute Walk Test, ft							
To 6 wk	17.66 (−21.53 to 56.86)	0.38	−10.15 (−48.30 to 27.99)	0.60	27.82 (−2.75 to 58.38)	0.24	0.07
To 6 mo	−27.33 (−79.66 to 25.01)	0.31	0.24 (−40.64 to 41.12)	0.99	−27.57 (−72.64 to 17.51)	−0.19	0.23
To 12 mo	−21.32 (−76.30 to 33.67)	0.45	2.19 (−40.78 to 45.16)	0.92	−23.51 (−77.10 to 30.09)	−0.19	0.39
Health Assessment Questionnaire							
To 6 wk	**−0.09 (−0.15 to −0.04)**	**0.0012**	−0.05 (−0.12 to 0.01)	0.12	−0.04 (−0.09 to 0.01)	−0.20	0.13
To 6 mo	−0.05 (−0.12 to 0.01)	0.10	−0.02 (−0.10 to 0.05)	0.50	−0.03 (−0.09 to 0.03)	−0.12	0.37
To 12 mo	−0.06 (−0.12 to 0.01)	0.11	−0.02 (−0.08 to 0.04)	0.54	−0.04 (−0.10 to 0.03)	−0.25	0.29
Total MVPA log transformed							
To 6 wk	0.10 (−0.18 to 0.38)	0.47	−0.08 (−0.37 to 0.21)	0.58	0.19 (−0.06 to 0.43)	0.18	0.14
To 6 mo	−0.15 (−0.44 to 0.14)	0.32	0.03 (−0.30 to 0.36)	0.85	−0.18 (−0.48 to 0.12)	−0.14	0.24
To 12 mo	**−0.16 (−0.62 to 0.00)**	**0.05**	−0.01 (−0.32 to 0.29)	0.94	**−0.30 (−0.58 to −0.01)**	−0.26	**0.04**
Steps, steps/day							
To 6 wk	366.88 (−115.82 to 849.59)	0.14	87.89 (−327.60 to 503.38)	0.68	278.99 (−139.28 to 697.27)	0.19	0.19
To 6 mo	90.60 (−411.45 to 592.65)	0.72	66.28 (−384.96 to 517.52)	0.77	24.32 (−388.16 to 436.79)	0.01	0.91
To 12 mo	−225.44 (−759.13 to 308.24)	0.41	−39.21 (−455.74 to 377.33)	0.85	−186.24 (−607.26 to 234.79)	−0.12	0.39
Arthritis management self-efficacy							
To 6 wk	**1.55 (0.93 to 2.17)**	**<0.0001**	0.49 (−0.11 to 1.08)	0.11	**1.06 (0.61 to 1.52)**	0.58	**<0.0001**
To 6 mo	**1.17 (0.55 to 1.80)**	**0.0002**	0.52 (−0.10 to 1.14)	0.10	**0.65 (0.16 to 1.15)**	0.36	**0.01**
To 12 mo	**1.42 (0.73 to 2.11)**	**<0.0001**	0.65 (−0.04 to 1.34)	0.06	**0.77 (0.09 to 1.44)**	0.38	**0.03**
Depression (CES-D score)							
To 6 wk	**−1.57 (−2.69 to −0.45)**	**0.01**	−0.91 (−1.94 to 0.13)	0.08	−0.66 (−1.79 to 0.47)	−0.16	0.25
To 6 mo	**−1.51 (−2.68 to −0.34)**	**0.01**	−0.40 (−1.50 to 0.69)	0.47	−1.11 (−2.37 to 0.15)	−0.24	0.08
To 12 mo	**−1.53 (−2.80 to −0.26)**	**0.02**	−0.08 (−1.33 to 1.17)	0.91	−1.45 (−2.91 to 0.01)	−0.40	0.05
Weight, kg							
To 6 wk	−0.21 (−1.35 to 0.93)	0.71	0.13 (−0.84 to 1.10)	0.80	−0.34 (−1.03 to 0.35)	−0.15	0.33
To 6 mo	−0.19 (−1.17 to 0.79)	0.70	0.05 (−1.07 to 1.17)	0.93	−0.24 (−1.18 to 0.70)	−0.07	0.62
to 12 mo	−1.27 (−2.88 to 0.34)	0.12	−0.46 (−2.02 to 1.09)	0.56	−0.81 (−2.72 to 1.11)	−0.16	0.41
Body mass index							
To 6 wk	−0.01 (−0.46 to 0.45)	0.98	0.08 (−0.31 to 0.46)	0.69	−0.08 (−0.36 to 0.19)	−0.09	0.55
To 6 mo	−0.02 (−0.42 to 0.37)	0.92	0.07 (−0.35 to 0.49)	0.75	−0.09 (−0.46 to 0.27)	−0.07	0.62
To 12 mo	−0.35 (−0.98 to 0.27)	0.27	−0.14 (−0.74 to 0.47)	0.66	−0.22 (−0.94 to 0.51)	−0.12	0.56
Systolic blood pressure, mm Hg							
To 6 wk	−0.59 (−4.60 to 3.43)	0.77	0.07 (−3.67 to 3.80)	0.97	−0.65 (−4.47 to 3.16)	−0.04	0.74
To 6 mo	−1.22 (−5.30 to 2.87)	0.56	−1.72 (−5.92 to 2.48)	0.42	0.50 (−3.72 to 4.73)	0.03	0.82
To 12 mo	−3.51 (−8.18 to 1.16)	0.14	−1.49 (−5.91 to 2.93)	0.51	−2.02 (−7.28 to 3.24)	−0.14	0.45
Diastolic blood pressure, mm Hg							
To 6 wk	−0.46 (−3.01 to 2.10)	0.73	−0.23 (−2.81 to 2.35)	0.86	−0.23 (−2.61 to 2.16)	−0.02	0.85
To 6 mo	−0.67 (−3.37 to 2.02)	0.62	−1.16 (−3.84 to 1.52)	0.40	0.49 (−2.16 to 3.14)	0.05	0.72
To 12 mo	−2.52 (−5.62 to 0.57)	0.11	−0.40 (−3.21 to 2.41)	0.78	−2.12 (−5.43 to 1.19)	−0.22	0.21
Work time missed, %							
To 6 wk	0.56 (−4.59 to 5.72)	0.83	−0.50 (−4.36 to 3.35)	0.80	1.07 (−3.92 to 6.05)	0.16	0.67
To 6 mo	−2.27 (−5.44 to 0.91)	0.16	0.85 (−3.76 to 5.46)	0.72	−3.11 (−6.71 to 0.48)	−0.25	0.09
To 12 mo	2.79 (−5.81 to 11.40)	0.52	−0.29 (−4.67 to 4.08)	0.89	3.09 (−5.44 to 11.62)	0.27	0.48

(Continued)

Table 4. (*Cont'd*)

Variable	Walk With Ease by telephone		Wait list control		Difference between groups (95% CI)	Effect size, *d*	*P* value
	Change from baseline (95% CI)	*P* value	Change from baseline (95% CI)	*P* value			
Impairment while working, %							
To 6 wk	−3.20 (−14.79 to 8.40)	0.59	−8.24 (−18.62 to 2.15)	0.12	5.04 (−5.04 to 15.12)	0.28	0.33
To 6 mo	−1.92 (−13.22 to 9.38)	0.74	−6.51 (−17.65 to 4.63)	0.25	4.59 (−6.66 to 15.84)	0.17	0.42
To 12 mo	0.96 (−13.55 to 15.46)	0.90	−7.49 (−16.23 to 1.24)	0.09	8.45 (−6.00 to 22.90)	0.55	0.25
Overall work impairment, %							
To 6 wk	−3.58 (−15.51 to 8.35)	0.55	−8.14 (−19.05 to 2.77)	0.14	4.56 (−5.79 to 14.91)	0.24	0.39
To 6 mo	−2.59 (−14.30 to 9.11)	0.66	−5.71 (−17.27 to 5.85)	0.33	3.12 (−8.76 to 14.99)	0.11	0.61
To 12 mo	1.72 (−15.05 to 18.49)	0.84	−7.11 (−16.38 to 2.16)	0.13	8.83 (−8.03 to 25.69)	0.51	0.30
Activity impairment, %							
To 6 wk	**−11.98 (−19.16 to −4.80)**	**0.0011**	−3.91 (−10.91 to 3.08)	0.27	**−8.06 (−14.21 to −1.92)**	−0.38	**0.01**
To 6 mo	**−9.52 (−17.39 to −1.65)**	**0.02**	−4.32 (−11.92 to 3.28)	0.27	−5.20 (−12.46 to 2.06)	−0.21	0.16
To 12 mo	**−10.51 (−19.49 to −1.54)**	**0.02**	−5.64 (−12.81 to 1.53)	0.12	−4.88 (−13.66 to 3.91)	−0.27	0.28
Physician visits							
To 6 wk	−0.48 (−1.27 to 0.31)	0.23	**−1.14 (−1.95 to −0.33)**	**0.01**	0.66 (−0.07 to 1.39)	0.24	0.08
To 6 mo	**1.21 (0.40 to 2.02)**	**0.0034**	**1.12 (0.10 to 2.15)**	**0.03**	0.09 (−0.96 to 1.14)	0.02	0.87
To 12 mo	**1.89 (0.80 to 2.99)**	**0.0007**	0.99 (−0.39 to 2.38)	0.16	0.90 (−0.68 to 2.48)	0.19	0.26
Telephone calls made to doctors or staff							
To 6 wk	−0.26 (−0.74 to 0.21)	0.28	−0.45 (−1.07 to 0.17)	0.15	0.19 (−0.36 to 0.74)	0.07	0.50
To 6 mo	**0.73 (0.18 to 1.27)**	**0.01**	0.22 (−0.46 to 0.90)	0.53	0.51 (−0.18 to 1.20)	0.16	0.15
To 12 mo	**0.83 (0.09 to 1.56)**	**0.03**	0.19 (−0.57 to 0.95)	0.62	0.64 (−0.28 to 1.55)	0.24	0.17
Prescription medicines taken regularly							
To 6 wk	0.25 (−0.20 to 0.70)	0.27	0.03 (−0.38 to 0.44)	0.89	0.22 (−0.13 to 0.58)	0.17	0.22
To 6 mo	0.26 (−0.22 to 0.75)	0.29	0.40 (−0.13 to 0.92)	0.14	−0.13 (−0.58 to 0.31)	−0.07	0.56
To 12 mo	0.70 (−0.02 to 1.42)	0.06	0.41 (−0.14 to 0.97)	0.14	0.29 (−0.45 to 1.03)	0.18	0.45
Nonprescription medicines taken regularly							
To 6 wk	0.33 (−0.15 to 0.81)	0.18	0.20 (−0.21 to 0.61)	0.34	0.13 (−0.28 to 0.54)	0.09	0.54
To 6 mo	0.43 (−0.04 to 0.91)	0.07	0.26 (−0.22 to 0.75)	0.29	0.17 (−0.31 to 0.66)	0.10	0.49
To 12 mo	0.03 (−0.54 to 0.61)	0.91	0.48 (−0.21 to 1.18)	0.17	−0.45 (−1.14 to 0.25)	−0.22	0.21

* Bold text indicates significant difference between baseline and the follow-up or between groups ($P < 0.05$). CES-D, Center for Epidemiologic Studies Depression Scale; CI, confidence interval; MVPA, moderate- to vigorous-intensity physical activity.

18 in-person sessions (47%).[12] Since the COVID-19 pandemic, there has been more of a need for remotely delivered programs, particularly among older adults and those at higher risk of complications from COVID-19.[29] A telephone-delivered program provides supportive accountability from a group and leader yet also removes the traditional barriers to participation in an in-person program. Meeting participants' needs and preferences can lead to greater adherence to the program, which should lead to subsequent greater improvements in outcomes. Further, increased adherence to the program could result in an even greater incremental cost-effectiveness ratio than that recently demonstrated by a cost-effectiveness analysis of the self-directed version of WWE.[30]

From an implementation perspective, WWE-T may be a preferred and more feasible modality for organizations to offer as compared to in-person groups. Barriers that organizations face include poor weather conditions for outdoor walking and a lack of physical space for meetings and indoor walking.[31] During the WWE-T program, participants walk on their own time, eliminating the need for staff and walking space and substantially reducing the resources necessary to implement. Additionally, from a qualitative evaluation of organizations that implemented WWE, program managers recommended that WWE sessions should be offered two times per week instead of three times per week.[31] The frequency of two sessions per week aligns with the WWE-T and may be more feasible for both program organizations and participants.

This study had several strengths and limitations. First, this is one of the first studies to use objective monitoring to assess changes in physical activity from the WWE program. Second, previous examinations of WWE were with predominantly White and highly educated participants.[10,11] Sixty percent of the current study sample was Black, and 52% was below the US median household income; thus, the results provide insight on the effects of this program in underrepresented populations. Also, remotely delivered programs may be preferred among underrepresented populations and align with previous research that when given the option, more Black adults with arthritis opted to participate in the self-directed version of WWE as compared to the in-person group format.[32] Finally, retention was high, with 83.8% to 93.6% of participants completing some aspect of the assessment. Previous studies have had low survey or postprogram evaluation completion rates, particularly in self-directed versions of WWE.[11,27,33]

Although this study highlighted specific strengths, several limitations could impact the generalizability of the findings. First,

the study sample was predominantly female, which coincides with previous research regarding limited engagement of men in physical activity research.[34] Further, adults with arthritis from the southeastern United States were exclusively recruited into the study. Findings from this sample may not be as translatable to others residing in different regions of the country or internationally. Nonetheless, we focused on an important population because the Southeast has a high prevalence of arthritis, poor health, and severe joint pain.[35] Further, intervention fidelity was not evaluated in this community-delivered program beyond call delivery and duration, limiting our understanding on whether the program was implemented as intended. Additionally, although the results suggest the program is effective when delivered by telephone, a direct comparison of the effectiveness of this delivery modality as compared to group or self-directed WWE formats was not examined. Future research could examine differences in adherence, preferences, and behavioral and health-related outcomes across the three formats of WWE.

In summary, a WWE-T program led to short-term improvements in physical function, self-efficacy, and impairment related to daily activities in adults with arthritis. Long-term changes were observed at 12 months in physical function, self-efficacy, depression symptoms, and impairment related to daily activities. Although changes were not seen in all outcomes, this remotely delivered program may be an effective alternative to improve arthritis symptoms in adults with arthritis and other underrepresented populations who may face barriers to in-person community-based physical activity programs.

ACKNOWLEDGMENTS

We thank our collaborators from the Supplemental Nutrition Assistance Program Education, including Farrah Wigand, RDN, LD, CSOWM; Crystal Connor, MS, RD, LD; and Leah Price, MS, RD, LD, as implementation of this project would not have been possible without them. We also thank the Arthritis Foundation for their encouragement and input on the modification of the WWE program. We would also like to thank all study participants, research staff, and students for their important contributions to this project.

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 Pellegrini 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. Fallon EA, Boring MA, Foster AL, et al. Prevalence of diagnosed arthritis - United States, 2019-2021. MMWR Morb Mortal Wkly Rep 2023;72(41):1101–1107.
2. Barbour KE, Helmick CG, Boring M, et al. Vital signs: prevalence of doctor-diagnosed arthritis and arthritis-attributable activity limitation - United States, 2013-2015. MMWR Morb Mortal Wkly Rep 2017;66(9):246–253.
3. Murphy LB, Cisternas MG, Pasta DJ, et al. Medical expenditures and earnings losses among US adults with arthritis in 2013. Arthritis Care Res (Hoboken) 2018;70(6):869–876.
4. Office of Disease Prevention and Health Promotion. Healthy People 2030. 2020. Accessed January 27, 2021. https://health.gov/healthypeople
5. 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.
6. Katz P, Andonian BJ, Huffman KM. Benefits and promotion of physical activity in rheumatoid arthritis. Curr Opin Rheumatol 2020;32(3):307–314.
7. Kelley GA, Kelley KS, Hootman JM, et al. Effects of community-deliverable exercise on pain and physical function in adults with arthritis and other rheumatic diseases: a meta-analysis. Arthritis Care Res (Hoboken) 2011;63(1):79–93.
8. Conn VS, Hafdahl AR, Minor MA, et al. Physical activity interventions among adults with arthritis: meta-analysis of outcomes. Semin Arthritis Rheum 2008;37(5):307–316.
9. Kelley GA, Kelley KS, Hootman JM. Effects of exercise on depression in adults with arthritis: a systematic review with meta-analysis of randomized controlled trials. Arthritis Res Ther 2015;17(1):21.
10. Callahan LF, Shreffler JH, Altpeter M, et al. Evaluation of group and self-directed formats of the Arthritis Foundation's Walk With Ease Program. Arthritis Care Res (Hoboken) 2011;63(8):1098–1107.
11. Bruno M, Cummins S, Gaudiano L, et al. Effectiveness of two Arthritis Foundation programs: Walk With Ease, and YOU Can Break the Pain Cycle. Clin Interv Aging 2006;1(3):295–306.
12. Conte KP, Odden MC, Linton NM, et al. Effectiveness of a scaled-up arthritis self-management program in Oregon: Walk With Ease. Am J Public Health 2016;106(12):2227–2230.
13. Goode AD, Reeves MM, Eakin EG. Telephone-delivered interventions for physical activity and dietary behavior change: an updated systematic review. Am J Prev Med 2012;42(1):81–88.
14. Lamoureux NR, Phillips LA, DeShaw KJ, et al. Evaluating the feasibility and utility of telephonic motivational interviewing in older adults. PEC Innov 2024;5:100344.
15. Murphy LB, Cisternas MG, Greenlund KJ, et al. Defining arthritis for public health surveillance: methods and estimates in four US population health surveys. Arthritis Care Res (Hoboken) 2017;69(3):356–367.
16. Jake-Schoffman DE, Brown SD, Baiocchi M, et al. Methods-motivational interviewing approach for enhanced retention and attendance. Am J Prev Med 2021;61(4):606–617.
17. Dobson F, Hinman RS, Roos EM, et al. OARSI recommended performance-based tests to assess physical function in people diagnosed with hip or knee osteoarthritis. Osteoarthritis Cartilage 2013;21(8):1042–1052.
18. Choi L, Liu Z, Matthews CE, et al. Validation of accelerometer wear and nonwear time classification algorithm. Med Sci Sports Exerc 2011;43(2):357–364.
19. Troiano RP, Berrigan D, Dodd KW, et al. Physical activity in the United States measured by accelerometer. Med Sci Sports Exerc 2008;40(1):181–188.
20. Lorig K, Chastain RL, Ung E, et al. Development and evaluation of a scale to measure perceived self-efficacy in people with arthritis. Arthritis Rheum 1989;32(1):37–44.

21. Kohout FJ, Berkman LF, Evans DA, et al. Two shorter forms of the CES-D (Center for Epidemiological Studies Depression) depression symptoms index. J Aging Health 1993;5(2):179–193.
22. Bruce B, Fries JF. The Health Assessment Questionnaire (HAQ). Clin Exp Rheumatol 2005;23(5 suppl 39):S14–S18.
23. Reilly Associates. WPAI general information. 2019. Accessed October 13, 2020. http://www.reillyassociates.net/WPAI_General.html
24. Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. PharmacoEconomics 1993;4(5):353–365.
25. Singh JA, Bharat A, Khanna D, et al. Health care utilization in patients with gout: a prospective multicenter cohort study. BMC Musculoskelet Disord 2017;18(1):233.
26. Osteoarthritis Action Alliance. Walk With Ease post-program evaluation form. 2018. Accessed February 5, 2021. https://oaaction.unc.edu/wp-content/uploads/sites/623/2018/08/WWE-Post-Program-Survey-2.7.18.pdf
27. Silverstein RP, VanderVos M, Welch H, et al. Self-directed Walk With Ease workplace wellness program - Montana, 2015-2017. MMWR Morb Mortal Wkly Rep 2018;67(46):1295–1299.
28. Wen CP, Wai JPM, Tsai MK, et al. Minimum amount of physical activity for reduced mortality and extended life expectancy: a prospective cohort study. Lancet 2011;378(9798):1244–1253.
29. Steinman L, Chadwick K, Chavez Santos E, et al. Remote evidence-based programs for health promotion to support older adults during the COVID-19 pandemic and beyond: mixed methods outcome evaluation. JMIR Aging 2024;7:e52069.
30. Zimmerman ZE, Cleveland RJ, Kostic AM, et al. Walk with ease for knee osteoarthritis: a cost-effectiveness analysis. Osteoarthr Cartil Open 2023;5(3):100368.
31. Vilen LH, Altpeter M, Callahan LF. Overcoming barriers to *Walk With Ease* implementation in community organizations. Health Promot Pract 2022;23(4):708–717.
32. Wyatt B, Mingo CA, Waterman MB, et al. Impact of the Arthritis Foundation's Walk With Ease Program on arthritis symptoms in African Americans. Prev Chronic Dis 2014;11:E199.
33. Mazza NZ, Lanou AJ, Weisner S. Reach and impact of in-person and remote delivery formats of Walk with Ease. Inquiry 2023;60:469580231152314.
34. Waters LA, Galichet B, Owen N, et al. Who participates in physical activity intervention trials? J Phys Act Health 2011;8(1):85–103.
35. Duca LM, Helmick CG, Barbour KE, et al. State-specific prevalence of inactivity, self-rated health status, and severe joint pain among adults with arthritis - United States, 2019. Prev Chronic Dis 2022;19:E23.

Arthritis Care & Research
Vol. 77, No. 7, July 2025, pp 939–940

American College of Rheumatology
Empowering Rheumatology Professionals

LETTER

DOI 10.1002/acr.25499

Gout flare prophylaxis trials: comment on the article by Maher et al

To the Editor:

The risk of a gout flare is most significant when urate lowering is rapid, independent of the urate-lowering drug chosen.[1,2] The gradual dose escalation of urate-lowering therapy (ULT), regardless of the drug used, combined with anti-inflammatory prophylaxis, is recommended for the first three to six months to mitigate this risk.[3]

The article by Maher et al reviews the risk of gout flares when initiating or escalating urate-lowering drugs.[4] Several prophylaxis trials were included, including a trial of rilonacept, an interleukin-1 inhibitor.[5] However, the canakinumab trial—the largest randomized, controlled prophylaxis trial of 432 patients initiating ULT with allopurinol, comparing whether single doses of canakinumab ≥50 mg or 4 × 4 weekly doses of canakinumab provided superior prophylaxis compared with colchicine 0.5 mg/d, the current mainstay of prophylaxis—was omitted from this thoughtful review.[4,6] In this 16-week study, there was a 62% to 72% reduction in the mean number of flares per participant for patients in the canakinumab arms compared with the colchicine arm ($P \leq 0.0083$), and the proportion of participants who experienced ≥1 flare was significantly lower (15% to 27% vs 44%; $P < 0.05$). In addition, there was a 64% to 72% reduction in the risk of experiencing ≥1 flare for canakinumab doses ≥50 mg versus colchicine ($P \leq 0.05$). The canakinumab study findings support interleukin-1 as a pivotal mediator of gout inflammation and the role of interleukin-1 inhibitors in effectively preventing gout flares.[2,4–6]

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25499.

Naomi Schlesinger, MD
naomi.schlesinger@hsc.utah.edu
Jamie Dwyer, MD
University of Utah
Salt Lake City
Jeffrey Carson, MD
New Brunswick at Rutgers Biomedical and Health Sciences
New Brunswick, New Jersey
Luigi Brunetti, PhD, PharmD
Rutgers Ernest Mario School of Pharmacy
Piscataway, New Jersey

1. Schlesinger N, Etzel CJ, Greenberg J, et al. Gout prophylaxis evaluated according to the 2012 American College of Rheumatology Guidelines: analysis from the CORRONA Gout Registry. J Rheumatol 2016;43(5):924–930.
2. Schlesinger N. Treatment of chronic gouty arthritis: it is not just about urate-lowering therapy. Semin Arthritis Rheum 2012;42(2):155–165.
3. 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.
4. Maher D, Reeve E, Hopkins A, et al. Comparative risk of gout flares when initiating or escalating various urate-lowering therapy: a systematic review with network meta-analysis. Arthritis Care Res (Hoboken) 2024;76(6):871–881.
5. Mitha E, Schumacher HR, Fouche L, et al. Rilonacept for gout flare prevention during initiation of uric acid-lowering therapy: results from the PRESURGE-2 international, phase 3, randomized, placebo-controlled trial. Rheumatology (Oxford) 2013;52(7):1285–1292.
6. Schlesinger N, Mysler E, Lin HY, et al. Canakinumab reduces the risk of acute gouty arthritis flares during initiation of allopurinol treatment: results of a double-blind, randomized study. Ann Rheum Dis 2011l; 70(7):1264–1271.

DOI 10.1002/acr.25498

Reply

To the Editor:

We would like to thank Professor Schlesinger and colleagues for the interest and comment on our recently published systematic review with network meta-analysis of gout flares when initiating or escalating urate-lowering therapy (ULT).[1] As pointed out by the authors, trials of the interleukin-1 inhibitor canakinumab were not included in our review. Our strict inclusion and exclusion criteria (full details provided in the Supplementary Methods[1]) excluded phase 1 and 2 trials, such as the large randomized, controlled (phase 2) prophylaxis trial of canakinumab versus colchicine.[2] Various phase 3 and open-label extension trials investigating canakinumab for the treatment of acute flares and subsequent prevention of new flares did not meet our inclusion criteria.[3,4]

Canakinumab is currently approved by the US Food and Drug Administration and European Medicines Agency for the symptomatic treatment of gout flares in adults with frequent flares who cannot tolerate, are contraindicated, or do not benefit from nonsteroidal anti-inflammatory drugs and colchicine and for whom repeated courses of corticosteroids are not appropriate.[5,6] Although canakinumab prophylaxis trials are encouraging,[2] these somewhat limited approval criteria, which encompass treatment of an acute attack rather than prophylaxis after ULT initiation or intensification, coupled with the

findings of our review highlight that more research is required to better define both the benefits and harms of existing and newer medications for gout flare prophylaxis.[1] The unique dosing regimen of canakinumab may have significant implications for quality of life in a population for whom nonadherence is prevalent.[7]

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25498.

Dorsa Maher, BPharm Hons
dorsa.maher@mymail.unisa.edu.au
University of South Australia and SA Pharmacy, Adelaide South Australia, Australia
Emily Reeve, PhD
University of South Australia, Adelaide, South Australia, Australia and Monash University, Melbourne Victoria, Australia
Michael Wiese, PhD
University of South Australia Adelaide, South Australia, Australia

1. Maher D, Reeve E, Hopkins A, et al. Comparative risk of gout flares when initiating or escalating various urate-lowering therapy: a systematic review with network meta-analysis. Arthritis Care Research (Hoboken) 2024;76(6):871–81.
2. Schlesinger N, Mysler E, Lin HY, et al. Canakinumab reduces the risk of acute gouty arthritis flares during initiation of allopurinol treatment: results of a double-blind, randomised study. Ann Rheum Dis 2011; 70(7):1264–1271.
3. Schlesinger N, Alten RE, Bardin T, et al. Canakinumab for acute gouty arthritis in patients with limited treatment options: results from two randomised, multicentre, active-controlled, double-blind trials and their initial extensions. Ann Rheum Dis 2012;71(11):1839–1848.
4. Canakinumab in the treatment of gouty arthritis flare(s) and prevention of new flares in patients with chronic kidney disease (study withdrawn). ClinicalTrials.gov identifier: NCT01593527. Updated April 20, 2017. Accessed December 10, 2024. https://clinicaltrials.gov/study/NCT01593527
5. Novartis Pharmaceuticals Corporation. Medication guide: ILARIS® (canakinumab) injection. US Food and Drug Administration. 2023. Accessed December 20, 2024. https://www.accessdata.fda.gov/drugsatfda_docs/label/2023/125319s107lbl.pdf
6. Novartis Europharm Limited. Ilaris: EPAR - product information. European Medicines Agency. 2024. Accessed December 20, 2024. https://www.ema.europa.eu/en/documents/product-information/ilaris-epar-product-information_en.pdf
7. Scheepers LEJM, van Onna M, Stehouwer CDA, et al. Medication adherence among patients with gout: a systematic review and meta-analysis. Semin Arthritis Rheum 2018;47(5):689–702.