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. 11 November 2025

Arthritis Care & Research

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

Cover design: Sandra Pulmano

♾ This journal is printed on acid-free paper.

Arthritis Care & Research

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

VOLUME 77 • NOVEMBER 2025 • NO. 11

Cover image: The image on the cover (Choufani et al; pages 1293–1302) shows the global distribution of pivotal psoriatic arthritis clinical trials, with darker shades indicating greater trial activity.

Arthritis Care & Research
Vol. 77, No. 11, November 2025, pp 1285–1287
DOI 10.1002/acr.25578

American College of Rheumatology
Empowering Rheumatology Professionals

EDITORIAL

What Are the On-Ramps to Predict Psoriatic Arthritis? Does Monitoring C-Reactive Protein Levels Help?

Zheni Stavre[1] and Michael H. Weisman[2]

Psoriatic arthritis (PsA) is a chronic, immune-mediated disease with diverse clinical phenotypes affecting up to 1% of the population and up to 30% of patients with psoriasis (PsO).[1] PsA is associated with various comorbidities requiring recognition and monitoring by health care providers to improve outcomes and management.[2] Although effective treatments are available once PsA is diagnosed, there remains a lack of readily detectable biomarkers during the preclinical phase of the disease.

The preclinical phase of PsA is an entity that is extensively under investigation. Unlike rheumatoid arthritis, in which biomarkers in the form of cyclic citrullinated peptide antibody and rheumatoid factor are easily detectable in patient's blood up to 10 years before clinical diagnosis, PsA has no easily detectable methods to aid in prediction of disease development. This is despite an obvious heralding sign present in most patients with PsA, that of skin involvement in the form of PsO. Major efforts toward biomarker discovery to predict development of PsA have included microbiome analysis, metabolomic analysis, and gene analysis of whole blood and synovial fluid.[3–6] Efforts are also underway to identify a digital imaging biomarker for the detection of early PsA.[7,8] However, to date, no practical useful or accurate blood, synovium, or digital biomarkers exist to aid clinicians in early diagnosis of PsA. One serum biomarker, CXCL10, has been shown in a longitudinal study to be a predictive biomarker of PsA development, but it is not currently in clinical use because of high cost and practical technical considerations.[9]

The prospective cohort study by Eder et al[10] aimed to assess the utility of a widely used and available blood test, that of high-sensitivity C-reactive protein (hsCRP), in predicting development of PsA in patients with PsO. This study analyzed data collected from 589 patients with PsO who were observed prospectively from 2006 to 2019. Approximately 10% (n = 57) of these patients developed PsA during this period. hsCRP was measured only once at study enrollment in all patients; patients were then evaluated yearly by rheumatologists for development of PsA. The nested case-control study found significantly higher levels of baseline hsCRP in patients with arthralgias (mean hsCRP level 4.20 mg/L), patients with obesity (body mass index [BMI] >30, mean hsCRP level 4.75 mg/L), and female patients (mean hsCRP level 3.92 mg/L). Higher levels of hsCRP were predictive of future development of PsA in a multivariate analysis (hazard ratio 1.04, 95% confidence interval 1.01–1.07, $P = 0.007$). The authors concluded that higher levels of systemic inflammation, as measured through hsCRP, are associated with future development of PsA.[10]

These findings may appear intuitive or even minimally relevant considering the wide interpretation of an elevated value for CRP or the fact that no specific cutoff for hsCRP was found or proposed. However, we would like to argue that this study was a necessary and brilliant effort using information collected prospectively from an extremely well-characterized cohort using a simple blood test that is cheap and readily available. It is important to realize the difference between the more widely and clinically used CRP when assessing patients with PsA, as compared to hsCRP, to understand the significance that this study brings to the field. CRP is a marker of acute inflammation that rises and falls upon resolution of inflammation as well as a potential marker of chronic inflammation in rheumatic diseases. Nevertheless, in psoriatic disease, measurement of CRP alone does not always correlate with presence of disease or disease activity, as many patients with PsA will have a normal CRP value at diagnosis.[11] hsCRP is a more sensitive measurement of CRP, capable of detecting small changes at lower values and is currently primarily used by cardiologists to assess cardiovascular risk, with value cutoffs of <1 mg/L denoting low cardiovascular risk, 1 to 3 mg/L denoting average cardiovascular risk, and >3 mg/L denoting high cardiovascular risk.[12] For the practicing rheumatologist clinician, hsCRP occasionally finds its way into the medical records of patients with PsA after the wrong button under Epic orders is

[1]Zheni Stavre, MD: University of Massachusetts Chan Medical School, Worcester; [2]Michael H. Weisman, MD: Stanford University School of Medicine, Palo Alto, California.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25578.

Address correspondence via email to Zheni Stavre, MD, at zheni.stavre@umassmed.edu.

Submitted for publication May 18, 2025; accepted May 23, 2025.

clicked, but hsCRP is not routinely used or examined in the context of PsA. However, hsCRP does exactly what it promises; it detects with high sensitivity minor changes in chronic low-grade inflammation, such as those that can occur in patients with PsO developing PsA. It is known that patients with PsO have elevations in hsCRP levels compared to controls, but to date no other prospective longitudinal study has looked at rising hsCRP levels as predictive of PsA development.[13]

Situations known to cause prolonged elevations in CRP levels are also those that are associated with the development of cardiovascular disease, including elevated BMI, cigarette smoking, metabolic syndrome, diabetes mellitus, elevated triglyceride levels, and others.[12] It is possible, but clearly not known as of yet, that many of the lifestyle modifications, drugs, and nutrient supplementations that are widely disseminated in the public domain to influence the development of cardiac disease are also candidates for primary prevention of PsA. In light of this study, hsCRP becomes an obvious marker of low-grade systemic inflammation that can occur in patients with PsO as they progress to PsA; it could become a practical research tool to implement in carefully designed future studies for primary prevention of PsA in candidates with PsO.

In this issue of *Arthritis Care & Research*, the study by Eder et al,[10] despite hsCRP's apparent utility in predicting PsA from PsO, has its limitations. The first one is that the number of patients with PsA is too small to determine a cutoff value for PsA development. Measuring hsCRP over time, perhaps yearly or every six months in patients with PsO, and looking at its trajectory over time may turn out to be the way to go when it comes to early PsA diagnosis. Further, in this study, only 6% of the 568 patients were receiving biologic therapy at the time of hsCRP measurement, and of those who developed PsA, only 8.8% were taking a biologic. Is this reflective of the population of new patients with PsA that we see today? One could argue that the use of biologic therapy by dermatologists might mean that many patients with PsA would present to a rheumatologist after biologic therapy has failed. The absence of biologic management in this cohort leaves open the influence of this treatment tool to precisely identify the candidates for disease prevention. At the end of the day, the data from this study only give us a group perspective; how to use hsCRP in current decision-making for our clinic patients is still not there. In the future, it would be nice to see if hsCRP measurement that included a larger population of patients with PsO receiving biologic therapy would continue to be predictive of PsA development.

Lastly, food for thought: it is possible that with the boom in digital technology, machine learning, and data accessibility, hsCRP, if not on its own, will, along with one or a few other easy-to-collect markers, both common blood markers (neutrophil to lymphocyte ratio) and digital markers in the form of wearables or questionnaires (screening questionnaires for PsA and patient-reported outcome measures or other predictive tools such as the Psoriatic Arthritis Risk Estimation Tool,[14] to name a few), may turn out to be a key measure in predicting development of PsA in those with PsO. The field is moving very fast, and the current piece of data described in the article by Eder et al is an important component. For example, colleagues from the University of Rochester employing a humanized mouse model transgenic strain to examine PsO and PsA endotypes found that injection of both peripheral blood mononuclear cells and serum samples from patients with PsO and PsA was critical for the recapitulation of human skin and joint phenotypes.[15] Because serum factors were required for the experiments to succeed, the door remains open as to which and what kind of serum factors were necessary for the development of the PsA endotype. Much further work is needed to tie all of this together.

We would argue it is not quite time to start checking hsCRP in all patients with PsO for direct clinical care decision-making outside of the research experience. We could easily envision incorporating this value with screening questionnaires to identify those at high risk of PsA in carefully conducted observational studies in cohorts in addition to those of the Toronto experience. Dr Eder and colleagues have informed us as to the added value that measurement of hsCRP would bring to future research and to psoriatic disease diagnosis and treatment.

ACKNOWLEDGMENTS

The authors acknowledge limited use of large language models, including ChatGPT-4, for language editing of the first paragraph of this article and use of the Consensus AI Search Engine for Research to identify six references in paragraph two of this article.

AUTHOR CONTRIBUTIONS

Both authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr Stavre 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. Gladman DD, Antoni C, Mease P, et al. Psoriatic arthritis: epidemiology, clinical features, course, and outcome. Ann Rheum Dis 2005;64 (suppl 2):ii14–ii17.
2. Ogdie A, Schwartzman S, Husni ME. Recognizing and managing comorbidities in psoriatic arthritis. Curr Opin Rheumatol 2015;27(2): 118–126.
3. Scher JU, Ubeda C, Artacho A, et al. Decreased bacterial diversity characterizes the altered gut microbiota in patients with psoriatic arthritis, resembling dysbiosis in inflammatory bowel disease. Arthritis Rheumatol 2015;67(1):128–139.
4. Looby N, Roszkowska A, Reyes-Garcés N, et al. Serum metabolic fingerprinting of psoriasis and psoriatic arthritis patients using solid-

phase microextraction-liquid chromatography-high-resolution mass spectrometry. Metabolomics 2021;17(7):59.

5. Souto-Carneiro M, Tóth L, Behnisch R, et al. Differences in the serum metabolome and lipidome identify potential biomarkers for seronegative rheumatoid arthritis versus psoriatic arthritis. Ann Rheum Dis 2020;79(4):499–506.
6. Johnsson H, Cole J, McInnes IB, et al. Differences in transcriptional changes in psoriasis and psoriatic arthritis skin with immunoglobulin gene enrichment in psoriatic arthritis. Rheumatology (Oxford) 2024; 63(1):218–225.
7. Koehm M, Ohrndorf S, Foldenauer AC, et al. Fluorescence-optical imaging as a promising easy-to-use imaging biomarker to increase early psoriatic arthritis detection in patients with psoriasis: a cross-sectional cohort study with follow-up. RMD Open 2022;8(2):e002682.
8. Webster DE, Haberman RH, Perez-Chada LM, et al. Clinical validation of digitally acquired clinical data and machine learning models for remote measurement of psoriasis and psoriatic arthritis: a proof-of-concept study. J Rheumatol 2024;51(8):781–789.
9. Abji F, Lee KA, Pollock RA, et al. Declining levels of serum chemokine (C-X-C motif) ligand 10 over time are associated with new onset of psoriatic arthritis in patients with psoriasis: a new biomarker? Br J Dermatol 2020;183(5):920–927.
10. Eder L, Li X, Chandran V, et al. Association of higher levels of high-sensitivity C-reactive protein with future development of psoriatic arthritis in psoriasis: a prospective cohort study. Arthritis Care Res (Hoboken) 2025;77(11):1288–1292. https://doi.org/10.1002/acr.25539
11. Sokolova MV, Simon D, Nas K, et al. A set of serum markers detecting systemic inflammation in psoriatic skin, entheseal, and joint disease in the absence of C-reactive protein and its link to clinical disease manifestations. Arthritis Res Ther 2020;22(1):26.
12. Pearson TA, Mensah GA, Alexander RW, et al; Centers for Disease Control and Prevention; American Heart Association. Markers of inflammation and cardiovascular disease: application to clinical and public health practice: a statement for healthcare professionals from the Centers for Disease Control and Prevention and the American Heart Association. Circulation 2003;107(3):499–511.
13. Pirro M, Stingeni L, Vaudo G, et al. Systemic inflammation and imbalance between endothelial injury and repair in patients with psoriasis are associated with preclinical atherosclerosis. Eur J Prev Cardiol 2015;22(8):1027–1035.
14. Eder L, Lee KA, Chandran V, et al. Derivation of a multivariable Psoriatic Arthritis Risk Estimation Tool (PRESTO): a step towards prevention. Arthritis Rheumatol 2025;77(6):686–695.
15. Ritchlin CT, Rangel-Moreno J, Martino D, et al. Psoriatic arthritis subtypes are phenocopied in humanized mice. JCI Insight 2024;9(15): e178213.

Arthritis Care & Research
Vol. 77, No. 11, November 2025, pp 1288–1292
DOI 10.1002/acr.25539

American College of Rheumatology
Empowering Rheumatology Professionals

BRIEF REPORT

Association of Higher Levels of High-Sensitivity C-Reactive Protein With Future Development of Psoriatic Arthritis in Psoriasis: A Prospective Cohort Study

Lihi Eder,[1] Xianwei Li,[2] Vinod Chandran,[3] Cheryl F. Rosen,[4] Richard J. Cook,[2] and Dafna D. Gladman[3]

Objective. We aimed to assess whether high-sensitivity C-reactive protein (hsCRP) levels could predict the development of psoriatic arthritis (PsA) in patients with psoriasis.

Methods. We analyzed data from a prospective cohort of patients with psoriasis without PsA at enrollment. Participants were assessed annually by a rheumatologist for signs and symptoms of PsA. Information on patient demographics, psoriasis features, medications and musculoskeletal symptoms was collected. hsCRP levels were measured in serum samples collected at baseline using standard commercial assays. The association between hsCRP levels and risk of development of PsA was assessed using multivariable Cox proportional hazards model adjusted for age, sex, psoriasis severity and duration, nail lesions, body mass index (BMI), fatigue, and medication use.

Results. A total of 589 patients with psoriasis observed from 2006 to 2019 were analyzed. During the follow up period, 57 patients developed PsA. The mean level of hsCRP was 3.1 ± 5.5 mg/L (hsCRP levels in patients with incident PsA, 5.4 ± 13.1 mg/L). Significantly higher levels of hsCRP at baseline were found in patients with arthralgia, obesity, and in women. Higher hsCRP levels were associated with future development of PsA in multivariable analyses (hazard ratio 1.04; 95% confidence interval 1.01–1.07; $P = 0.007$). Similar effect size was seen in men and women. No significant interaction was found between hsCRP and sex or BMI.

Conclusion. Higher levels of systemic inflammation, as measured by hsCRP levels, are associated with future development of PsA.

INTRODUCTION

Approximately 2%–3% of patients with psoriasis develop psoriatic arthritis (PsA) each year.[1] However, despite advancement in understanding the preclinical phases of PsA and the development of a novel risk prediction score,[2,3] there remains a significant gap in the availability of simple and affordable biomarkers for predicting PsA development among patients with psoriasis. High-sensitivity C-reactive protein (hsCRP) is widely used in the clinical setting to assess systemic inflammation, particularly in evaluating cardiovascular risk. However, its potential to predict the progression to PsA in individuals with psoriasis remains unclear.

Research from various cohorts of patients with psoriasis suggests that systemic and local subclinical inflammation is present in many individuals long before they progress to PsA.[4] Several risk factors for PsA, such as obesity and extensive psoriasis, are indicative of systemic or local inflammation and immune system activation.[1,5–6] The presence of nail lesions, another risk factor for PsA, has been linked with musculoskeletal inflammation at the entheses because of their anatomic proximity.[7] Nonspecific musculoskeletal symptoms such as arthralgia, which predict progression to PsA, may reflect subclinical musculoskeletal inflammation that has not yet manifested as distinct physical examination findings.[8] Additional evidence from imaging studies reveals frequent subclinical inflammation at the entheses and joints in patients with psoriasis without PsA.[9,10] Given this growing body of evidence supporting the presence of systemic inflammation in preclinical PsA, a simple, affordable and accessible blood biomarker like hsCRP, could significantly enhance the identification of individuals at high risk for progression to PsA. We aimed to investigate the association between hsCRP levels and the risk of developing PsA in a prospective cohort of patients with psoriasis.

PATIENTS AND METHODS

Setting. This study was nested in the ongoing University of Toronto Psoriasis Cohort, which began in 2006 as a prospective longitudinal cohort study aimed at identifying risk factors for the development of PsA. All patients enrolled in the cohort have a

SIGNIFICANCE & INNOVATIONS

- Despite advancement in understanding the preclinical phases of psoriatic arthritis (PsA), simple and affordable biomarkers for predicting PsA development among patients with psoriasis are lacking.
- Our study demonstrated that elevated levels of high-sensitivity C-reactive protein (hsCRP), a widely used biomarker of systemic inflammation, are associated with future risk of PsA development in patients with psoriasis, independent of other known PsA risk factors.
- The study findings underscore the potential of hsCRP as a valuable biomarker for identifying patients at high risk of developing PsA in both dermatology and rheumatology settings.
- Early identification of at-risk populations could facilitate timely interventions that may prevent the progression to PsA.

dermatologist-confirmed diagnosis of psoriasis and undergo a pre-enrollment assessment by a rheumatologist to exclude any history or presence of inflammatory arthritis. The cohort has been described in detail in previous publications.[1] For this study, we used data collected from January 2006 to December 2019. Patients who lacked follow up study visits or did not have a serum sample in our biobank were excluded from the study. The study received approval from the University Health Network Research Ethics Board, and all participants provided informed consent.

Data collection. All study participants underwent an annual assessment by a rheumatologist, who evaluated whether patients had developed PsA since the previous visit and collected information on potential risk factors. Data collection followed standard protocols, capturing information on lifestyle habits, family medical history, musculoskeletal symptoms, comorbidities, medications, and findings from skin and nail examination.

Measurement of hsCRP. Blood samples were collected at each annual visit, and serum samples are stored in our biobank. The first available serum sample was identified, and hsCRP levels were measured in our hospital laboratory using standard commercial kits. For each patient, hsCRP levels were measured only once.

Case definition. At each visit, a comprehensive assessment of PsA symptoms and signs was performed by a rheumatologist expert in PsA. The diagnosis of PsA was based on clinical findings with imaging modalities ordered only when clinically necessary to investigate abnormalities suggestive of PsA. The diagnosis of PsA was determined by at least two rheumatologists after reviewing all available clinical, laboratory, and imaging data.

To address loss to follow up, we reached out to patients who missed two or more consecutive annual visits to determine if they were alive and whether they have consulted a physician for musculoskeletal symptoms. We also reviewed all relevant medical records from rheumatologists and other specialists (available in the patient electronic medical records) to ascertain if a new diagnosis of PsA had been made.

Statistical analysis. The empirical distribution of each continuous variable was described by its mean and SD and by frequencies and percentages for categorical variables. For descriptive purposes, differences in baseline hsCRP levels across predefined patient groups were analyzed using the *t*-test. The number of person-years at risk of developing PsA was calculated as the time between the date of enrollment in the cohort and the date of the last visit or the date of PsA diagnosis, whichever came first. The association between baseline levels of hsCRP levels and the risk of developing PsA was assessed using multivariable Cox proportional hazards models adjusted for age, sex, psoriasis duration, psoriasis area and severity index score, nail lesions, body mass index (BMI), functional assessment of chronic illness fatigue score, treatment with biologics, and treatment with systemic nonbiologics or phototherapy; the duration of psoriasis disease at the time the hsCRP level was assessed was treated as a left-truncation time. We adjusted for these variables because of their established link with PsA risk, with most included in the Psoriatic Arthritis Risk Estimation Tool (PRESTO) tool.[2] Subgroup analyses were also performed by sex to determine whether the effect size differed between men and women; this was assessed by testing the interaction term involving hsCRP and sex with a similar interaction assessed for BMI. $P > 0.05$ was considered statistically significant. Missing data were dealt with using multiple imputations (via PROC MI and PROC MIANALYZE in SAS) based on predictive mean matching using

Supported in part by a New Investigator Grant from the Physician Services Incorporated Foundation. Dr Eder's work was supported by Canada Research Chair in Inflammatory Rheumatic Diseases (Tier 2). Dr Chandran's work was supported by a salary award from the Department of Medicine, University of Toronto. The Gladman-Krembil Psoriatic Disease program at Toronto Western Hospital is supported by the Krembil Foundation.

[1]Lihi Eder, MD, PhD: Women's College Hospital and University of Toronto, Toronto, Ontario, Canada; [2]Xianwei Li, MMath, Richard J. Cook, PhD: University of Waterloo, Waterloo, Ontario, Canada; [3]Vinod Chandran, MBBS, MD, DM, PhD, Dafna D. Gladman, MD: University of Toronto and University Health Network, Toronto, Ontario, Canada; [4]Cheryl F. Rosen, MD: Toronto Western Hospital, 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.25539).

Author disclosures and graphical abstract are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25539.

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

Submitted for publication October 29, 2024; accepted in revised form March 19, 2025.

the full conditional specification; five complete datasets were generated by imputation.

RESULTS

A total of 702 patients with psoriasis without clinical evidence of PsA were enrolled in the Toronto Psoriasis Cohort and observed from January 1, 2006, to December 31, 2019. Of these, 67 patients were excluded from the analysis due to having only a single visit and contributing no follow up data. An additional 46 patients were excluded due to lack of serum samples. Ultimately, 589 patients were included in the analysis, with 57 patients developing PsA during the follow up period (Supplementary Figure 1). The mean duration of follow up was 7.5 years, yielding a crude PsA incidence rate of 1.2 patients per 100 patient years. The mean age of study participants was 47.3 ± 13.5 years (43.1% women, Table 1), and mean duration of psoriasis was 16.2 ± 14.4 years. The mean hsCRP level was 3.1 ± 5.5 mg/L, with those who developed PsA showing higher baseline hsCRP levels (5.4 ± 13.1 mg/L). hsCRP distribution by diagnosis can be found in Supplementary Figure 2.

Baseline hsCRP levels were significantly higher in patients with arthralgia (4.2 ± 8.53 mg/L vs 2.71 ± 3.67 mg/L), those with obesity (4.75 ± 4.95 mg/L vs 2.45 ± 5.61 mg/L), and in women (3.92 ± 7.13 mg/L vs 2.51 ± 3.68 mg/L) (Supplementary Table 1). No significant differences in hsCRP levels were observed across the remaining groups.

Univariate regression analysis revealed that higher hsCRP levels were significantly associated with future development of PsA, with a 3% increase in PsA risk for each 1 mg/L increase in hsCRP (hazard ratio [HR] 1.03; 95% confidence interval [CI] 1.01–1.05; P = 0.002; Table 2). This association remained significant in the multivariable regression analysis, with a 4% increase in PsA risk for each 1 mg/L increase in hsCRP (HR 1.04; 95% CI 1.01–1.07; P = 0.007). The effect size was consistent across men and women (Figure 1). No significant interaction was found between hsCRP and sex or BMI (data not shown).

DISCUSSION

Our study showed that a higher level of hsCRP, a biomarker of systemic inflammation, is independently associated with increased risk of the development of PsA among patients with psoriasis regardless of other known risk factors. This finding underscores the potential of hsCRP as a valuable biomarker for identifying patients at high risk of developing PsA in both dermatology and rheumatology settings.

Although rheumatologists often use CRP levels to assess disease activity, measuring hsCRP offers a more sensitive assessment of the protein level at the lower end of the spectrum. This distinction may be important because CRP levels tend to be lower in patients with PsA compared with other inflammatory rheumatic conditions, such as rheumatoid arthritis and polymyalgia rheumatica, which are more influenced by interleukin-6–driven inflammatory pathways. Notably, CRP levels are elevated in only

Table 1. Baseline patient characteristics*

Variable	All, n (%)	Developed PsA,[a] n (%)	No PsA,[b] n (%)
Age, mean ± SD, y	47.3 ± 13.5	48.6 ± 12.2	47.2 ± 13.7
Sex, female	254 (43.1)	23 (40.4)	231 (43.4)
Race and ethnicity			
White	449 (76.2)	48 (84.2)	401 (75.4)
South Asian	41 (7)	2 (3.5)	39 (7.3)
Chinese	28 (4.8)	3 (5.3)	25 (4.7)
Filipino	16 (2.7)	0	16 (3)
Middle Eastern	13 (2.2)	1 (1.8)	12 (2.3)
East Asian/Japanese	11 (1.9)	0	11 (2.1)
Black	10 (1.7)	1 (1.8)	9 (1.8)
Other	21 (3.6)	2 (3.5)	19 (3.6)
Duration of psoriasis, mean ± SD, y	16.2 ± 14.4	20.2 ± 15.4	15.7 ± 14.2
hsCRP, mean ± SD, mg/L	3.1± 5.5	5.4 ± 13.1	2.9 ± 3.8
PASI, mean ± SD	5.2 ± 5.8	5.2 ± 4.8	5.2 ± 5.9
Nail lesions, yes	272 (46.2)	31 (54.4)	241 (45.3)
BMI, mean ± SD	27.9 ± 5.9	28 ± 5.4	27.9 ± 5.9
Patient pain score (0–10), mean ± SD	1.5 ± 2.2	1.7 ± 2.2	1.5 ± 2.2
FACIT fatigue score, mean ± SD	44.7 ± 7.1	42.1 ± 8	45 ± 7
Current biologic therapy, yes	35 (5.9)	5 (8.8)	30 (5.6)
Current nonbiologic systemic therapy for psoriasis or UV therapy, yes	396 (67.2)	33 (57.9)	363 (68.2)

* N = 589 for total participants. BMI, body mass index; FACIT, functional assessment of chronic illness; hsCRP, high-sensitivity C-reactive protein; PASI, psoriasis area and severity index; PsA, psoriatic arthritis; UV, ultraviolet.
[a] n = 57.
[b] n = 532.

Table 2. The association between hsCRP levels and the development of PsA using Cox proportional hazards models*

Variable	Univariate model		Multivariate model	
	HR (95% CI)	*P* value	HR (95% CI)	*P* value
hsCRP level, mg/L	**1.03 (1.01–1.05)**	**0.002**	**1.04 (1.01–1.06)**	**0.008**
Age, y	1.01 (0.99–1.03)	0.48	1.00 (0.97–1.02)	0.76
Sex, female	1.12 (0.66–1.91)	0.66	1.23 (0.70–2.16)	0.46
Psoriasis duration, y	1.02 (1.01–1.03)	0.04	1.02 (1.00–1.04)	0.02
PASI	0.99 (0.95–1.04)	0.79	1.00 (0.95–1.05)	0.93
Nail lesions (yes)	1.30 (0.77–2.20)	0.32	1.33 (0.76–2.32)	0.32
BMI	1.00 (0.96–1.04)	0.95	0.98 (0.93–1.03)	0.37
Patient pain (0–10)	1.05 (0.94–1.19)	0.35	0.98 (0.86–1.12)	0.78
FACIT fatigue	0.95 (0.92, 0.99)	0.01	0.95 (0.91–0.99)	0.01
Biologic therapy	1.98 (0.79–4.97)	0.14	1.72 (0.66–4.43)	0.27
Nonbiologic systemic therapy or phototherapy	0.50 (0.29–0.85)	0.01	0.44 (0.25–0.78)	0.005

* Bold values indicate statistical significance. N = 589 for total participants, n = 57 for events. BMI, body mass index; CI, confidence interval; FACIT, functional assessment of chronic illness; HR, hazard ratio; hsCRP, high-sensitivity C-reactive protein; PASI, psoriasis area and severity index; PsA, psoriatic arthritis.

50% of the patients with PsA, despite clinical evidence of active disease.[11] This reflects the variability in the biologic mechanisms underlying PsA. Despite these limitations, CRP provides valuable prognostic information regarding the risk of progression of joint damage, response to therapy and development of comorbidities, such as cardiovascular events.[12,13]

To our knowledge, this is the first publication that demonstrated an association between hsCRP levels and risk of PsA conversion. Because of the scarcity of prospectively observed psoriasis cohorts that collect biosamples, few studies to date have evaluated other biomarkers as predictors of PsA conversion.[14] C-X-C motif chemokine ligand 10 is the only protein biomarker that was found to predict the conversion to PsA in a longitudinal cohort study.[15] Other biomarker studies that assessed biomarkers for psoriasis progression were cross-sectional, thus the identified biomarkers may merely reflect disease activity rather than serve as indicators for future development of PsA.[14] Our study adds hsCRP as a potential useful predictive marker, with the advantage of accessibility, standardization, and low cost.

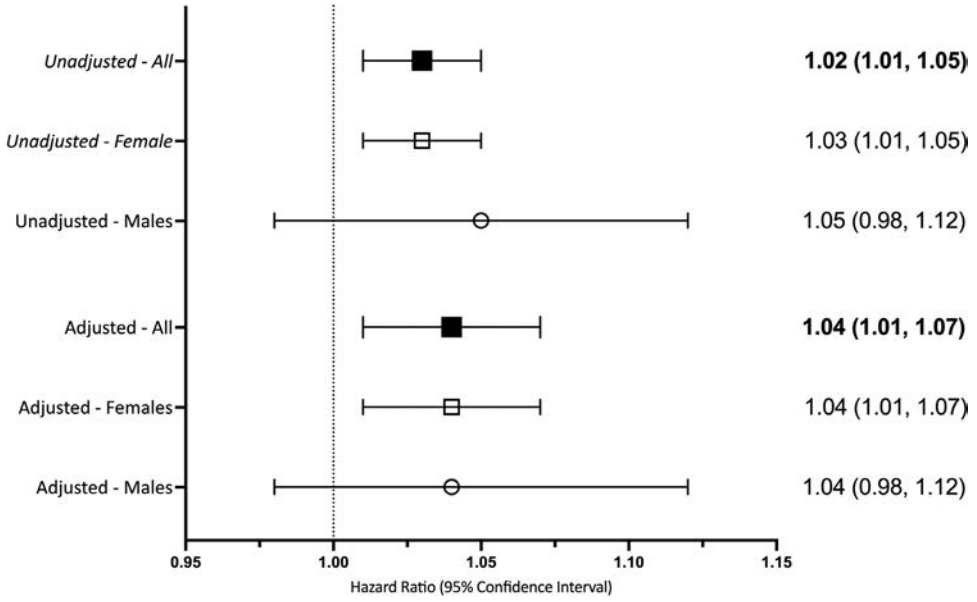


Figure 1. The association between hsCRP and the development of PsA by Cox proportional hazard models. The figure shows hazard ratios and their corresponding 95% confidence intervals of each model including univariate and multivariable models and sex-specific models. Each model was adjusted for age, sex, psoriasis duration, PASI, nail lesions, BMI, pain, FACIT fatigue, treatment with biologics, and treatment with nonbiologic systemic therapy or phototherapy. BMI, body mass index; FACIT, functional assessment of chronic illness; hsCRP, high-sensitivity C-reactive protein; PASI, psoriasis area and severity index; PsA, psoriatic arthritis.

However, it is important to note that our study provides only preliminary information on the predictive value of hsCRP levels. Because of the small number of patients developing PsA and the study design, we could not evaluate important predictive metrics, such as discriminative ability (eg, area under the receiver operating characteristic curve), or determine a useful cut-off for hsCRP levels. In addition, we could only evaluate hsCRP at a single time point—a Cox model involving time-dependent covariates of hsCRP levels over time would be preferred. It is possible that hsCRP levels change as a function of time before the onset of PsA, which would affect its prognostic value. It is unlikely that a single biomarker such as hsCRP will offer sufficient predictive capacity on its own. Adding hsCRP to a clinical prediction tool for PsA like PRESTO[2] may improve its performance. Unfortunately, because of the limited number of converters with available serum samples within the PRESTO prediction window, we could not explore this question. Nevertheless, our study demonstrates that hsCRP levels predict conversion to PsA independently of established risk factors, including those encompassed in the PRESTO tool.

In summary, elevated hsCRP levels are associated with an increased risk of progression to PsA in patients with psoriasis. This finding underscores the presence of subclinical inflammation before the development of objective PsA signs and highlights the potential of hsCRP as a biomarker for identifying at-risk populations. This could facilitate early interventions that may prevent a patient's progression to PsA.

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. Eder L, Haddad A, Rosen CF, et al. The incidence and risk factors for psoriatic arthritis in patients with psoriasis: a prospective cohort study. Arthritis Rheumatol 2016;68(4):915–923.
2. Eder L, Lee KA, Chandran V, et al. Derivation of a multivariable psoriatic arthritis risk estimation tool (PRESTO): a step towards prevention. Arthritis Rheumatol Published online August 9, 2023. https://doi.org/10.1002/art.42661.
3. Perez-Chada LM, Haberman RH, Chandran V, et al. Consensus terminology for preclinical phases of psoriatic arthritis for use in research studies: results from a Delphi consensus study. Nat Rev Rheumatol 2021;17(4):238–243.
4. Faustini F, Simon D, Oliveira I, et al. Subclinical joint inflammation in patients with psoriasis without concomitant psoriatic arthritis: a cross-sectional and longitudinal analysis. Ann Rheum Dis 2016;75(12):2068–2074.
5. Love TJ, Zhu Y, Zhang Y, et al. Obesity and the risk of psoriatic arthritis: a population-based study. Ann Rheum Dis 2012;71(8):1273–1277.
6. Ogdie A, Shin DB, Love TJ, et al. Body surface area affected by psoriasis and the risk for psoriatic arthritis: a prospective population-based cohort study. Rheumatology (Oxford) 2022;61(5):1877–1884.
7. Aydin SZ, Castillo-Gallego C, Ash ZR, et al. Ultrasonographic assessment of nail in psoriatic disease shows a link between onychopathy and distal interphalangeal joint extensor tendon enthesopathy. Dermatology 2012;225(3):231–235.
8. Eder L, Polachek A, Rosen CF, et al. The development of psoriatic arthritis in patients with psoriasis is preceded by a period of nonspecific musculoskeletal symptoms: a prospective cohort study. Arthritis Rheumatol 2017;69(3):622–629.
9. Eder L, Jayakar J, Thavaneswaran A, et al. Is the Madrid Sonographic Enthesitis Index useful for differentiating psoriatic arthritis from psoriasis alone and healthy controls? J Rheumatol 2014;41(3):466–472.
10. Zabotti A, McGonagle DG, Giovannini I, et al. Transition phase towards psoriatic arthritis: clinical and ultrasonographic characterisation of psoriatic arthralgia. RMD Open 2019;5(2):e001067.
11. Sokolova MV, Simon D, Nas K, et al. A set of serum markers detecting systemic inflammation in psoriatic skin, entheseal, and joint disease in the absence of C-reactive protein and its link to clinical disease manifestations. Arthritis Res Ther 2020;22(1):26.
12. Houttekiet C, de Vlam K, Neerinckx B, et al. Systematic review of the use of CRP in clinical trials for psoriatic arthritis: a concern for clinical practice? RMD Open 2022;8(1):e001756.
13. Elmamoun M, Leung YY, O'Sullivan D, et al. Using acute-phase reactants to inform the development of instruments for the updated psoriatic arthritis core outcome measurement set. J Rheumatol 2019;46(3):266–273.
14. Ramessur R, Corbett M, Marshall D, et al. Biomarkers of disease progression in people with psoriasis: a scoping review. Br J Dermatol 2022;187(4):481–493.
15. Abji F, Lee KA, Pollock RA, et al. Declining levels of serum chemokine (C-X-C motif) ligand 10 over time are associated with new onset of psoriatic arthritis in patients with psoriasis: a new biomarker? Br J Dermatol 2020;183(5):920–927.

Arthritis Care & Research
Vol. 77, No. 11, November 2025, pp 1293–1302
DOI 10.1002/acr.25574

American College of Rheumatology
Empowering Rheumatology Professionals

Race, Ethnicity, and Geographic Diversity in Pivotal Psoriatic Arthritis Clinical Trials: Further Progress Needed

Mathieu Choufani,[1] Joerg Ermann,[2] and Niti Goel[3]

Objective. We assessed race, ethnicity, and geographic diversity in pivotal trials of biologic and targeted synthetic disease-modifying antirheumatic drugs approved for psoriatic arthritis (PsA) in the United States.

Methods. We conducted a descriptive epidemiologic study, examining the reporting and representation of race, ethnicity, and geographic distribution of trial sites in pivotal PsA trials using data from journal publications, the Drugs@FDA database, and ClinicalTrials.gov.

Results. We identified 29 pivotal PsA trials for 16 targeted therapies with start dates between 2000 and 2019. Race data were reported in 93% of trials. Race reporting was highest in journal publications (86%); however, among these, 46% reported only the proportion of White participants. Ethnicity data were available for 41% of trials, primarily from ClinicalTrials.gov, with improved reporting in recent years. Among 14,165 participants in 27 trials with race data, 92% were White, 7% were Asian, and fewer than 1% were Black, American Indian/Alaska Native, or Native Hawaiian/Pacific Islander. Among 8,105 participants in 12 trials with ethnicity data, 11% were Hispanic or Latino individuals. Location data were available for 26 trials. In the United States, trial activity was highest in Texas (24 trials), Florida and Pennsylvania (22 each), and California (20), strongly correlating with state population size ($r = 0.78$, $P < 0.0001$). Globally, trial sites were identified in 55 countries, primarily in North America and Europe, with fewer sites in Asia, Africa, and Latin America.

Conclusion. Racial and ethnic diversity in pivotal PsA trials remains limited. Tailored and multifaceted strategies are needed to improve participant representativeness in future trials.

INTRODUCTION

Randomized controlled trials (RCTs) are the cornerstone of evidence-based medicine, providing critical data on the efficacy and safety of new treatments.[1,2] Participant diversity in RCTs is crucial to ensure the generalizability of findings and understand potentially varied responses in patient subsets defined by age, sex, race, or ethnicity.[3] However, RCTs often fail to include diverse populations.[4]

Psoriatic disease, including psoriatic arthritis (PsA), is a significant worldwide health concern affecting individuals across all racial and ethnic groups.[5] In the United States, a recent claims-based study reported PsA prevalence rates ranging from 0.13% to 0.19% in Asian individuals, 0.04% to 0.19% in Black individuals, 0.09% to 0.30% in Hispanic individuals, and 0.19% to 0.34% in White individuals.[6] PsA prevalence is likely underestimated in non-White populations due to disparities in health care access and higher rates of undiagnosed disease.[7] Beyond prevalence, different racial and ethnic groups also experience varying disease manifestations and disease burdens in PsA.[8–10] There remains a paucity of data on the therapeutic effects of PsA treatments across different demographic groups.

Recognizing the importance of diversity in RCTs to ensure that results are applicable across diverse populations in clinical practice, organizations such as the Institute of Medicine and the US Food and Drug Administration (FDA) have made significant efforts to improve the recruitment of more diverse populations in clinical trials.[3,11–13] However, there are limited studies that examine the racial and ethnic diversity of participants in PsA trials. In this study, we analyzed race and ethnicity reporting for pivotal PsA trials using data from three primary sources: journal publications, the FDA's online database (Drugs@FDA), and

[1]Mathieu Choufani, MD: Brigham and Women's Hospital, Boston, Massachusetts; [2]Joerg Ermann, MD: Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts; [3]Niti Goel, MD: Caduceus Biomedical Consulting, LLC, and Duke University School of Medicine, Durham, North Carolina.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25574).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25574.

Address correspondence via email to Niti Goel, MD, at cadbio2019@gmail.com.

Submitted for publication February 24, 2025; accepted in revised form May 8, 2025.

SIGNIFICANCE & INNOVATIONS

- Significant improvements in race and ethnicity reporting were demonstrated in recent years on ClinicalTrials.gov. Journal publications and the US Food and Drug Administration need to make similar efforts to ensure complete and consistent reporting of participant racial and ethnic demographics.
- White participants constituted the overwhelming majority in pivotal psoriatic arthritis (PsA) trials, with other racial and ethnic groups remaining significantly underrepresented.
- Recent increases in trial sites in East and Southeast Asia may have contributed to improved Asian participant representation. The rise in Eastern European trial sites increased overall participation, though it has not substantially diversified overall trial populations.
- Tailored and multifaceted recruitment efforts addressing patient-, provider-, and system-level barriers are needed to improve participant diversity in PsA clinical trials.

ClinicalTrials.gov. Additionally, we assessed the representation of racial and ethnic groups in these studies and explored the geographic distribution of PsA trial sites from both US and global perspectives. Our objective was to identify gaps in race and ethnicity reporting and highlight the need for greater diversity in PsA RCTs.

MATERIALS AND METHODS

Study design and data sources. We conducted a descriptive epidemiologic study to assess the reporting of race and ethnicity and to describe the racial, ethnic, and geographic diversity of participants in pivotal clinical trials of biologic disease-modifying antirheumatic drugs (bDMARDs) and targeted synthetic DMARDs (tsDMARDs) approved for the treatment of PsA in the United States. Trial data were obtained from three primary public sources: journal publications, Drugs@FDA, and ClinicalTrials.gov. State population and demographic data were obtained from Census.gov.

Identification of FDA-approved therapies and pivotal trials. We first established a list of all bDMARDs and tsDMARDs that received FDA approval for the treatment of PsA up to October 2024 (Suppl. Figure 1). For each drug, we searched Drugs@FDA, reviewing the FDA Summary Basis of Approval (SBoA), if available, and the FDA label specific to the PsA indication to identify the pivotal trials. SBoAs are structured based on drug indications rather than individual trials, meaning that a single SBoA may include data from one or multiple trials. We did not submit direct data requests to the FDA if an SBoA was not publicly available on Drugs@FDA. Trials were considered "pivotal" if they were explicitly labeled as such in the FDA SBoA or if they were referenced in the FDA label for the PsA indication. We also used Drugs@FDA to obtain the year of approval of each drug for the PsA indication and to identify the relevant prescribing information, clinical trials, and trial characteristics, which were then used to locate the corresponding National Clinical Trial number for each pivotal RCT on ClinicalTrials.gov. We then reviewed the trial's associated "Publications" section to identify the primary publication for each trial, examining the full text of the listed publications to identify the publication that reported the primary analysis of the RCT. A second reviewer, who has extensive knowledge of pivotal PsA trials, independently verified the selection of pivotal trials and related publications to ensure consistency and accuracy.

Data extraction and synthesis. For each pivotal trial, we manually extracted data on age, sex, race, and ethnicity from the three primary sources: journal publications, Drugs@FDA (specifically FDA SBoAs), and ClinicalTrials.gov (Suppl. Figure 1). Location data were extracted from ClinicalTrials.gov because this source offered the most detailed information on trial site locations. Data from the three sources were compiled and curated to create a comprehensive data set for each trial. Discrepancies fell into three categories: (1) missing data in one or more sources, (2) differences in how categories were grouped (eg, combining Black and Native American as "other" vs. reporting separately), and (3) rare differences in participant counts across sources. These discrepancies were resolved by cross-checking all sources and selecting the most consistent and granular data based on author consensus. Data procurement was performed by MC and reviewed by JE and NG. This process took place from September to December 2024.

Variables of interest. We recorded age as mean ± SD, along with the number of participants aged ≥65 years when reported. Sex, as recorded in the original sources, was reported as a binary variable (female or male) across all trials. We categorized race following the 1997 Office of Management and Budget (OMB) minimum categories,[14] which include American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Pacific Islander, and White. We combined "multiple" and "other" into a single category in our data set because some of the source data had grouped these categories together. Ethnicity data were recorded as "Hispanic or Latino" and "not Hispanic or Latino" when available. The updated 2024 OMB[14] reporting policy was not applied because all pivotal trials analyzed predated this update. When enumerating participants, we recognize that these participants may not be unique because some individuals could have participated in more than one trial. For geographic data, we recorded the number of trials with at least one enrollment site in each US state, and for countries other than the United States, we recorded the number of trials with at least one enrollment site in each country. States and countries were

grouped into four categories: high trial activity (14–26 trials), moderate trial activity (7–13 trials), low trial activity (1–6 trials), and no trial activity (0 trials). High activity (14–26 trials) corresponded to participation in more than 50% of the trials, moderate activity (7–13 trials) corresponded to participation in 25% to 50%, and low activity (1–6 trials) corresponded to participation in fewer than 25%.

Analytic methods. We used descriptive statistics to summarize trial characteristics, including participant demographics, geographic distribution, and reporting practices for race and ethnicity. The data were grouped into four time periods based on the start dates of the trials: 2000 to 2005, 2006 to 2010, 2011 to 2015, and 2016 to 2020. We assessed the correlation between state population and clinical trial activity using Spearman correlation analysis to account for the nonnormal distribution of the data (Shapiro–Wilk $P < 0.05$). The Spearman correlation coefficient (r), 95% confidence interval (CI), and P value were reported. Descriptive statistics were performed using Microsoft Excel, and the Spearman correlation coefficient and bar charts were generated using GraphPad Prism (version 10.4.1). We created the US and world map figures using MapChart (https://www.mapchart.net/).

Ethical considerations. This study used publicly available, deidentified data from journal publications, Drugs@FDA, and ClinicalTrials.gov, with no direct interaction with human participants. Consequently, ethics approval or informed consent was not required. The study complies with the data use policies of journal publishers, Drugs@FDA, and ClinicalTrials.gov.

RESULTS

Summary of pivotal PsA trials. A total of 29 pivotal PsA trials with start dates from April 2000 to April 2019 were identified and analyzed (Suppl. Table 1), encompassing 14,845 participants. These trials investigated the therapeutic efficacy of 16 drugs with different mechanisms of action; there were seven trials (24%) for tumor necrosis factor inhibitors, seven trials for interleukin-17A (IL-17A) or IL-17A/F inhibitors, four trials (14%) each for JAK inhibitors and IL-23 inhibitors, three trials (10%) for phosphodiesterase 4 inhibitors, and two trials each for CTLA-4Ig and IL-12/23 inhibitors. Nearly all were phase 3 trials (28 trials, 97%), and only one trial (3%) was classified as a phase 2 trial.

Race and ethnicity reporting in pivotal trials. Race data were available from at least one of the three data sources (journal publications, Drugs@FDA, ClinicalTrials.gov) for 93% (27 of 29) of trials, with reporting remaining relatively consistent over the 20-year period (Table 1). Ethnicity data were less frequently reported and available for only 41% (12 of 29) of all trials. However, a notable improvement in ethnicity reporting was observed over time, reaching 100% in the 2016 to 2020 period.

Among the three data sources, journal publications reported race data most frequently (86%, 24 of 28 trials) (Figure 1A), though 46% of them (11 of 24 trials) mentioned only the proportion of White participants. Ethnicity reporting in journal publications was limited to 7% (2 of 28 trials) overall, both instances occurring in the 2016 to 2020 period. FDA SBoAs were available online for 55% (16 of 29) of trials, all of which included race data (Figure 1A). Ethnicity reporting in SBoAs was sparse—only 5 of

Table 1. Race and ethnicity reporting by data source*

	All trials	2000–2005[a]	2006–2010	2011–2015	2016–2020[b]
Trials, n	29	5	7	9	8
Data from any source[c]					
Race reported	27 (93)	4 (80)	7 (100)	8 (89)	8 (100)
Ethnicity reported	12 (41)	1 (20)	0	3 (33)	8 (100)
Journal publications (n = 28 trials; 1 trial unpublished[d])					
Race reported	24 (86)	3 (75)[d]	5 (71)	8 (89)	8 (100)
Ethnicity reported	2 (7)	0	0	0	2 (25)
Only White race reported	11/24 (46)	3/3 (100)	2/5 (40)	3/8 (38)	3/8 (38)
Drugs@FDA					
Race reported	16 (55)	4 (80)	6 (86)	2 (22)	4 (50)
Ethnicity reported	5 (17)	1 (20)	0	0	4 (50)
ClinicalTrials.gov					
Race reported	13 (45)	0	1 (14)	4 (44)	8 (100)
Ethnicity reported	11 (38)	0	0	3 (33)	8 (100)

* Values are expressed as n (%) unless otherwise indicated. FDA, US Food and Drug Administration.
[a] Year ranges are based on the start dates of the trials.
[b] All trial start dates in this group were 2017 and later, aligning with the implementation of the FDA "Final Rule."
[c] Any race or ethnicity data reported in any of the three sources (journal publication, FDA.gov, ClinicalTrials.gov).
[d] One pivotal trial (M02-570) from the 2000–2005 period was not published, reducing the total number of trials in the publication subgroup from 29 to 28. For the 2000–2005 period in this subgroup, the denominator is adjusted from 5 to 4.

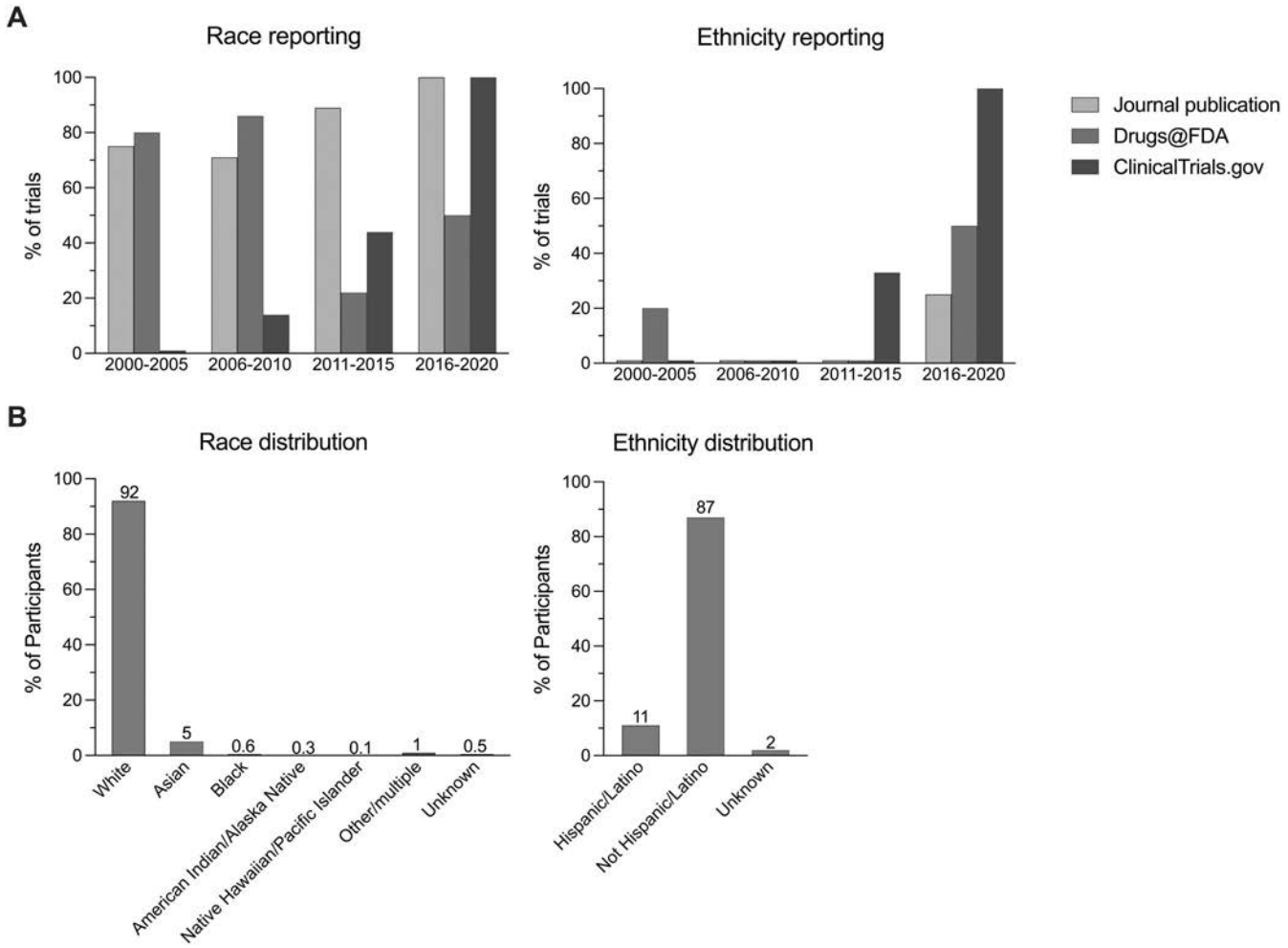


Figure 1. Race and ethnicity reporting and distribution in pivotal PsA trials. (A) Race and ethnicity reporting in 29 pivotal PsA trials across three data sources: journal publications (n = 28, one trial unpublished), Drugs@FDA database (specifically the US Food and Drug Administration Summary Basis of Approvals), and ClinicalTrials.gov. (B) Race and ethnicity distribution in pivotal PsA trials. Numbers above bars indicate the percentage of participants in each category. The total number of participants was n = 14,165 for race and n = 8,105 for ethnicity. PsA, psoriatic arthritis.

the 16 trials with available FDA SBoAs included ethnicity data, accounting for just 17% (5 of 29) of all pivotal trials. Race data were available on ClinicalTrials.gov for 45% (13 of 29) of trials, and ethnicity data were available for 38% (11 of 29). Race and ethnicity reporting improved over time, both reaching 100% (8 of 8 trials) on ClinicalTrials.gov in 2016 to 2020 (Figure 1A).

Race and ethnicity reporting by journal. A primary publication could be identified for 28 of the 29 pivotal trials, published in 1 of 5 journals (Table 2): *Annals of the Rheumatic Diseases* (*Ann Rheum Dis*), 11 (39%); *The Lancet*, 7 (25%); *Arthritis & Rheumatology* (*Arthritis Rheumatol*), 5 (18%); *New England Journal of Medicine* (*NEJM*), 4 (14%); and *Journal of Rheumatology* (*J Rheumatol*), 1 (4%). Race reporting varied across these journals: 82% (9 of 11 trials) in *Ann Rheum Dis*, 86% (6 of 7) in *The Lancet*, 80% (4 of 5) in *Arthritis Rheumatol*, and 100% in *NEJM* (4 of 4) and *J Rheumatol* (1 of 1). Race reporting was frequently limited to the fraction of White participants. Ethnicity data were rarely included in the primary publication,

Table 2. Race and ethnicity reporting by journal*

	Ann Rheum Dis	*The Lancet*	*Arthritis Rheumatol*	*NEJM*	*J Rheumatol*
Journal publications, n	11	7	5	4	1
Race reported	9 (82)	6 (86)	4 (80)	4 (100)	1 (100)
Only White race reported	2/9 (22)	2/6 (33)	4/4 (100)	3/4 (75)	0
Ethnicity reported	2 (18)	0	0	0	0

* Values are expressed as n (%) unless otherwise indicated.

Table 3. Race and ethnicity representation

	All	2000–2005[a]	2006–2010	2011–2015	2016–2020
Number of trials	29	5	7	9	8
Race of participants[b] (n = 27 trials)					
Number of trials	27	4	7	8	8
Total number of participants, n	14,165	1,023	2,999	4,019	6,124
White, n (%)	13,076 (92)	974 (95)	2,863 (95)	3,556 (88)	5,683 (93)
Asian, n (%)	680 (5)	12 (1)	68 (2)	279 (7)	326 (5)
Black, n (%)	79 (0.6)	10 (1)	15 (0.5)	15 (0.4)	39 (0.6)
American Indian/Alaska Native, n (%)	35 (0.3)	1 (0.1)	2 (0.1)	21 (0.5)	11 (0.2)
Native Hawaiian/Pacific Islander, n (%)	16 (0.1)	0	1 (0.03)	3 (0.1)	12 (0.2)
Other/multiple, n (%)	205 (1)	15 (1)	49 (2)	93 (2)	48 (0.8)
Unknown or not reported, n (%)	69 (0.5)	11 (1)	1 (0.03)	52 (1)	5 (0.1)
Ethnicity of participants (n = 12 trials)					
Number of trials	12	1	0	3	8
Total number of participants, n	8,105	205	0	1,776	6,124
Hispanic or Latino, n (%)	876 (11)	11 (5)	0	192 (11)	673 (11)
Not Hispanic or Latino, n (%)	7,084 (87)	194 (95)	0	1,445 (81)	5,445 (89)
Unknown, n (%)	145 (2)	0	0	139 (8)	6 (0.01)

[a] Year ranges are based on the start dates of the trials.
[b] White race was reported in 27 trials, Asian was reported in 24 trials, Black was reported in 24 trials, American Indian/Alaska Native was reported in 16 trials, and Native Hawaiian/Pacific Islander was reported in 11 trials.

with *Ann Rheum Dis* being the only journal to report ethnicity data in 18% (2 of 11) of trials.

Race and ethnicity representation in pivotal trials. Race data were available for 27 of the 29 trials, encompassing 14,165 participants (Table 3, Figure 1B). White participants accounted for 92% of the total population, with consistently high representation across time periods: 95% in 2000 to 2005 and 2006 to 2010, 88% in 2011 to 2015, and 93% in 2016 to 2020. Asian participants made up 5% of participants overall, with a slight increase over time: 1% in 2000 to 2005, 2% in 2006 to 2010, 7% in 2011 to 2015, and 5% in 2016 to 2020. Black participants made up 0.6% of the participants, with consistently low representation across all periods: 1% in 2000 to 2005 and between 0.4% and 0.6% in subsequent periods. American Indian/Alaska Native participants made up 0.3% of all trial participants, and Native Hawaiian/Pacific Islander participants accounted for 0.1%. Ethnicity data were available for 12 of the

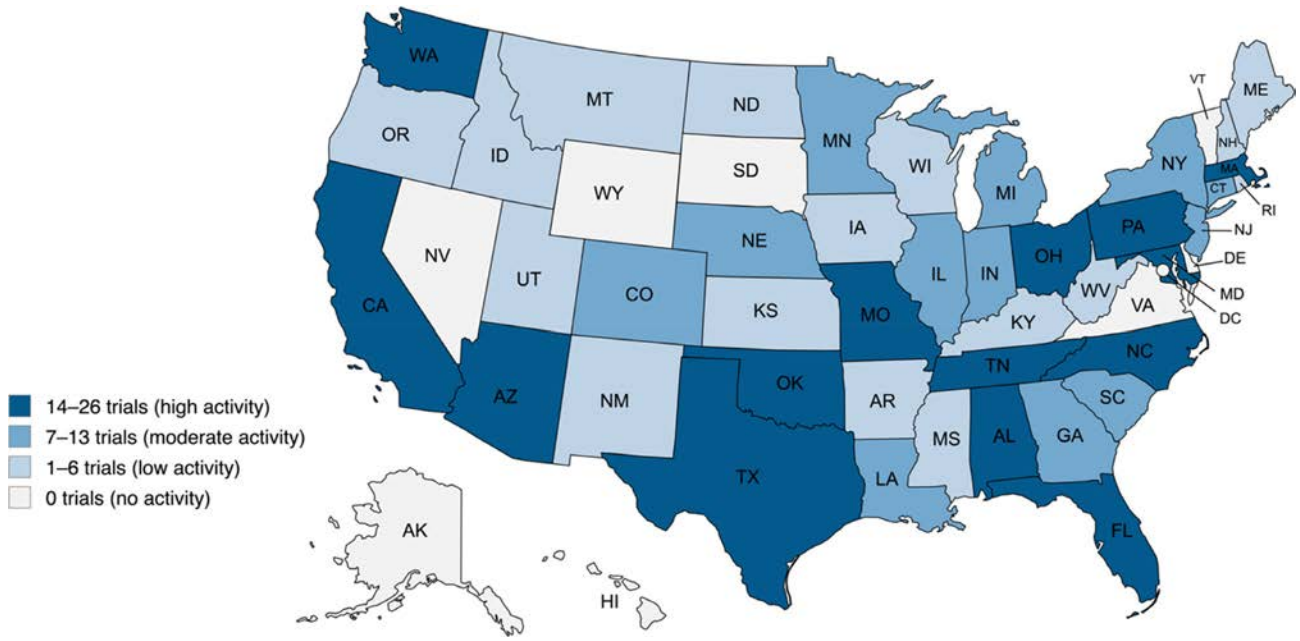


Figure 2. Geographic distribution of pivotal PsA trials in the United States. This map illustrates the geographic distribution of enrollment sites across US states for 26 pivotal PsA trials (location data were missing for 3 of 29 trials). The number of trials in which a state had at least one enrollment site determines its activity level. Darker shades indicate higher clinical trial activity. PsA, psoriatic arthritis.

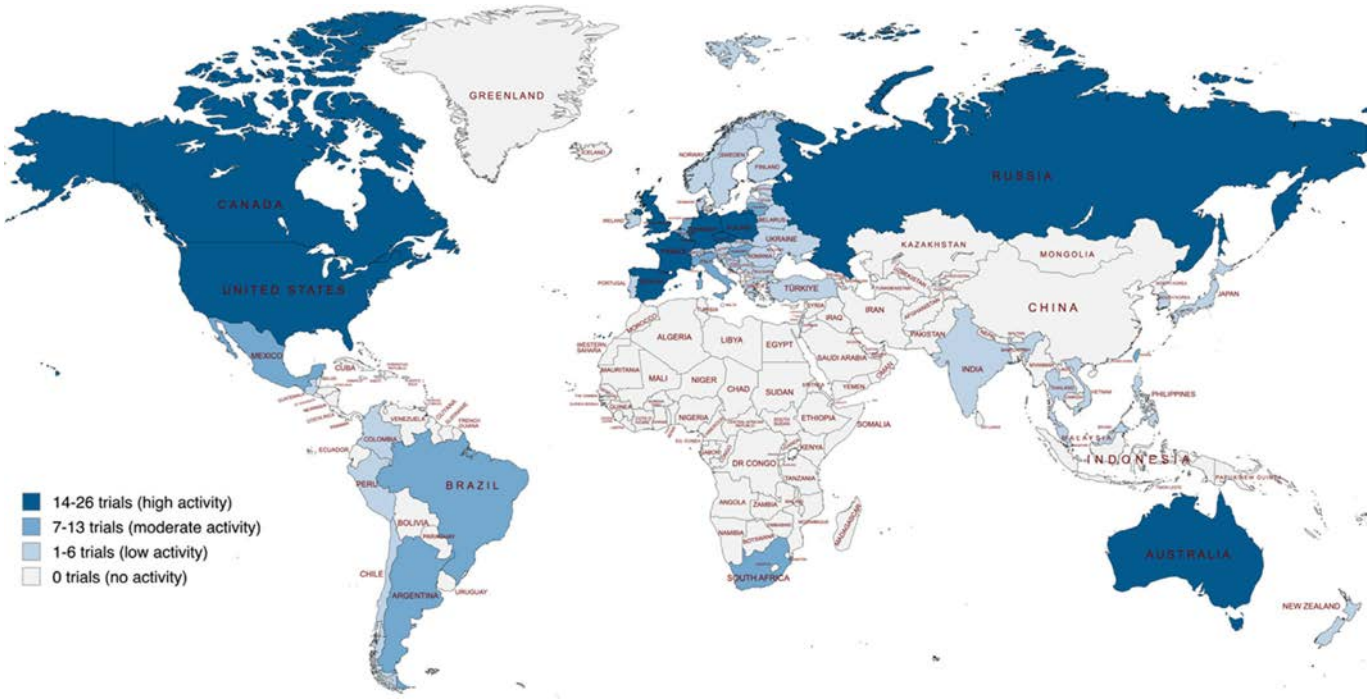


Figure 3. Geographic distribution of pivotal PsA trials in the world. This map illustrates the global distribution of enrollment sites for 26 pivotal PsA trials (3 trials had missing location data). The number of trials in which a country had at least one enrollment site determines its activity level. Darker shades represent higher clinical trial activity. PsA, psoriatic arthritis.

29 trials, including 8,105 participants. Hispanic or Latino individuals represented 11% of participants in these studies, with consistent representation of 11% in both the 2011 to 2015 and 2016 to 2020 periods; a single trial reporting ethnicity data before 2011 included 5% Hispanic or Latino participants.

Geographic representation of pivotal trials. Detailed location data were available for 26 pivotal PsA trials, all of which enrolled patients in the United States. Clinical trial activity was highest in Texas (24 trials), followed by Florida and Pennsylvania (22 each), California (20), and Alabama (18) (Figure 2, Suppl. Table 2). Overall, 14 states had high trial activity (14–26 trials), 12 states had moderate activity (7–13 trials), 16 states had low activity (1–6 trials), and 8 states had no participation in clinical trials. We noticed a strong positive correlation between state population size and clinical trial activity ($r = 0.78$, 95% CI 0.63–0.87, $P < 0.0001$) (Suppl. Figure 2).

Globally, enrollment sites were identified in 55 unique countries (Figure 3). The median number of countries per trial increased from 4 in 2000 to 2005 to 15 in 2016 to 2020 (Suppl. Table 3). A total of 10 countries had high trial activity (14–26 trials), 10 had moderate activity (7–13 trials), and 35 had low activity (1–6 trials), with no enrollment sites in all other countries. Clinical trial activity was highest in the United States (26 trials), followed by Poland (22 trials), Canada and Spain (21 trials each), and Germany and the United Kingdom (20 trials each). Enrollment sites were concentrated in North America and Europe, with comparatively fewer sites in Asia, Africa, and Latin America. There has been a substantial increase in trial site representation in Eastern Europe and East and Southeast Asia from 2006–2010 to 2016–2020 (Suppl. Table 2). In Eastern Europe, Czechia (28.6% to 62.5%), Ukraine (0% to 50.0%), and Latvia (14.3% to 37.5%) saw the largest increases, with notable growth also observed in Lithuania, Estonia, Hungary, Bulgaria, and Romania. In Asia, trial site representation expanded considerably in Taiwan (14.3% to 50.0%), Japan (0% to 50.0%), and Malaysia (0% to 50.0%), but also in the Republic of Korea, Singapore, and Thailand.

DISCUSSION

In this study, we investigated race, ethnicity, and geographic diversity in pivotal PsA drug trials over the past 20 years, focusing on reporting practices and participant representation. Reporting of race and ethnicity varied across journal publications, the Drugs@FDA database, and ClinicalTrials.gov. Journal publications frequently reported race data, though this was often limited to the proportion of White participants and rarely included ethnicity. ClinicalTrials.gov demonstrated substantial improvements in race and ethnicity reporting in recent years, whereas the Drugs@FDA database provided critical yet inconsistent data due to the variable availability of FDA SBoAs. PsA trial populations were predominantly White participants, with underrepresentation

of other racial and ethnic groups and no clear trend of improvement over the last 20 years other than in the recruitment of Asian individuals in the last 10 years. Our analysis of trial site locations revealed a high representation of enrollment sites in Western Europe, North America, and, increasingly with time, Eastern Europe and Asia, with the latter regions potentially contributing to the increased Asian representation in clinical trial populations. In the United States, trial site activity correlated with state population size, with notable activity in racially and ethnically diverse regions such as Texas, Florida, California, and Alabama (Suppl. Table 3).

Race and ethnicity reporting in journal publications, as demonstrated by our findings and supported by other studies, has remained inconsistent and inadequate despite some improvements over time.[15,16] Although limited guidance exists in the literature and in author guidelines, there is no standard for the reporting of race and ethnicity in journal publications.[17,18] For instance, the International Committee of Medical Journal Editors recommended in 2013 including descriptive data on race and ethnicity and clarifying how these variables were measured but did not endorse a specific categorization system, such as the OMB standards.[18]

FDA SBoAs were inconsistently available on the Drugs@FDA database. When accessible, FDA SBoAs provided detailed demographic breakdowns, filling gaps left by other data sources. Demographic data in FDA SBoAs may also be more reliable than other sources due to the rigorous independent review process conducted by the FDA during drug approval. Although the FDA aims to publish these SBoAs within 30 days of the initial approval of a new drug,[19] SBoAs were accessible online for only half of the pivotal trials analyzed, likely in part because PsA was often not the initial indication approved for a specific therapeutic. This resulted in missing demographic data from Drugs@FDA, a limitation we mitigated by integrating data from ClinicalTrials.gov and journal publications. Ensuring the timely and consistent publication of FDA SBoAs for all indications would greatly improve the accessibility and completeness of available demographic data in pivotal trials for PsA and other conditions.

ClinicalTrials.gov has shown substantial progress in reporting race and ethnicity data for pivotal PsA trials over the years, reaching 100% reporting in the 2016 to 2020 period. These improvements in race and ethnicity reporting align with trends observed in other studies.[15,20–22] This progress is likely attributable to key regulatory milestones, including the 2007 FDA Amendments Act, the 2016 FDA guidance on the collection of race and ethnicity data, and the 2017 implementation of the FDA "Final Rule,"[23–25] which established ClinicalTrials.gov as a critical resource for race and ethnicity data. Nonetheless, data gaps persist in older trials and those trials that did not fully adhere to these regulations, underscoring the need for ongoing compliance efforts and monitoring.

Our findings are consistent with prior studies demonstrating that White participants consistently made up the overwhelming majority of PsA trial participants, whereas American Indian/Alaska Native, Asian, Black, Native Hawaiian/Pacific Islander, and Hispanic/Latino groups were underrepresented.[26–28] This underrepresentation of racial and ethnic groups has also been observed in RCTs in other rheumatic diseases, such as axial spondyloarthritis, systemic sclerosis, rheumatoid arthritis, gout, and systemic lupus erythematosus,[22,29–32] as well as in other medical fields, such as dermatology, neuroscience, pediatrics, ophthalmology, and oncology.[16,20,33–35]

Expanding recruitment efforts to previously underrepresented regions worldwide can enhance both geographic and demographic diversity of study populations.[36] Pivotal PsA trials have increasingly included sites in Eastern Europe and Asia, marking progress toward global representation. Despite these shifts in geographic site distribution, White participants have continued to compose the overwhelming majority of PsA trial participants (93% in 2016–2020), highlighting the need for additional efforts to ensure greater racial and ethnic diversity in trial enrollment.

Several factors may explain why trial sponsors favor certain recruitment sites over others.[37–40] Preferred sites often offer a combination of experienced investigators, established research infrastructure, and streamlined regulatory processes that facilitate faster trial start-up timelines. Cost-efficiency also plays a role, particularly in Eastern Europe and parts of Asia, where lower operational costs and faster recruitment may offer advantages over Western Europe and North America. Site selection may also reflect anticipated participant availability—particularly access to treatment-naive patients—and whether sponsors intend to seek approval and market the drug in that country.

In the United States, participants were enrolled into the pivotal PsA trials across the entire country. Only eight states did not contribute a single recruitment site. We observed a strong positive correlation between state population size and trial activity, that is, states with large populations also had the highest representation in the pivotal PsA trials (Suppl. Figure 2). Some of these states, such as California, Texas, and Florida, are also among the most racially and ethnically diverse states in the country (Suppl. Table 3). Nevertheless, this diversity is not reflected in the representation of non-White groups in the overall study populations. Because of a lack of more granular recruitment data, we cannot tell whether this is due to selective recruitment or the limited contribution of study sites in these states to the overall study populations. However, these findings suggest that the presence of trial sites in racially and ethnically diverse states alone may not suffice to address enrollment disparities.

Barriers to the participation of underrepresented racial and ethnic populations in clinical trials in the United States can be divided into three main groups.[41–43] Patient-level barriers include transportation challenges, limited access to trials, child care obligations, lack of paid time off, language barriers, cultural differences, and mistrust of the health care system. Provider-level

barriers include limited awareness of trials, implicit biases toward patients of certain racial or ethnic backgrounds, and concerns about potentially damaging the patient–provider relationship. System-level barriers, which disproportionately affect minority populations, include the limited availability of clinical trials, inadequate research infrastructure, insufficient community engagement, financial burdens on institutions, and restrictive eligibility criteria.[42] These barriers often manifest differently across racial and ethnic groups; for instance, Black individuals may distrust medical research because of historical abuses, whereas Native American, Native Hawaiian, and Pacific Islander populations may be underenrolled partly because trial sites are often not located in areas where these populations are concentrated.[41]

Addressing these barriers to participation is essential for improving the enrollment of underrepresented groups and requires a tailored, multifaceted approach.[42,44] Strategies include cultural sensitivity training, hiring culturally diverse and bilingual staff, translating materials, and engaging trusted community leaders.[41,44] Additional measures include flexible scheduling, accessible study locations, and monetary incentives.[44] The VISIBLE trial on psoriasis demonstrated the effectiveness of implementing a broad range of such measures in recruiting a diverse participant population.[45] Recent FDA guidance documents also provide recommendations and emphasize structured plans to increase the participation of underrepresented groups in clinical research.[3,11]

To implement these strategies effectively, PsA trial sponsors and regulatory agencies could consider specific interventions. Diversity action plans, as previously recommended by the FDA, may provide a structured framework for improving enrollment.[3] Trial sponsors could provide more comprehensive support to trial sites through financial incentives (to the extent permitted by institutional review boards and ethics committees), technical assistance, and oversight to ensure effective implementation of diversity-focused recruitment efforts.[46] Expanding decentralized and hybrid clinical trial models, including home-based monitoring and telemedicine visits, may reduce logistical barriers that disproportionately affect racial and ethnic minority populations.[47] Strengthening community partnerships—with local leaders, primary care networks, and patient advocacy groups—can help build trust and engagement in research participation.[44] Collectively, these and other targeted strategies could help improve equitable representation in future PsA trials.

Our study has several limitations. Race and ethnicity are social constructs, and their US-centric OMB categorizations may be less meaningful in an international context. The lack of standardized definitions and data collection procedures represents a challenge in race and ethnicity reporting and may result in inconsistent or inaccurate data across sources and trials.[48] Our data sources often did not specify how race and ethnicity data were collected, and there were sometimes inconsistencies in how these categories were defined. The data sources lack granular enrollment data for individual trial sites, precluding a more detailed analysis of enrollment trends and factors affecting trial population representation in the United States or other countries. Our analysis focuses exclusively on pivotal trials of those therapeutics that were FDA approved. We focused on these trials reviewed by the FDA because international and national regulatory guidelines specify that participants in clinical trials should be representative of the population for whom the medicine will be indicated in clinical practice, as defined in the final label.[49,50] We believe these trials are the most relevant and rigorously reviewed because they were critical for FDA approval. However, the total number of interventional RCTs performed in PsA is substantially higher and includes phase 2, additional phase 3, and phase 4 studies of FDA-approved drugs as well as RCTs originally intended as pivotal for therapeutics that failed to achieve FDA approval.

Our study has several strengths. We conducted a comprehensive analysis of race, ethnicity, and geographic diversity in pivotal FDA trials of PsA bDMARDs and tsDMARDs, integrating data from three key sources: journal publications, the Drugs@FDA database, and ClinicalTrials.gov. This novel multisource approach allowed us to identify reporting gaps and inconsistencies that would be overlooked in single-source analyses. By focusing exclusively on pivotal trials, our findings are directly relevant to the most rigorously reviewed and clinically impactful trials in PsA.

In conclusion, this study underscores the persistent underrepresentation of racial and ethnic minority groups in pivotal PsA trials and highlights geographic disparities in participant inclusion. Recent improvements in race and ethnicity reporting through ClinicalTrials.gov are promising, but significant gaps remain in journal publications and FDA data. Standardizing and mandating race and ethnicity reporting across data sources is crucial to ensure transparency and equity. Tailored and multifaceted recruitment strategies are essential to increase diversity and ensure the generalizability of trial results, ultimately aiding clinicians to provide the most effective and informed treatment options for all patients.

ACKNOWLEDGMENT

We used ChatGPT (OpenAI) to assist with spelling, grammar, and language improvement during the preparation of this manuscript.

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 Goel 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. Bothwell LE, Podolsky SH. The emergence of the randomized, controlled trial. N Engl J Med 2016;375(6):501–504.
2. Bothwell LE, Greene JA, Podolsky SH, et al. Assessing the gold standard—lessons from the history of RCTs. N Engl J Med 2016; 374(22):2175–2181.
3. US Food and Drug Administration. Diversity action plans to improve enrollment of participants from underrepresented populations in clinical studies. Accessed January 29, 2025. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/diversity-action-plans-improve-enrollment-participants-underrepresented-populations-clinical-studies
4. Schwartz AL, Alsan M, Morris AA, et al. Why diverse clinical trial participation matters. N Engl J Med 2023;388(14):1252–1254.
5. Lembke S, Macfarlane GJ, Jones GT. The worldwide prevalence of psoriatic arthritis-a systematic review and meta-analysis. Rheumatology (Oxford) 2024;63(12):3211–3220.
6. Ogdie A, Matthias W, Thielen RJ, et al. Racial differences in prevalence and treatment for psoriatic arthritis and ankylosing spondylitis by insurance coverage in the USA. Rheumatol Ther 2021;8(4): 1725–1739.
7. Takeshita J, Chau J, Duffin KC, et al. Promoting diversity, equity, and inclusion for psoriatic diseases. J Rheumatol 2022;49(6 suppl 1): 48–51.
8. Ahmed F, Ogdie A, Fitzsimmons R, et al. AB0922 psoriatic arthritis disease activity differs by race/ethnicity [abstract]. Ann Rheum Dis 2022;81(suppl 1):1590.
9. Haberman RH, Ahmed T, Um S, et al. Racial and ethnic determinants of psoriatic arthritis phenotypes and disease activity. Rheumatology (Oxford) 2025;64(2):574–580.
10. Lin I, Krupsky K, Way N, et al. Patient-reported and economic racial and ethnic disparities in patients with psoriatic arthritis: results from the National Health and Wellness Survey. Rheumatol Ther 2024; 11(6):1569–1590.
11. US Food and Drug Administration. Enhancing the diversity of clinical trial populations—eligibility criteria, enrollment practices, and trial designs; guidance for industry; availability. Accessed January 29, 2025. https://www.federalregister.gov/documents/2020/11/10/2020-24881/enhancing-the-diversity-of-clinical-trial-populations-eligibility-criteria-enrollment-practices-and
12. Food and Drug Omnibus Reform Act (FDORA) of 2022. Accessed August 2, 2024. https://www.fda.gov/regulatory-information/selected-amendments-fdc-act/food-and-drug-omnibus-reform-act-fdora-2022
13. George E, Baker McDowell A, Vozza M, et al. Regulatory landscape with U.S. patient requirements and clinical trial diversity expectations. Contemp Clin Trials Commun 2024;42:101331
14. Office of Management and Budget. Revisions to OMB's Statistical Policy Directive No. 15: standards for maintaining, collecting, and presenting federal data on race and ethnicity. Accessed August 2, 2024. https://www.federalregister.gov/documents/2024/03/29/2024-06469/revisions-to-ombs-statistical-policy-directive-no-15-standards-for-maintaining-collecting-and
15. Fain KM, Nelson JT, Tse T, et al. Race and ethnicity reporting for clinical trials in ClinicalTrials.gov and publications. Contemp Clin Trials 2021;101:106237.
16. Chen V, Akhtar S, Zheng C, et al. Assessment of changes in diversity in dermatology clinical trials between 2010-2015 and 2015-2020: a systematic review. JAMA Dermatol 2022;158(3):288–292.
17. Flanagin A, Frey T, Christiansen SL; AMA Manual of Style Committee. Updated guidance on the reporting of race and ethnicity in medical and science journals. JAMA 2021;326(7):621–627.
18. International Committee of Medical Journal Editors. Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals. Accessed November 25, 2024. https://www.icmje.org/icmje-recommendations.pdf
19. Center for Biologics Evaluation and Research. SOPP 8401.7: action package for posting. US Food and Drug Administration. 2022. https://www.fda.gov/media/82426/download
20. Rees CA, Stewart AM, Mehta S, et al. Reporting of participant race and ethnicity in published US pediatric clinical trials from 2011 to 2020. JAMA Pediatr 2022;176(5):e220142.
21. Paleoudis EG, Han Z, Gelman S, et al. Improved clinical trial race/ethnicity reporting and updated inclusion profile, 2017-2022: a New Jersey snapshot. Glob Epidemiol 2023;7:100134.
22. Choufani M, Ghusn W, Dubreuil M, et al. Diversity in axial spondyloarthritis drug trials: enrollment by sex, race, ethnicity, and geographic region. J Rheumatol Published online March 15, 2025. https://doi.org/10.3899/jrheum.2024-1013
23. Food and Drug Administration Amendments Act of 2007, HR 3580, 110th Congress (2007). Pub L No. 110–85. Accessed July 20, 2024. https://www.govinfo.gov/content/pkg/PLAW-110publ85/pdf/PLAW-110publ85.pdf#page=82
24. US Food and Drug Administration. Collection of race and ethnicity in clinical trials. Accessed August 2, 2024. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/collection-race-and-ethnicity-data-clinical-trials
25. 42 CFR part 11—clinical trials registration and results information submission. Accessed August 2, 2024. https://www.ecfr.gov/current/title-42/chapter-I/subchapter-A/part-11
26. Shwe S, Nguyen C, Bhutani T. Racial disparities in clinical trials of biologic treatments for psoriatic arthritis. J Am Acad Dermatol 2022; 87(4):910–912.
27. Zhu W, Ayoub S, Morand E, et al. The evolving demographics of participants in psoriatic arthritis phase III randomised controlled trials of b/tsDMARDs: a systematic review. Semin Arthritis Rheum 2023;60: 152175.
28. Kaur D, Zulfiqar F, Batra A, et al. Racial and gender disparities in psoriatic arthritis clinical trials [abstract]. Arthritis Rheumatol 2024;76 (suppl 9). Accessed January 29, 2025. https://acrabstracts.org/abstract/racial-and-gender-disparities-in-psoriatic-arthritis-clinical-trials/
29. Falasinnu T, Chaichian Y, Bass MB, et al. The representation of gender and race/ethnic groups in randomized clinical trials of individuals with systemic lupus erythematosus. Curr Rheumatol Rep 2018; 20(4):20.
30. Strait A, Castillo F, Choden S, et al. Demographic characteristics of participants in rheumatoid arthritis randomized clinical trials: a systematic review. JAMA Netw Open 2019;2(11):e1914745.
31. Fogacci F, Borghi C, Di Micoli A, et al. Inequalities in enrollment of women and racial minorities in trials testing uric acid lowering drugs. Nutr Metab Cardiovasc Dis 2021;31(12):3305–3313.
32. Chaix E, Mongin D, Gabay C, et al. Reporting and representativeness of race, ethnicity, and socioeconomic status in systemic sclerosis randomized trials: an observational study. Arthritis Care Res (Hoboken) 2023;75(8):1698–1705.
33. Oyer RA, Hurley P, Boehmer L, et al. Increasing racial and ethnic diversity in cancer clinical trials: an American Society of Clinical Oncology and Association of Community Cancer Centers joint research statement. J Clin Oncol 2022;40(19):2163–2171.
34. Rutten-Jacobs L, McIver T, Reyes A, et al. Racial and ethnic diversity in global neuroscience clinical trials. Contemp Clin Trials Commun 2024;37:101255.
35. Montazeri F, Wang M, Atkuru A, et al. Racial, ethnic, and gender diversity in United States ophthalmology clinical trials. Ophthalmol Sci 2023;4(1):100402.

36. Tečić Vuger A, Separovic R, Tolaney SM, et al. Globalization of clinical research in oncology: status, challenges, and future directions. J Cancer Policy 2024;42:100500.

37. Thiers FA, Sinskey AJ, Berndt ER. Trends in the globalization of clinical trials. Nat Rev Drug Discov 2008;7:13–14.

38. Berthon-Jones N, Courtney-Vega K, Donaldson A, et al. Assessing site performance in the Altair study, a multinational clinical trial. Trials 2015;16:138.

39. Dombernowsky T, Haedersdal M, Lassen U, et al. Clinical trial allocation in multinational pharmaceutical companies - a qualitative study on influential factors. Pharmacol Res Perspect 2017;5(3):e00317.

40. Jeong S, Sohn M, Kim JH, et al. Current globalization of drug interventional clinical trials: characteristics and associated factors, 2011-2013. Trials 2017;18(1):288.

41. George S, Duran N, Norris K. A systematic review of barriers and facilitators to minority research participation among African Americans, Latinos, Asian Americans, and Pacific Islanders. Am J Public Health 2014;104(2):e16–e31.

42. Hamel LM, Penner LA, Albrecht TL, et al. Barriers to clinical trial enrollment in racial and ethnic minority patients with cancer. Cancer Control 2016;23(4):327–337.

43. Sheikh SZ, Wanty NI, Stephens J, et al. The state of lupus clinical trials: minority participation needed. J Clin Med 2019;8(8):8.

44. Wali N, Huda MN, Gill T, et al. A systematic review of recruitment and retention of ethnic minorities and migrants in obesity prevention randomised controlled trials. Int J Obes (Lond) 2024;48(8):1065–1079.

45. Alexis A, McMichael A, Vashi N, et al. Improving diversity in a novel psoriasis study: VISIBLE as a framework for clinical trial quality improvement. JAMA Dermatol 2025;161(3):256–264.

46. Yousafi S, Rangachari P, Holland ML. Barriers to recruitment and retention among underrepresented populations in cancer clinical trials: a qualitative study of the perspectives of clinical trial research coordinating staff at a cancer center. J Healthc Leadersh 2024;16:427–441.

47. Jean-Louis G, Seixas AA. The value of decentralized clinical trials: Inclusion, accessibility, and innovation. Science 2024;385(6711):eadq4994.

48. Kaneshiro B, Geling O, Gellert K, et al. The challenges of collecting data on race and ethnicity in a diverse, multiethnic state. Hawaii Med J 2011;70(8):168–171.

49. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. General considerations for clinical studies E8(R1). https://database.ich.org/sites/default/files/E8-R1_Guideline_Step4_2021_1006.pdf

50. US Food and Drug Administration. E8(R1) general considerations for clinical studies. Accessed April 10, 2025. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/e8r1-general-considerations-clinical-studies

Arthritis Care & Research
Vol. 77, No. 11, November 2025, pp 1303–1312
DOI 10.1002/acr.25575

American College of Rheumatology
Empowering Rheumatology Professionals

Association of Elevated Serum S100A8/A9 Levels and Cognitive Impairment in Patients With Systemic Lupus Erythematosus

Carolina Muñoz-Grajales,[1] Michelle L. Barraclough,[2] Juan P. Diaz-Martinez,[3] Jiandong Su,[3] Kathleen Bingham,[4] Mahta Kakvan,[3] Roberta Pozzi Kretzmann,[3] Maria Carmela Tartaglia,[5] Lesley Ruttan,[6] May Y. Choi,[7] Simone Appenzeller,[8] Sherief Marzouk,[4] Dennisse Bonilla,[3] Patricia Katz,[9] Dorcas Beaton,[10] Robin Green,[5] Dafna D. Gladman,[3] Joan Wither,[11] and Zahi Touma[3]

Objective. Cognitive impairment (CI) is common in patients with systemic lupus erythematosus (SLE). Despite its prevalence, the immune mechanisms are not well understood. We previously reported elevated serum levels of S100A8/A9 and matrix metalloproteinase 9 (MMP-9) in patients with SLE and CI. This study aims to validate those findings by examining the relationship between serum levels and CI in patients with SLE at baseline and after one year.

Methods. We assessed cognitive function in 112 patients with SLE using the adapted American College of Rheumatology–Neuropsychological Battery, defining CI as impairment in two or more domains. Serum S100A8/A9 and MMP-9 levels were measured by enzyme-linked immunosorbent assay. We compared serum levels between CI and non-CI groups, evaluated cognitive domain performance at baseline and one year, and explored associations between serum changes and cognitive status changes.

Results. At baseline, 48 patients (42.8%) had CI. After one year, the cognitive funtion remained stable in 55%, improved in 31.2%, and worsened in 13% of patients. Serum S100A8/A9 levels were significantly higher in CI patients at baseline (P = 0.0007, r = 0.413) and one year (P = 0.0045, r = 0.359), correlating inversely with multiple CI domains. The worsened group showed a significant increase in S100A8/A9 levels, whereas the improved group exhibited a reduction.

Conclusion. In this large cohort of patients with well-characterized SLE, serum S100A8/A9 levels were elevated in those with CI and showed an inverse relationship with cognitive performance across multiple domains. Changes in S100A8/A9 levels corresponded with changes in cognitive status over one year. These findings warrant further investigation into the role of S100A8/A9 in CI within the context of SLE.

INTRODUCTION

Cognitive impairment (CI) is one of the most common neuropsychiatric systemic lupus erythematosus (NPSLE) syndromes (estimated prevalence of 38%, 95% confidence interval 33–43%).[1,2] Symptoms of CI include declines in memory, processing speed, attention, and planning abilities, which significantly impact patients' daily functioning,

Supported by grants from the Arthritis Society of Canada (Y10-15-323 and Y5-15-081), Canadian Institutes of Health Research (PJ5 169668), Physician's Services Incorporated Foundation (16-10), the Ontario Ministry of Research & Innovation - Early Research Awards, and the Lupus Research Alliance (702951). Dr Muñoz-Grajales's work was supported by the Gary S. Gilkeson Career Development Award from the Lupus Foundation of America. Drs Wither and Touma's work was supported by the Department of Medicine, University of Toronto. The Toronto Lupus Program was supported by Lupus Ontario, the Schroeder Arthritis Institute, and donations from the Kathi and Peter Kaiser family, the Lou and Marissa Rocca family, and the Stacey and Mark Krembil family.

[1]Carolina Muñoz-Grajales, MD, PhD, Schroeder Arthritis Institute and Krembil Research Institute, University Health Network and University of Toronto Lupus Clinic, Toronto Western Hospital, Toronto, Ontario, Canada, and University of Manitoba, Winnipeg, Manitoba, Canada; [2]Michelle L. Barraclough, MD, PhD: Schroeder Arthritis Institute and Krembil Research Institute, University Health Network, Toronto, Ontario, Canada, and The University of Manchester, NIHR Manchester Biomedical Research Centre, Manchester University NHS Foundation Trust, and Manchester Academic Health Science Centre, Manchester, United Kingdom; [3]Juan P. Diaz-Martinez, PhD(c), MSc, Jiandong Su, BSc, Mahta Kakvan, MA, MEd, PhD(c), Roberta Pozzi Kretzmann, MSc, Dennisse Bonilla, BSc (Hons), Dafna D. Gladman, MD, Zahi Touma, MD, PhD: Schroeder Arthritis Institute and Krembil Research Institute, University Health Network and University of Toronto Lupus Clinic, Toronto Western Hospital, Toronto, Ontario, Canada; [4]Kathleen Bingham, MD, PhD, Sherief Marzouk, MD: Centre for Mental Health, University Health Network and University of Toronto, Toronto, Ontario, Canada; [5]Maria Carmela Tartaglia, MD, Robin Green, PhD, C. Psych: University of Toronto Krembil Neurosciences Centre, Toronto, Ontario, Canada; [6]Lesley Ruttan, MA, PhD, C. Psych: Toronto Rehabilitation Institute, University Health Network, Toronto, Ontario, Canada; [7]May Y. Choi, MD, MPH: Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada; [8]Simone Appenzeller, MD, PhD: University of Campinas, São Paulo, Brazil; [9]Patricia Katz, MA, PhD: University of California, San Francisco; [10]Dorcas Beaton, MSc, PhD: Institute for Work and Health, University of Toronto, Toronto, Ontario, Canada; [11]Joan Wither, MD, PhD: Schroeder Arthritis Institute and Krembil Research Institute, University Health Network and University of Toronto, Toronto, Ontario, Canada.

SIGNIFICANCE & INNOVATIONS

- Previous studies have highlighted alterations in cytokines and proteases in patients with systemic lupus erythematosus (SLE) with various neuropsychiatric SLE syndromes, implicating neuroinflammation. However, research specifically focused on cognitive impairment (CI) remains limited, often constrained by small sample sizes.
- Our findings emphasize the potential role of S100A8/A9 in the immunopathogenesis of CI among adult patients with SLE. Nevertheless, further investigation is crucial to fully elucidate the underlying mechanisms and validate the significance of S100A8/A9 in CI within patients with SLE.
- This study enhances our understanding of CI in SLE, offering insights that could inform future research directions and therapeutic strategies aimed at addressing cognitive dysfunction in patients with SLE.

social role participation, and health-related quality of life.[3–7] Despite recognition of its negative impact in patients with SLE, the underlying mechanisms of CI remain poorly understood.[8]

Studies have reported that patients with SLE with different NPSLE syndromes exhibit alterations in the levels of various molecules, including proinflammatory mediators (interleukin-6 [IL-6], interferon-γ [IFNγ], tumor necrosis factor α [TNFα], etc) and proteases (neutrophil gelatinase–associated lipocalin [NGAL], matrix metalloproteinase 9 [MMP-9]), which can affect intrinsic brain components such as the blood–brain barrier (BBB), the neurovascular interface, and resident microglia, leading to neuroinflammation.[9–15] However, studies specifically focusing on CI are scarce, and those available had a limited sample size, leaving the role of these molecules in CI in question.

Recently, we published results from a large, well-characterized cohort of patients with SLE (n = 290, 116 patients [40%] with CI), in which we investigated the potential role of serum analytes (not including autoantibodies) in distinguishing patients with SLE with and without CI.[16] Among the nine analytes measured (IL-6, IL-10, IFNγ, NGAL, MMP-9, S100A8/A9, S100B, TNFα, and TWEAK), we found that patients with SLE and CI had significantly higher serum levels of S100A8/A9, and to a lesser extent,[16] MMP-9. Mechanistically, this observation aligns with known processes. For instance, S100A8/A9 or calprotectin, is involved in recruiting immune cells to sites of inflammation and activating the NF-κB pathway, which leads to the production of proinflammatory cytokines[17] such as TNFα and IL-6. Specifically, in neurologic disorders, it has been shown that S100A8/A9 induces the production of reactive oxygen species and activates microglia.[18] Similarly, it is known that MMP-9 contributes to the breakdown of the extracellular matrix and therefore may participate in the remodeling of the BBB, thus facilitating the infiltration of peripheral immune cells and proinflammatory mediators into the central nervous system (CNS) with the subsequent activation of resident cells such as microglia.[19] Additionally, S100A8/A9 may promote cell expression of metalloproteinases,[20] including MMP-9.

In this new study, we aimed to validate our previous findings by investigating the relationship between serum S100A8/A9 and MMP-9 levels and CI at two different time points: baseline and one-year follow-up. We also assessed the relationship with performance across the different CI domain tests and determined whether changes in serum levels from baseline to one-year follow-up correlated with changes in cognition over the same time period.

MATERIALS AND METHODS

Participants and data collection. Included in this study were all individuals from the ongoing larger study at the Toronto Lupus Clinic, Toronto Western Hospital/University Health Network (UHN) who, by February 2022, had completed one year of follow-up. The parent study cohort, as previously described,[21] consisted of adults aged 18 to 65 years meeting the 2019 EULAR/American College of Rheumatology (ACR) SLE classification criteria,[22] demonstrating adequate English proficiency for cognitive task completion, and providing informed consent. Unlike some other NPSLE manifestations, CI can occur without other SLE clinical symptoms and independently of disease activity,[23] making its attribution to SLE particularly difficult. Consequently, in our cohort, CI was linked to SLE if patients had no other illnesses unrelated to SLE that could cause a decline in cognition.

A standardized data retrieval form was used to collect demographic and clinical data. SLE disease activity and damage were assessed using the Systemic Lupus Erythematosus Disease Activity Index-2000 (SLEDAI-2K)[24] and the Systemic Lupus International Collaborating Clinics/ACR Damage Index (SDI),[25] respectively. The study was approved by the Research Ethics Board of the UHN (CAPCR-ID no. 15-9582). Patients and the public were not involved in the design, conduct, reporting, or dissemination plans of our research.

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.25575).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25575.

Address correspondence via email to Zahi Touma, MD, PhD, at zahi.touma@uhn.ca.

Submitted for publication August 13, 2024; accepted in revised form April 8, 2025.

Classification of cognitive status. Cognitive assessments of the study participants were conducted by psychometrists, supervised by neuropsychologists using a modified comprehensive one-hour ACR–Neuropsychological Battery (ACR-NB),[26] as previously outlined.[21] This assessment includes 19 cognitive tests covering six cognitive domains: domain 1 (D1), manual motor speed; D2, simple attention and processing speed; D3, visual-spatial construction; D4, language processing; D5, learning and memory; and D6, executive functioning (see Supplementary Table 1 for tests including in each domain).

Because of missing data from D1 in >10% of participants who were unable to complete the dominant and nondominant hand tapping due to hand pain or joint deformities, the motor speed domain scores were excluded from the analysis. The ACR-NB scores for the remaining five domains were normalized with data stratified by age and sex to classify patients into three groups, as previously described[27]: CI (two or more domains impaired), indeterminate CI (only one domain impaired), and non-CI (no domain impaired). A domain was considered impaired if a Z score of −1.5 or less was recorded in at least one test in D2, D3, and D4 and in two or more tests[27] in D5 and D6.

Serum S100A8/A9 and MMP-9 measurement. Blood samples were collected and processed to obtain serum on the same day as the neuropsychologic assessment. These serum samples were stored at −80°C until use. Evaluation of the effect of duration of storage on serum S100A8/A9 and MMP-9 levels, using Spearman's rank correlation, showed no significant impact.

Serum S100A8/A9 and MMP-9 levels were measured by enzyme-linked immunosorbent assay using DuoSets from R&D Systems. All measurements were performed in duplicate according to the manufacturer's instructions, with dilutions set to 1:2,000 for S100A8/A9 the heterodimer (catalog number DY8226-05; dynamic range 94–6,000 pg/mL) and 1:1,500 for MMP-9 (catalog number DY911-05; dynamic range 31.2–2,000 pg/mL). The intra-assay coefficient of variation (CV) was less than 5% for both, and the interassay CVs were 11.96% and 13.93%, respectively.

Other laboratory measurements. Anti–double-stranded DNA (anti-dsDNA), antiphospholipid antibodies (APLA) (anti-cardiolipin, anti-β2 glycoprotein I, and lupus anticoagulant), and complement levels (C3 and C4) were measured as part of the patients' routine clinical assessments at the UHN laboratory. The results closest to the cognitive assessment, within a 30-day window, were used for analysis.

Statistical analysis. *Comparisons among groups.* For continuous data, the Mann-Whitney U-test was used to compare two groups (CI and non-CI), and the Kruskall-Wallis test with Bonferroni correction was used to compare multiple groups (CI, indeterminate CI, and non-CI). For categorical variables, a chi-square test was used. Correlations were calculated using Spearman's rank correlation coefficient.

Regression analysis. The relationship between serum levels and study participant performance in each test domain, categorized by visit, was investigated using multivariable regression analyses controlling for covariates deemed clinically relevant (sex, age, race, SLEDAI-2K, and the use of antimalarials, glucocorticoids, and immunosuppressants) and adjusting for multiple comparisons. The variables were scaled a priori to better satisfy the assumptions of the regression model.

Variations over time. Patients' cognitive status was categorized as stable, improved, or worsened based on the ACR-NB cognitive assessment from baseline to the one-year visit. Stability was defined as unchanged cognitive status, improvement as a transition from CI or indeterminate CI status to non-CI or from CI to indeterminate CI, and worsening as a transition from non-CI or indeterminate CI status to CI or from non-CI to indeterminate CI. The Wilcoxon matched-paired signed-rank test was used to evaluate the differences in serum levels of S100A8/A9 and MMP-9 between baseline and the one-year visit within groups (stable, improved, and worsened). This was operationalized by first computing the differences in serum levels of S100A8/A9 and MMP-9 between baseline and the one-year visit for each participant within the stable, improved, and worsened cognitive status groups. These differences were then ranked in absolute terms, and signed ranks were assigned based on the direction of change. The test statistic was derived from the sum of these ranks, and significance was determined based on its distribution under the null hypothesis.

The relationship between serum S100A8/A9 and MMP-9 levels and the interaction between group (improved or worsened) and time was investigated using a linear mixed-effects model (LMM). The LMM included fixed effects for group, time, and their interaction and random intercepts for participants to account for repeated measures within participants. A logarithmic transformation was applied to the analyte levels to stabilize variance and meet the assumptions of the model. All statistical analyses were performed using GraphPad Prism version 9.1.5 (GraphPad Software) or RStudio version 1.3.1073 (Integrated Development Environment for R. RStudio, PBC).

RESULTS

At baseline, 48 of the 112 study participants (42.8%) had CI, 32 study participants (28.6%) had indeterminate CI, and 32 study participants (28.6%) did not have CI. At the one-year follow-up, the distribution was 36 study participants (32%) with CI, 34 study participants (30.3%) with indeterminate CI, and 42 study participants (37.5%) without CI. The demographic

and general characteristics at baseline and the one-year follow-up visit are displayed in Tables 1 and 2, respectively. Most study participants were female and of White self-reported race, with no statistically significant difference in the distribution of races among the groups (non-CI, indeterminate CI, and CI) at either baseline or the one-year follow-up visit. Similarly, the median age, disease duration, SLEDAI-2K, and SDI were comparable among groups at both visits. Additionally, the frequency of hypocomplementemia, anti-dsDNA, APLA, and the medications that patients were taking did not differ significantly (chi-square $P > 0.05$). The clinical manifestations were also comparable among groups at both time points (Supplementary Tables 2 and 3). Only two patients had a history of cerebrovascular disease, both of which belonged to the CI group at both visits.

Comparison of S100A8/A9 and MMP-9 serum levels in patients with SLE with and without CI. At both the baseline and one-year follow-up visits, serum levels of S100A8/A9 and MMP-9 showed a moderate correlation with each other ($\rho = 0.52$, $P < 0.0001$ and $\rho = 0.406$, $P < 0.0001$, respectively). However, only the S100A8/A9 levels were significantly higher in patients with CI compared to those with indeterminate CI ($P = 0.0036$ and $P = 0.0045$ at baseline and one-year follow-up, respectively) and non-CI ($P = 0.028$ and $P = 0.0007$). No differences in the serum levels of S100A8/A9 and MMP-9 were found between patients without CI and those with indeterminate CI, at any time point. These findings are presented in Figure 1. Notably, the serum levels of S100A8/A9 did not correlate with the clinical SLEDAI-2K at any visit ($\rho = 0.004$, $P = 0.9$ at baseline and $\rho = 0.15$, $P = 0.11$ at the one-year visit).

Relationship between S100A8/A9 and MMP-9 serum levels and individual tasks in each domain. Among patients with CI, D5 (learning and memory) had the highest proportion of impaired patients at both visits (83.3% at baseline and 80.6% at the one-year follow-up, respectively), followed by D3 (visual-spatial construction, 70.8% and 64%, respectively) and D2 (simple attention and processing speed, 52% and 55.6%, respectively) (Figure 2A). For D4 (language processing), the proportion of patients with CI at baseline was 29% and 25% at the one-year visit. D6 (executive functioning) was slightly different between the baseline and the one-year visit, with rates of impairment increasing from 25% to 39%. Please note that these percentages do not sum to 100% because patients with CI had impairments in two or more domains. As depicted in

Table 1. Baseline study population characteristics*

Variable	Study population (N = 112)	No CI (n = 32, 28.6%)	Indeterminate CI (n = 32, 28.6%)	CI (n = 48, 42.8%)
Female, n (%)	97 (87)	30 (93)	27 (84)	40 (83)
Age at assessment, median (IQR), y	42.9 (30.4–50.7)	39.1 (32.2–50.2)	40.4 (28.3–51.9)	45.1 (31.7–51.67)
Self-reported race, n (%)				
White	69 (61.6)	21 (65.6)	25 (78)	25 (52)
Black	26 (23.2)	8 (25)	2 (6.2)	16 (33)
Chinese	8 (7.4)	2 (6.2)	1 (3)	5 (10.4)
Other[a]	9 (6.25)	1 (3)	4 (12.5)	2 (4.2)
Disease duration at assessment, median (IQR), y	12.5 (5.7–22.4)	12 (5.1–24.4)	3.8 (6–22.5)	12.9 (4.5–21.7)
SDI score, median (IQR)	1 (0–2)	1 (0–1.25)	0 (0–1)	1 (0–2)
Clinical SLEDAI-2K score, median (IQR)	2 (0–4)	2 (0–4)	2 (0–4)	2 (0–5)
Minimum to maximum	0–22	0–14	0–9	0–22
Positive anti-dsDNA, n (%)	45 (40)	15 (46)	13 (40.6)	17 (35.4)
Low C3 or C4, n (%)	41 (36.6)	16 (50)	11 (34.4)	14 (29)
APLA, n (%)	19 (17)	7 (21.8)	4 (12.5)	8 (16.6)
Current medication use				
Antimalarial, n (%)	89 (79)	25 (78)	28 (87.5)	36 (75)
Glucocorticoids, n (%)	139 (47)	13 (40.6)	15 (46.8)	27 (56)
Glucocorticoid dose (prednisone or equivalent), median (SD)	5.23 (9.8)	5.11 (11.7)	3.12 (5.8)	6.75 (10.5)
Immunosuppressant, n (%)	71 (63.4)	18 (56)	23 (71)	30 (62)
Azathioprine	16 (22.5)	7 (38.8)	5 (21)	4 (13.3)
Methotrexate	7 (9.8)	1 (5.5)	4 (17.4)	2 (6.6)
Mycophenolate	45 (63.4)	9 (50)	14 (61)	22 (73)
Other	3 (4.2)	1 (5.5)	0 (0)	2 (6.6)
Biologics	7 (6.2)	2 (6.2)	3 (9.3)	2 (4.2)

* anti-dsDNA, anti–double-stranded DNA; APLA, antiphospholipid antibodies; C3, complement component 3; CI, cognitive impairment; IQR, interquartile range; SDI, Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index; SLEDAI-2K: Systemic Lupus Erythematosus Disease Activity Index-2000.
[a] Other: Native North American, Filipino, and Mixed. Categories for self-reported race were collapsed into Black, White, and Other (including Chinese and other groups), whereas immunosuppressant use was categorized as "yes" or "no" for analysis.

Table 2. Study population characteristics at the one-year follow-up*

Variable	Study population (N = 112)	No CI (n = 42, 37.5%)	Indeterminate CI (n = 34, 30.3%)	CI (n = 36, 32%)
Female, n (%)	97 (87)	36 (85.7)	30 (88)	31 (86)
Age at assessment, median (IQR), y	44 (31.5–52)	40 (32–51)	46.7 (31–53)	45.46 (32–53)
Self-reported race, n (%)				
White	69 (61.6)	27 (64.3)	24 (66.6)	18 (50)
Black	26 (23.2)	6 (14.3)	7 (20.6)	13 (36)
Chinese	8 (7.4)	5 (12)	0 (0)	3 (8.3)
Other[a]	9 (6.25)	4 (9.5)	3 (8.8)	2 (5.5)
Disease duration at assessment, median (IQR), y	13.76 (6.7–23.7)	13.21 (6.76–24)	16.49 (7.35–26)	13.11 (5–23.48)
SDI score, median (IQR)	1 (0–2)	0.5 (0–2)	0 (0–1.25)	1 (0–3)
Clinical SLEDAI-2K score, median (IQR)	2 (0–4)	2 (0–4)	2 (0–4)	2 (0–4)
Minimum to maximum	0–18	0–18	0–10	0–14
Positive anti-dsDNA, n (%)	48 (42.8)	20 (47.6)	13 (38.3)	15 (41.6)
Low C3 or C4, n (%)	42 (37.5)	19 (45)	11 (32.3)	12 (33.3)
APLA, n (%)	22 (19.6)	11 (26)	5 (14.7)	6 (16.6)
Current medication use				
Antimalarial, n (%)	93 (83)	36 (85.7)	29 (85.3)	28 (77)
Glucocorticoids, n (%)	44 (39)	18 (42.8)	12 (50)	14 (38.8)
Glucocorticoid dose (prednisone or equivalent), median (SD)	3.49 (6.1)	4.38 (7.05)	3.41(6.6)	2.51(4.1)
Immunosuppressant, n (%)	66 (59)	25 (59.5)	21 (61.7)	20 (55.5)
Azathioprine	13 (19.7)	6 (24)	3 (14.3)	4 (20)
Methotrexate	6 (9.5)	1 (4)	3 (14.3)	2 (10)
Mycophenolate	45 (68.2)	17 (68)	15 (67)	13 (65)
Other	2 (3)	1 (4)	0 (0)	1 (5)
Biologics	7 (6.2)	2 (6.2)	3 (9.3)	2 (4.2)

* anti-dsDNA, anti–double-stranded DNA; APLA, antiphospholipid antibodies; C3, complement component 3; CI, cognitive impairment; IQR, interquartile range; SDI, Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index; SLEDAI-2K, Systemic Lupus Erythematosus Disease Activity Index-2000.
[a] Other: Native North American, Filipino, and Mixed. Categories for self-reported race were collapsed into Black, White, and Other (including Chinese and other groups), whereas immunosuppressant use was categorized as "yes" or "no" for analysis.

Supplementary Figure 1, at both visits, the serum levels of S100A8/A9 were consistently higher in the group with impairment in the most frequent affected domains (D5 and D3). No differences were observed for MMP-9.

At baseline, of the 48 patients with CI, 25 patients (52%) had two impaired domains, 19 patients (39.6%) had three impaired domains, and 4 patients (8.3%) had four or five impaired domains. Similar proportions were observed at the one-year visit: of the 36 patients with CI, 19 patients (52.7%) had two impaired domains, 12 patients (33.3%) had three impaired domains, and 5 patients (13.8%) had four or five impaired domains. At both time points, the serum levels of S100A8/A9 or MMP-9 did not correlate with the number of impaired domains in the group of patients with CI.

Figure 2B provides a comprehensive summary of the regression analyses evaluating the relationship between serum levels of S100A8/A9 and MMP-9 and task performance across each cognitive domain at baseline and the one-year visit. This analysis showed that S100A8/A9 exhibited a negative relationship with the performance in multiple cognitive tests across different domains at both visits. In contrast, there were no statistically significant findings for MMP-9. Supplementary Figure 2 illustrates the interaction between S100A8/A9 and the cognitive tests that were statistically significant.

Changes over time. A year after the first cognitive assessment, the cognitive status remained unchanged in 62 patients (55%; 27 patients remained as CI, 14 patients had indeterminate CI, and 21 patients did not have CI), improved in 35 patients (31.2%) and worsened in 15 patients (13.4%) (Supplementary Table 4 summarizes the worsened and improved group characteristics). The Wilcoxon matched-pairs signed-rank test revealed no significant changes for S100A8/A9 in the groups that remained unchanged (Supplementary Figure 3) and statistically significant difference within the improved and worsened groups (Figure 3A).

An LMM was used to examine the interaction between group (improved vs worsened cognitive status) and time (baseline vs one-year visit), while accounting for serum levels of S100A8/A9 and MMP-9. This analysis revealed a significant interaction effect for S100A8/A9 (estimate = 1.090, SE = 0.309, P = 0.00092), suggesting that changes in S100A8/A9 levels over time differ significantly between the two groups. Specifically, the improved group

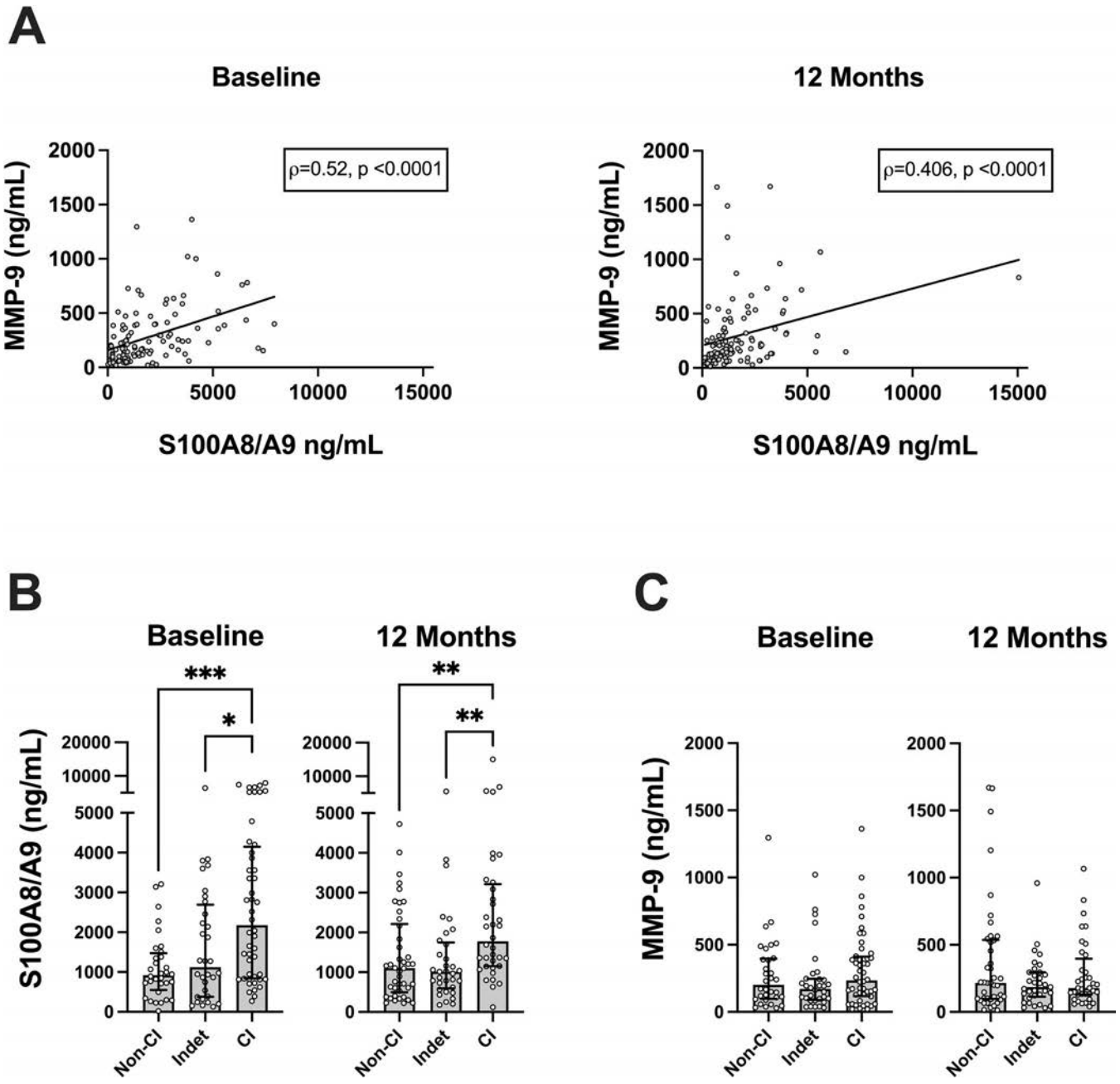


Figure 1. Serum levels of S100A8/A9 and MMP-9 discriminated by visit. (A) Correlations between S100A8/A9 and MMP-9 serum levels at (left) baseline and (right) one-year visit. Correlations were calculated using the Spearman's correlation coefficient. (B and C) Strip plots with median bars showing, from left to right, levels for patients without CI, Indet, and definitive CI for (B) S100A8/A9 and (C) MMP-9. Each circle represents a single study participant, the middle line indicating the median for the study participants and error bars denoting the interquartile ranges. Statistical significance was determined using the Kruskal-Wallis test corrected for multiple comparisons with significant differences indicated by asterisks (*$P \leq 0.05$, **$P \leq 0.01$, ***$P \leq 0.001$). Nonsignificant differences are no displayed. CI, cognitive impairment; Indet, indeterminate cognitive impairment; MMP-9, matrix metalloproteinase 9.

showed a significant median decrease of –562 (95% confidence interval –93 to –903; $P = 0.0214$), whereas the worsened group exhibited a significant median increase of +878 (95% confidence interval 554–2,020; $P = 0.0015$) (Figure 3B) for S100A8/A9. The MMP-9 serum levels did not significantly change between the two time points within either the improved or the worsened group. The main effect of time (baseline vs one-year visit) alone was not significant (estimate = –0.277, SE = 0.173, $P = 0.11553$), indicating that overall changes in S100A8/A9 levels across all participants, irrespective of group differences, were not statistically significant.

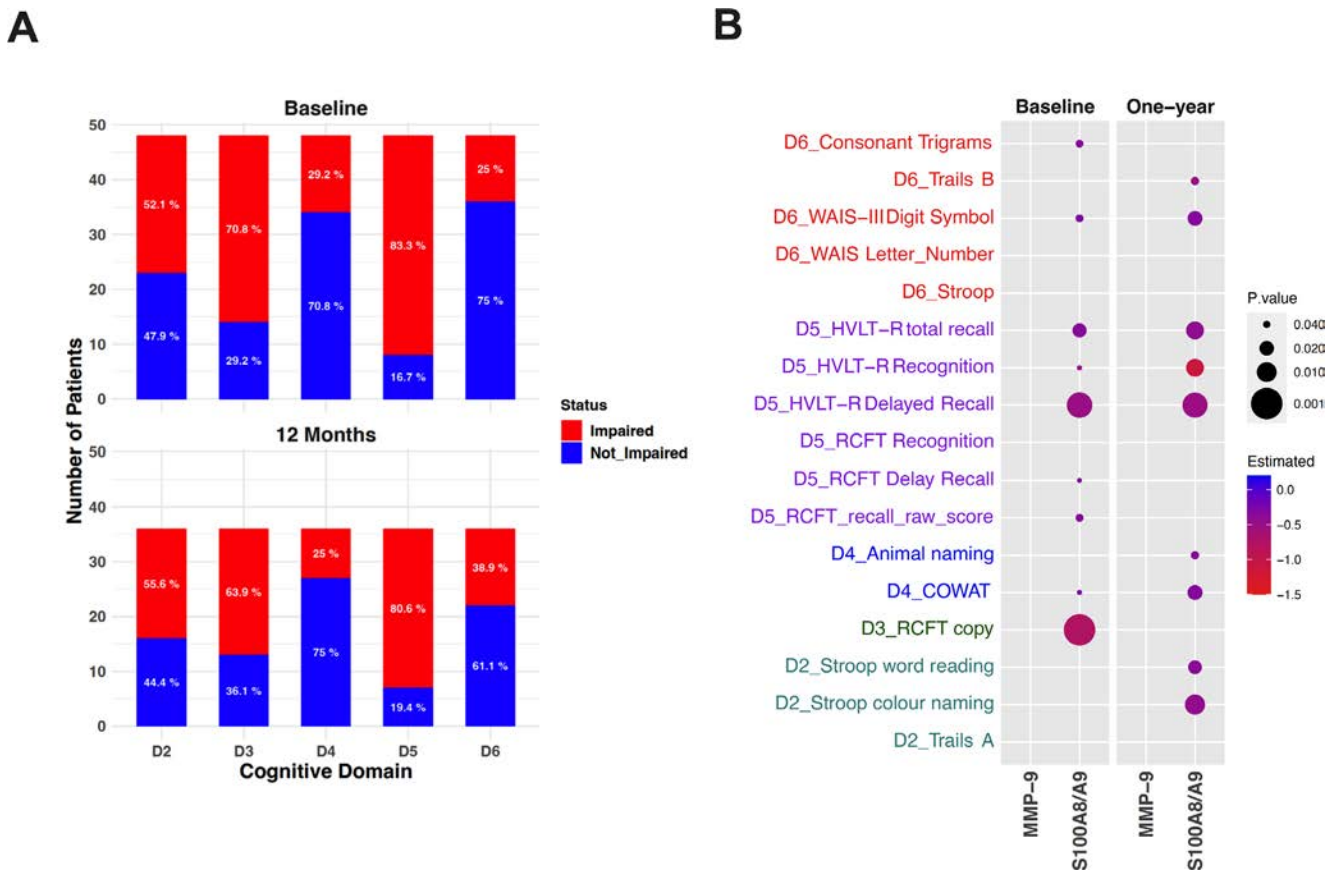


Figure 2. (A) Proportion of patients with impairment in each American College of Rheumatology–Neuropsychological Battery domain discriminated by visit. At baseline, 48 patients had cognitive impairment, and at the one-year visit 36 patients had cognitive impairment. (B) Relationship between the serum MMP-9 and S100A8/A9 levels and each cognitive test Z score. Results from multivariable analysis controlled by sex, age, race, Systemic Lupus Erythematosus Disease Activity Index-2000, and use of antimalarials, glucocorticoids, or immunosuppressants discriminated by visit. Numerical variables were scaled to improve the model fitting. The effect size is indicated by color, and *P* values adjusted for multiple comparisons, indicated by dot size. COWAT, Controlled Oral Word Association Test; D2, domain 2; HVLT-R, Hopkins Verbal Learning Test Revised; MMP-9, matrix metalloproteinase 9; RCFT, Ray-Osterrieth Complex Figure Test; WAIS, Wechsler Adult Intellegence Scale.

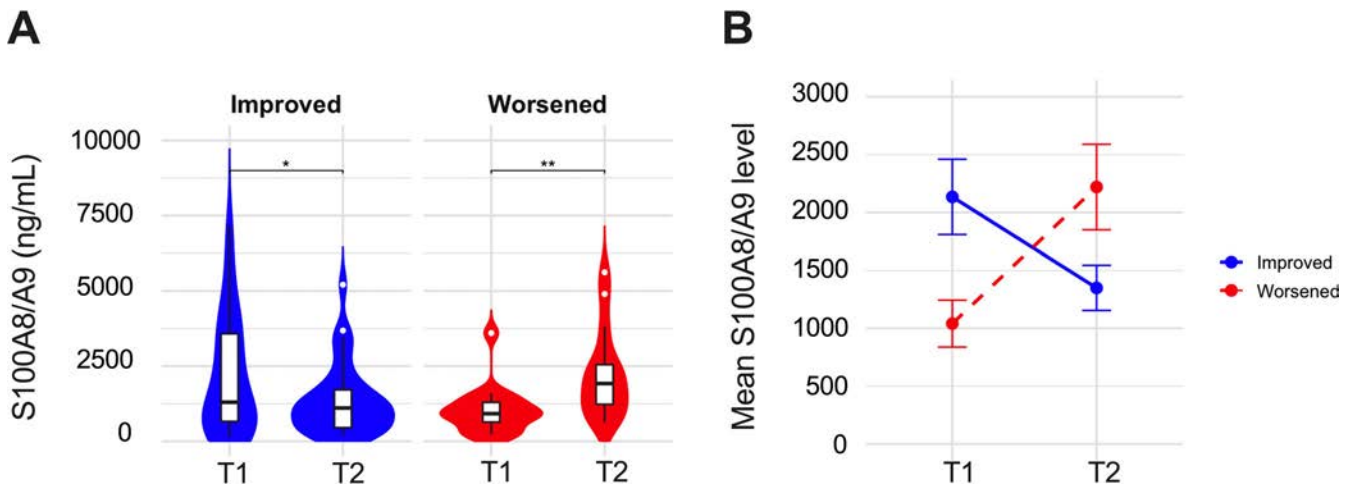


Figure 3. Comparison of S100A8/A9 Levels in improved and worsened groups over time. (A) Violin plots of levels (ng/mL) in the improved and worsened groups at two time points (T1 and T2). The plots show the distribution, median, and interquartile range for each group and time point. The blue color represents the improved group, and the red color represents the worsened group. (B) Interaction plot of mean levels (ng/mL) over time for the improved and worsened groups. Error bars represent the SEM. The solid line represents the improved group, and the dashed line represents the worsened group. T1, baseline; T2, one-year visit. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25575/abstract.

DISCUSSION

Motivated by our previous findings,[16] this study investigated the relationship between serum levels of S100A8/A9 and MMP-9 with CI in patients with well-characterized SLE at baseline and the one-year follow-up. Overall, the findings validate our previous observations regarding S100A8/A9. They also align with a cross-sectional study involving a smaller sample size (n = 72 patients with SLE, 26 patients with CI) that examined the levels of S100A8/A9 in patients with SLE with and without NPSLE, in which higher serum concentrations of S100A8/A9 were seen in patients with SLE with CI compared to those without NPSLE.[28] Although MMP-9 did not show significant differences among groups, as in our initial study with a larger sample size, it positively correlated with S100A8/A9 at both visits. This suggests that MMP-9 may contribute, to some extent, to the mechanisms by which S100A8/A9 influences cognitive decline in patients with SLE.

Cognitive function relies on the intricate interplay between various brain regions,[29] which is partly reflected in the different cognitive domains assessed.[30] Consistent with our previous findings, we observed that elevated levels of S100A8/A9 were negatively correlated with multiple cognitive tests across various domains. Furthermore, the results from the longitudinal part of this study showed that although some participants did not exhibit a significant change in S100A8/A9 levels over time, the direction and extent of change were distinctly different between the worsened and improved groups with a notably greater increase in S100A8/A9 levels in the worsened group.

Studies have shown that patients with SLE have higher circulating levels of S100A8/A9 compared to healthy controls.[31–33] S100A8/A9 is a heterodimer composed of the S100 proteins S100A8 and S100A9, which is the most stable and biologically active form in circulation, although homodimers and heterotetramers also exist.[34] S100A8/A9 constitutes up to 60% of the cytosolic proteins in neutrophils and monocytes. However, other cell types, including endothelial and epithelial cells, can also produce and secrete this cytokine under certain conditions.[34] Notably, neutrophils secrete S100A8/A9 actively in a calcium-dependent manner and during the formation of neutrophil extracellular traps (NETs).[34] Increased NET formation with reduced degradation is a well-recognized phenomenon in SLE.[35,36] As discussed in the Introduction, S100A8/A9 has multiple potential mechanisms by which it can induce neuroinflammation.[17,18,20] Thus, persistent elevated levels due to monocyte and neutrophil activation could ultimately result in overt CI.

Although higher levels of S100A8/A9 have been reported in patients with SLE with active disease, particularly in those with severe manifestations,[34,37] not all studies found a correlation between S100A8/A9 levels and overall disease activity.[32,34] Indeed, and in agreement with our previous findings, we did not observe any correlation between serum S100A8/A9 levels and the SLEDAI-2K or any association with specific clinical characteristics at baseline or during the one-year follow-up. However, it is important to note that in general our study population had low disease activity, which may account for the lack of such a correlation.

One of the primary strengths of our study is the large sample size coupled with comprehensive neuropsychological and clinical assessments, which enabled us to derive robust and reliable conclusions from the data. We meticulously considered the potential confounding effects of medications on serum levels, enhancing the validity of our findings. Our focused approach on CI adds to the study's specificity. Consistency was ensured by measuring all serum concentrations using the same matrix, and our evaluation of storage effects on serum levels showed no significant impact, thereby reinforcing the reliability of our measurements. A significant strength of our study is its longitudinal design, incorporating a follow-up assessment at one year. This enabled us to draw more precise conclusions about the relationships between serum levels of S100A8/A9 and MMP-9 with CI. Extending the duration of the follow-up period in future studies would solidify our findings and provide deeper insights into the temporal relationship between the levels of these analytes and CI.

Our study has some limitations. First, we measured concentrations solely in serum samples. Although serum measurement is more participant-friendly and clinically feasible compared to cerebrospinal fluid (CSF) measurement, it does not allow us to directly determine whether S100A8/A9 contributes to neuroinflammation. Additionally, several cytokines and chemokines relevant to NPSLE have been found to be significantly higher in CSF than in serum.[38] Therefore, our study cannot ascertain whether circulating S100A8/A9 migrates to the CNS and induces local inflammation or if the circulating levels are merely an epiphenomenon. Additionally, our study was conducted at a single center, potentially limiting the generalizability of the findings to other populations or settings. The study population also lacked diversity in terms of disease severity, which may affect the applicability of the results to patients with varying degrees of disease activity. However, this relatively homogeneous low disease activity may also be considered a strength, as it limits the influence of inflammation in other organs on S100A8/A9 levels, which could have otherwise impacted our ability to detect its association with CI. Finally, despite adjusting for potential confounders, there may still be unmeasured variables that could influence the relationship between S100A8/A9 and MMP-9 serum levels and CI.

In conclusion, our study underscores the possible involvement of S100A8/A9 in the immunopathogenesis of CI in adult patients with SLE. Further research is essential to elucidate the precise mechanism.

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 Touma 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. Hanly JG. Diagnosis and management of neuropsychiatric SLE. Nat Rev Rheumatol 2014;10(6):338–347.
2. Rayes HA, Tani C, Kwan A, et al. What is the prevalence of cognitive impairment in lupus and which instruments are used to measure it ? A systematic review and meta-analysis. Semin Arthritis Rheum 2018; 48(2):240–255.
3. Bingham KS, DiazMartinez J, Green R, et al. Longitudinal relationships between cognitive domains and depression and anxiety symptoms in systemic lupus erythematosus. Semin Arthritis Rheum 2021;51(6): 1186–1192.
4. Appenzeller S, Cendes F, Costallat LTL. Cognitive impairment and employment status in systemic lupus erythematosus: a prospective longitudinal study. Arthritis Rheum 2009;61(5):680–687.
5. Tam LS, Wong A, Mok VC, et al. The relationship between neuropsychiatric, clinical, and laboratory variables and quality of life of Chinese patients with systemic lupus erythematosus. J Rheumatol 2008; 35(6):1038–1045.
6. Calderón J, Flores P, Aguirre JM, et al. Impact of cognitive impairment, depression, disease activity, and disease damage on quality of life in women with systemic lupus erythematosus. Scand J Rheumatol 2017;46(4):273–280.
7. Mendelsohn S, Khoja L, Alfred S, et al. Cognitive impairment in systemic lupus erythematosus is negatively related to social role participation and quality of life: a systematic review. Lupus 2021;30(10): 1617–1630.
8. Seet D, Allameen NA, Tay SH, et al. Cognitive dysfunction in systemic lupus erythematosus: immunopathology, clinical manifestations, neuroimaging and management. Rheumatol Ther 2021;8(2):651–679.
9. Jeltsch-David H, Muller S. Neuropsychiatric systemic lupus erythematosus : pathogenesis and biomarkers. Nat Rev Neurol 2014;10(10): 579–596.
10. Santer DM, Yoshio T, Minota S, et al. Potent induction of IFN-alpha and chemokines by autoantibodies in the cerebrospinal fluid of patients with neuropsychiatric lupus. J Immunol 2009;182(2):1192–1201.
11. Yoshio, T. et al. IL-6, IL-8, IP-10, MCP-1 and G-CSF are significantly increased in cerebrospinal fluid but not in sera of patients with central neuropsychiatric lupus erythematosus. Lupus 2016;25(9):997–1003.
12. Okamoto H, Kobayashi A, Yamanaka H. Cytokines and chemokines in neuropsychiatric syndromes of systemic lupus erythematosus. J Biomed Biotechnol 2010;2010:268436.
13. Fragoso-Loyo H, Atisha-Fregoso Y, Nuñez-Alvarez CA, et al. Utility of TWEAK to assess neuropsychiatric disease activity in systemic lupus erythematosus. Lupus 2016;25(4):364–369.
14. Hirohata S, Kikuchi H. Role of serum IL-6 in neuropsychiatric systemic lupus erythematosus. ACR Open Rheumatol 2021;3(1):42–49.
15. Zarfeshani A, Carroll KR, Volpe BT, et al. Cognitive impairment in SLE: mechanisms and therapeutic approaches. Curr Rheumatol Rep 2021;23(4):25.
16. Muñoz-Grajales C, Barraclough ML, Diaz-Martinez JP, et al. Serum S100A8/A9 and MMP-9 levels are elevated in systemic lupus erythematosus patients with cognitive impairment. Front Immunol 2024;14:1326751.
17. Wang S, Song R, Wang Z, et al. S100A8/A9 in inflammation. Front Immunol 2018;9:1298.
18. Wu M, Xu L, Wang Y, et al. S100A8/A9 induces microglia activation and promotes the apoptosis of oligodendrocyte precursor cells by activating the NF-κB signaling pathway. Brain Res Bull 2018;143: 234–245.
19. Könnecke H, Bechmann I. The role of microglia and matrix metalloproteinases involvement in neuroinflammation and gliomas. Clin Dev Immunol 2013;2013:914104.
20. Hu W, Tao Z, Zhou Q. et al. Effects of S100 calcium-binding protein A8 (S100A8) and S100 calcium-binding protein A9 (S100A9) on matrix metalloproteinase (MMP) expression in nasopharyngeal carcinoma CNE-2 cells. Transl Cancer Res 2021;10(4):1874–1884.
21. Barraclough M, Erdman L, Diaz-Martinez JP, et al. Systemic lupus erythematosus phenotypes formed from machine learning with a specific focus on cognitive impairment. Rheumatology (Oxford) 2023; 62(11):3610–3618.
22. 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.
23. Langensee L, Mårtensson J, Jönsen A, et al. Cognitive performance in systemic lupus erythematosus patients: a cross-sectional and longitudinal study. BMC Rheumatol 2022;6(1):22.
24. Gladman DD, Ibañez D, Urowitz MB. Systemic lupus erythematosus disease activity index 2000. J Rheumatol 2002;29(2):288–291.
25. Gladman D, Ginzler E, Goldsmith C, et al. The development and initial validation of the Systemic Lupus International Collaborating Clinics/American College of Rheumatology damage index for systemic lupus erythematosus. Arthritis Rheum 1996;39(3): 363–369.
26. Kozora E, Ellison MC, West S. Reliability and validity of the proposed American College of Rheumatology Neuropsychological Battery for systemic lupus erythematosus. Arthritis Rheum 2004;51(5):810–818.
27. Tayer-Shifman OE, Green R, Beaton DE, et al. Validity evidence for the use of automated neuropsychologic assessment metrics as a screening tool for cognitive impairment in systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2020;72(12):1809–1819.
28. Zervides KA, Jern A, Nystedt J, et al. Serum S100A8/A9 concentrations are associated with neuropsychiatric involvement in systemic lupus erythematosus: a cross-sectional study. BMC Rheumatol 2022;6(1):38.
29. Bertolero MA, Yeo BTT, D'Esposito M. The modular and integrative functional architecture of the human brain. Proc Natl Acad Sci USA 2015;112(49):E6798–E6807.
30. Harvey PD. Domains of cognition and their assessment. Dialogues Clin Neurosci 2019;21(3):227–237.
31. Soyfoo MS, Roth J, Vogl T, et al. Phagocyte-specific S100A8/A9 protein levels during disease exacerbations and infections in systemic lupus erythematosus. J Rheumatol 2009;36(10):2190–2194.
32. Šumová B, Cerezo LA, Szczuková L, et al. Circulating S100 proteins effectively discriminate SLE patients from healthy controls: a cross-sectional study. Rheumatol Int 2019;39(3):469–478.
33. Homa-Mlak I, Mazurek M, Majdan A, et al. Serum calprotectin - a NET product - as a biomarker of disease activity in patients with systemic lupus erythematosus: a single-center case-control study from Poland. Med Sci Monit 2022;28:e936534.
34. Manfredi M, Van Hoovels L, Benucci M, et al. Circulating calprotectin (cCLP) in autoimmune diseases. Autoimmun Rev 2023;22(5):103295.

35. Salemme R, Peralta LN, Meka SH, et al. The role of NETosis in systemic lupus erythematosus. J Cell Immunol 2019;1(2):33–42.
36. Dieker J, Tel J, Pieterse E et al. Circulating apoptotic microparticles in systemic lupus erythematosus patients drive the activation of dendritic cell subsets and prime neutrophils for NETosis. Arthritis Rheumatol 2016;68(2):462–472.
37. Davies JC, Midgley A, Carlsson E, et al. Urine and serum S100A8/A9 and S100A12 associate with active lupus nephritis and may predict response to rituximab treatment. RMD Open 2020;6(2):e001257.
38. Lee J, Dhillon N, Pope J. All-cause hospitalizations in systemic lupus erythematosus from a large Canadian referral centre. Rheumatology (Oxford) 2013;52(5):905–909.

Arthritis Care & Research
Vol. 77, No. 11, November 2025, pp 1313–1322
DOI 10.1002/acr.25572

American College of Rheumatology
Empowering Rheumatology Professionals

Screening for Social Determinants of Health in Patients With Systemic Lupus Erythematosus: A Point-of-Care Feasibility Study

S. Sam Lim,[1] Vijay R. Nadipelli,[2] Michelle Bruno,[3] Daphne Lew,[4] Bernard Rubin,[5] Charmayne M. Dunlop-Thomas,[6] Kelsey Hulcher,[4] Kristi R. Mitchell,[3] Lydia Demetriou,[7] Jeff Berko,[3] Jason Bao,[8] and Alfred H. J. Kim[4]

Objective. Social determinants of health (SDoH) can impact outcomes but are not routinely screened for in US outpatient rheumatology clinics. This study determined the feasibility of routine point-of-care SDoH screening among patients with systemic lupus erythematosus (SLE) and associated barriers and facilitators at the physician, care team, and patient levels.

Methods. This observational, prospective, mixed-methods pilot study (GSK Study 219011) conducted in two large academic tertiary lupus clinics in the United States screened adults with SLE for SDoH over two weeks (institution 1, July to August 2023) and seven weeks (institution 2, August to October 2023). Patient demographics and patient-reported responses to questions covering up to eight SDoH domains chosen by participating institutions were collected, and an optional patient experience survey was conducted afterward. Participating physicians and care teams were asked questions on screening implementation, tool usability and comprehension, clinician acceptance, and facilitators of use. Transcripts were analyzed using thematic analysis.

Results. The study included 69 patients with SLE across both institutions; 65 completed the patient experience survey. SDoH screening was successfully implemented with minimal disruption to clinical workflow and was viewed as valuable by physicians, care teams, and patients. Reported facilitators to successful SDoH screening included institutional leadership buy-in to address health equity and a brief screening tool format (five minutes or less). Barriers included limited resources and insufficient time or training.

Conclusion. With an appropriately resourced and trained care team, successful routine SDoH screening in lupus clinics is feasible, valuable to clinicians and care teams, and effective for connecting patients to needed social resources.

INTRODUCTION

Systemic lupus erythematosus (SLE) is a chronic, multisystem autoimmune disease affecting more than 200,000 people in the United States.[1] Despite advances in treatment, SLE can lead to widespread organ damage, resulting in considerable morbidity and early mortality.[2] There is concern that health disparities result in differences in the incidence, treatment, and outcomes of patients with SLE.[3–5] SLE is nine times more prevalent among women than men, and the prevalence of SLE is more than two times higher in Black women compared with White women.[1] In addition, SLE disproportionally affects patients from underrepresented groups and of lower socioeconomic status, with ethnic minority groups and those with public insurance more likely to

Presented in part at the American College of Rheumatology Convergence 2024, Washington, DC, November 14–19, 2024.

Supported by GSK (GSK Study 219011). Atlas Clarity LLC operationalized the study; Atlas Clarity LLC's operationalization of the study was funded by GSK.

[1]S. Sam Lim, MD, MPH: Emory University School of Medicine and Grady Health System, Atlanta, Georgia; [2]Vijay R. Nadipelli, BPharm, MS, MHP: GSK, Philadelphia, Pennsylvania; [3]Michelle Bruno, MPP, Kristi R. Mitchell, MPH, Jeff Berko, MPH: Atlas Clarity LLC, San Francisco, California; [4]Daphne Lew, PhD, MPH, Kelsey Hulcher, RD, LDN, Alfred H. J. Kim, MD, PhD: Washington University School of Medicine, St Louis, Missouri; [5]Bernard Rubin, DO, MPH: GSK, Durham, North Carolina GSK, Durham, North Carolina, USA; [6]Charmayne M. Dunlop-Thomas, MS, MPH: Emory University School of Medicine, Atlanta, Georgia; [7]Lydia Demetriou, MSc: GSK, London, United Kingdom; [8]Jason Bao: Emory University School of Medicine, Atlanta, Georgia, and Duke University, Durham, North Carolina.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25572).

Author disclosures and graphical abstract are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25572.

Address correspondence via email to Vijay R. Nadipelli, BPharm, MS, MHP, at vijay.r.nadipelli@gsk.com.

Submitted for publication October 10, 2024; accepted in revised form April 24, 2025.

SIGNIFICANCE & INNOVATIONS

- It is known that social determinants of health (SDoH) can have a major impact on disparities in systemic lupus erythematosus (SLE) incidence and outcomes. This study examined the feasibility of, and barriers associated with, introducing a routine point-of-care screening among patients with SLE for SDoH in US outpatient rheumatology clinics.
- Screening for SDoH among patients with SLE can be successfully implemented in tertiary lupus specialty clinics with minimal disruption to clinical workflow.
- Organizational championing and leadership buy-in to address health equity were among the most important facilitators to successful SDoH screening, and appropriate resources in the form of lupus navigators and/or social workers are critical for addressing the social needs identified through the SDoH screening process.
- Study findings offer support for the wider adoption of SDoH screening at other specialty lupus clinics and serve as a first step for addressing the social risks that underpin health disparities present in SLE.

receive fragmented care, thus leading to worse clinical outcomes.[4,6] An increased risk of developing lupus nephritis and higher mortality rates also exist within these marginalized groups.[7,8]

Although the underlying causes of these disparities include genetic, epigenetic, and socioeconomic factors,[5] research suggests that challenges related to social determinants of health (SDoH) are prevalent among those with SLE and have a major impact on disparities in SLE incidence and outcomes.[3,9] SDoH, defined by the World Health Organization as "the conditions in which people are born, grow, live, work, and age," account for 30% to 55% of health outcomes and contribute to health inequalities across the globe.[10] In multiple countries, adverse SDoH conditions, such as lower socioeconomic status, lack of medical insurance, and racial segregation, have been associated with reduced medication adherence, poor retention of SLE care, increased SLE activity, and higher SLE-related organ damage.[9]

National health policy organizations call for health care providers, health plans, and wider community-based organizations to implement SDoH screenings and perform more comprehensive SDoH assessments.[11,12] Patients are also supportive of being screened for SDoH[13]; however, despite widespread acceptance of the role of SDoH in health outcomes and the call for increased SDoH screening, SDoH are not adequately or routinely screened for at the point of care in the United States. In a survey performed in 2017 to 2018, less than 25% of hospitals and 16% of physician practices in the US adequately screened for SDoH.[14]

Although SDoH collection occurs in a range of settings, from specialty clinics to large integrated health care systems,[15–17] few studies have examined SDoH screening processes in an SLE patient population. Results from a multihospital primary care–based integrated care management program showed that screening patients for SDoH effectively addressed barriers to health care that may have previously prevented them from accessing care and adhering to treatment.[18] Although this study demonstrated the utility of screening for SDoH in an SLE patient population, the program lacked a structured screening tool and was not specific to a lupus clinic. Given the impact of SDoH on the health outcomes of patients with SLE, efforts should be made to examine the feasibility of routine SDoH screening and collect data on SDoH for this patient population.

To address this gap, this pilot study aimed to determine the feasibility of routine point-of-care SDoH screening among patients with SLE in tertiary lupus clinics; the associated barriers and facilitators on the physician, care team, and patient levels; and the strengths and limitations of such screening. This real-world study also examined clinician and care team perceptions of the SDoH screening to better understand patient experience and perceptions of having their social risks assessed.

MATERIALS AND METHODS

Study design. This was an observational, prospective, mixed-methods study (GSK Study 219011) that collected quantitative and qualitative data in two large, academic tertiary lupus clinics in the United States (institutions 1 and 2). The study population consisted of the clinical care team at each of the two participating institutions and the patients with SLE who had clinic appointments during the implementation period. Care teams at each institution were asked to screen for SDoH among all adult patients with SLE by asking the patients to complete a questionnaire for the SDoH screening tool selected (described in the SDoH screening tools section). The SDoH screening tool was implemented over two clinic days for each physician, and the study was completed over a two-week period for institution 1 (July to August 2023) and a seven-week period for institution 2 (August to October 2023) to accommodate individual physician schedules.

Institutions and patients. The study selected two large, academic medical centers affiliated with safety net hospitals (ie, hospitals that provide significant levels of care to people who are uninsured or underinsured and cannot cover the cost of care from their own resources[19]). Institution 1, located in the Southeast region of the United States, is a public hospital system serving a population that is more than 40% uninsured. Institution 2, located in the Midwest United States, is a private nonprofit institution providing significant financial assistance and charity care. According to the Community Health Needs Assessment report published by each institution in 2022, poverty, lack of health care insurance, and crime and violence were major SDoH factors in the

surrounding communities.[20,21] Both institutions specialized in treating lupus, represented a diverse population of adult patients with lupus, had a volume of ≥30 patients with lupus per week, and included two or more rheumatologists treating adult patients with lupus. Additionally, both institutions did not already routinely use an SDoH screening tool in their lupus clinics. Care teams serving patients with lupus at the institutions participated in the study. Patients ≥18 years of age with a physician-confirmed SLE diagnosis and an appointment at one of the two participating institutions during the study period were eligible for study participation and were screened for SDoH.

Site assessments. A site profile questionnaire was completed by the principal investigators (PIs) before study implementation to gather basic characteristics of the institutions, such as clinic type, number and type of practicing clinicians, and average number of patients with lupus seen per week. During onboarding, the PIs also completed a 60-minute assessment interview to collect information about the institution's current (baseline) clinical workflow for patients with lupus before SDoH screening implementation and any workflow modifications to accommodate the use of the SDoH screening tool.

SDoH screening tools. Although there is no single gold standard SDoH screening tool,[22] the validated Accountable Health Communities Health-Related Social Needs screening tool by the Center for Medicare and Medicaid Innovation of the US Department of Health and Human Services Centers for Medicare and Medicaid Services[23] was the basis for the screeners used by the two participating institutions because this was the screening tool available in the Epic electronic medical record (EMR). The institutions adapted the SDoH screening templates available in their clinic systems in ways they felt were most appropriate for their patient population and workflow needs. The screening tool at institution 1 included four SDoH domains: financial stability, housing and utilities, transportation, and food security. The screening tool at institution 2 included eight domains: physical activity, stress, social connections, and depression in addition to the four domains included in the screening tool at institution 1. The screening tools used at each institution were both fixed-choice questionnaires (Supplementary Figure 1).

A study data collection form was used to document anonymized patient responses to the SDoH screening tool questions, along with patient demographics obtained from the EMR. Patients who completed the SDoH screening were provided with a brief, optional patient survey to gather their perceptions and experiences with the SDoH screening tool. This survey included questions on screening tool readability, comprehension, comfortability, and reactions to being asked SDoH questions. Patients were offered a nominal financial incentive (in the form of a gift card to a maximum value of $25 in accordance with site protocols) for completing the optional survey. Supplementary Figure 2 outlines the clinic workflow for routine SDoH screening implemented during this study.

Postscreening interviews. Semistructured 30-minute interviews with participating clinicians (ie, rheumatologists, medical assistants, nurses) and care team members involved in SDoH tool implementation (eg, research coordinators, lupus navigators, and social workers) were conducted. The majority of interviews were conducted within two weeks of study implementation; only a few interviews occurred more than four weeks after implementation due to scheduling conflicts. Participating members were asked to describe their experience using the tool, providing perceptions on workflow, timing, value, and patient reactions to being screened. Semistructured interviews with institutional leaders, such as quality improvement officers, were also conducted to gain a broader perspective on how implementation aligned with institution-wide goals.

Statistical analysis. This study used a mixed-methods approach. Quantitative data were collected using the site data collection form and optional patient experience survey. Descriptive statistics for quantitative measures are presented as frequencies and percentages. For survey questions, the number of valid responses to a given question was used as the denominator for calculating percentages. Stata (StataCorp, College Station, Texas, US) was used to conduct quantitative analyses. For qualitative data, content analysis was conducted on the semistructured interview responses. Interview recordings were transcribed, and transcripts were reviewed for accuracy and clarity before they were imported into a qualitative interview coding program (Dedoose, US) to analyze findings and categorize responses. Interview responses were coded for analysis to include the following domains: feasibility of SDoH screening, clinician acceptance of SDoH screening, patient experience, and organizational culture. Transcripts were coded, and the data were then transferred into a data matrix to facilitate thematic analysis.

Ethics statement and data availability. This study was conducted in accordance with the Declaration of Helsinki 2008. The study protocol and all the data collection tools were submitted to a central institutional review board (WCG IRB, formerly Western IRB) in line with the requirements and received "exempt" status because the study was a noninterventional screening/tool evaluation. Institution 1 determined that the study did not require institutional review board review because it did not meet the definition of human subject research (US Department of Health and Human Services) or clinical investigation (US Food and Drug Administration). Informed consent was not required because the study was noninterventional, and the study results and participant data were anonymized. For requests for access to anonymized subject-level data, please contact the corresponding author.

RESULTS

Participating institutions. At the time of study data collection, institution 1 had 714 patients with lupus, and institution 2 had 349 patients with lupus actively engaged in care. The lupus patient population at each institution was predominantly Black (82.1% [586 of 714] at institution 1, 61.0% [213 of 349] at institution 2) and female (90.1% [643 of 714] at institution 1, 87.1% [304 of 349] at institution 2). Institution 2 had more patients covered by commercial health insurance compared with institution 1 (77.1% [269 of 349] vs 21.0% [150 of 714]; Table 1). A total of 16 care team members across two institutions participated in the study, including six rheumatologists, two medical assistants, one nurse, four research coordinators, two lupus navigators, and one social worker (Supplementary Table 1).

Patient characteristics. During the study period, 69 (institution 1, n = 40; institution 2, n = 29) of a total of 75 patients with SLE who attended the clinic were screened for SDoH and included in the study (Table 2). At institution 1, two patients were not screened because they did not speak or read English. Four patients at institution 2 were not screened for the following reasons: declined because of time constraints (n = 2), did not speak or read English (n = 1), or had physical limitations (n = 1). Both institutions had similar proportions of female study participants (institution 1, 92.5% [37 of 40]; institution 2, 93.1% [27 of 29]); institution 1 had a higher proportion of study participants who were 18 to 39 years of age (55.0% [22 of 40]) and Black (80.0% [32 of 40]) compared with institution 2 (44.8% [13 of 29] and 65.5% [19 of 29], respectively; Table 2).

Overall, 77.5% (31 of 40) of study participants at institution 1 screened positive for at least one of four SDoH domains, and 93.1% (27 of 29) of participants at institution 2 screened positive for at least one of eight SDoH domains. The most commonly indicated social risks were financial strain (71.8% [28 of 39]) at institution 1 and social connection challenges (79.3% [23 of 29]) at institution 2 (Table 2).

SDoH screening feasibility. *Preimplementation planning: Baseline workflow practices.* Based on the preimplementation clinical workflow assessment interview, a few modifications to baseline workflow were required to accommodate the use of the

Table 1. Characteristics of the total lupus population at each participating institution

	Institution 1 (N = 714)	Institution 2 (N = 349)
Characteristics[a]		
Female, n (%)	643 (90.1)	304 (87.1)
Age, n (%)		
18–33 y	79 (11.1)	157 (45.0)
34–49 y	307 (43.0)	157 (45.0)
50+ y	328 (45.9)	35 (10.0)
Race and ethnicity, n (%)		
Black	586 (82.1)	213 (61.0)
White	57 (8.0)	133 (38.1)
Other	71 (9.9)	3 (0.9)
Health insurance, n (%)		
Government sponsored[b]	564 (79.0)	80 (22.9)
Commercial[c]	150 (21.0)	269 (77.1)
Uninsured	0 (0)	0 (0)

[a] Data were collected through the site profile questionnaire that each institution's principal investigator completed at the beginning of the study.
[b] For institution 1, "government sponsored" included Medicare/Medicaid/Dual eligible, and other (military or Veterans Affairs insurance); for institution 2, "government sponsored" included Managed Medicare/Medicare or Managed Medicaid/Medicaid.
[c] For institution 2, "commercial" included commercial insurance, managed care, worker's compensation, and other (eg, generic supplemental, ministry, and other private nonmajor commercial insurance).

Table 2. Study participant characteristics*

Characteristics	Institution 1 (N = 40)	Institution 2 (N = 29)
Female, n (%)[a]	37 (92.5)	27 (93.1)
Race, n (%)[a]		
Black	32 (80.0)	19 (65.5)
≥2 races	3 (7.5)	0 (0)
White	2 (5.0)	8 (27.6)
American Indian/Alaska Native	1 (2.5)	0 (0)
Asian	1 (2.5)	2 (6.9)
Other	1 (2.5)	0 (0)
Age, n (%)[a]		
18–29 y	10 (25.0)	3 (10.3)
30–39 y	12 (30.0)	10 (34.5)
40–49 y	6 (15.0)	10 (34.5)
50–59 y	5 (12.5)	3 (10.3)
60+ y	7 (17.5)	3 (10.3)
Positive screening for SDoH domain, n (%)[b]		
Patients screened positive for ≥1 SDoH domain	31 (77.5)	27 (93.1)
Financial resource strain	28[c] (71.8)	17 (58.6)
Food insecurity	20[c] (51.3)	9 (31.0)
Housing instability	14 (35.0)	9 (31.0)
Transportation challenges	6 (15.0)	2 (6.9)
Social connection challenges[d]	–	23 (79.3)
Physical inactivity[d]	–	22 (75.9)
Stress[d]	–	9 (31.0)
Depression[d]	–	9 (31.0)

* EMR, electronic medical record; SDoH, social determinants of health.
[a] Data were obtained from the study data collection log completed by each institution. Because of rounding of data to one decimal place, the percentages for different age categories add up to 99.9% for institution 2.
[b] For institution 1, if a patient marked a response on the questionnaire for any of the questions in each domain that indicated a social need, they were considered as having a "positive" screen result. For institution 2, the present study used the same scoring algorithm as the one used in institution 2's electronic medical record. Responses to all SDoH domains were optional; therefore, the denominator varies due to missing responses.
[c] N = 39.
[d] Domains that were only screened at institution 2.

SDoH screening tool. The biggest changes involved assigning individuals to distribute the screening tool and reminding physicians to review SDoH results in the EMRs and to discuss the results with patients.

Implementation: Staff resources to implement the SDoH screening tool. Staff resources used to implement the SDoH screening tool at both institutions minimally burdened the clinic practice. Institution 1 elected two medical assistants to administer a paper copy of the SDoH questionnaire to patients upon check-in and enter the responses into the EMR upon completion, reporting that the time needed to enter questionnaire responses to the EMR was "a couple of minutes." Institution

(A) Implementation: Staff resources to implement the SDoH screening tool
Care team member (Institution 1): "I think it was pretty easy and straightforward. Our patients are very much used to filling out surveys and questionnaires for other studies. So it went actually pretty well."
Medical assistant (Institution 1): "We do have a lot of patients that cannot read. They cannot comprehend. They cannot remember. So, it's a little bit difficult for them, but I helped them get through it as best I could to help them understand so at least they would answer the question[s]."
Research coordinator (Institution 2) reported that they "did not have any trouble with specifically administering [the tool], and every patient filled it out without trouble."

(B) Implementation: Impact of SDoH screening on clinical workflow
Physician (Institution 2): "It didn't impact my workflow. It was easy. It took a very short amount of time. There weren't any major time constraints."
Physician (Institution 2): "Sometimes you're waiting five-plus minutes [for a patient to finish the screener]. And if this happens with every patient, that ends up being about 50 minutes, which would be 2 additional patients that I could have squeezed in [to the day's clinic]."
Research coordinator (Institution 2): "Overall, the flow was good. [There were technical hiccups when the screener was first rolled out] but [after those were resolved], I don't think it disrupted workflow that much."
Medical assistant (Institution 1): "It really depends on the patient. With the older patient, it's going to take a little longer, but we were still able to do it in a timely manner."

(C) Implementation: Impact of SDoH screening on patient-provider interactions
Physician (Institution 1): "'If you don't mind, I'll just take a look at your answers.' And if I saw something concerning, I'd say, 'Well, I see that you mentioned you have some problem with housing. Would it be okay if we had someone from our team reach out to you about this?'"
Physician (Institution 1): "This whole experience has brought up discussions from both the patient and the clinician side that we would not have had before. And so there were certain instances when those issues were really heightened, that I don't think would have been brought up as early or if ever before, and that does impact kind of how people engage with the system and get access to different things, like medicines and stuff."
Physician (Institution 1): "There are the few that had very extreme, significant [social determinant] issues where we did spend most of the encounter on those issues, and I'm just guessing like maybe 5% of the encounters then over the two weeks. But I mean, even though that's an absolute small number, I mean, that was a very big issue for them that may not have otherwise come up in that way or as directly. So, I think it was helpful for those few individuals."

Figure 1. Care team and physician interview responses to different aspects of SDoH screening tool implementation. (A) Staff resources to implement the SDoH screening tool. (B) Impact of SDoH screening on clinical workflow. (C) Impact of SDoH screening on patient–provider interactions. SDoH, social determinants of health.

2 primarily had two nonclinician research coordinators assigned and trained to administer the SDoH questionnaire to patients via a tablet device, and the information was automatically uploaded into the EMR. In general, care team members administering the tool reported that it was easy to use and not burdensome, and they were able to assist any patients who had difficulty completing the questionnaire (Figure 1A).

Implementation: Impact of SDoH screening on clinical workflow and patient–provider interactions. Although physicians reported differing degrees of workflow disruption with SDoH screening, most care team members reported minimal impact on clinical workflow (Figure 1B). Although patients were not timed, care teams reported that patients typically spent 3 to 10 minutes completing the screening tool. This time was reportedly dependent on the patient's age, familiarity with technology (if a tablet was used), and health literacy level. For a few older adults or those with difficulty comprehending the questions, care team members were present to help them complete the questionnaire. Overall, physicians at both institutions reported that, on average, an additional 5 to 10 minutes were spent reviewing SDoH results with patients or waiting for patients to complete the questionnaire. Some physicians were able to direct patients to other care team members who could respond to their social risks and reported that the screening results provided deeper context about the challenges faced by patients with SLE (Figure 1C).

Facilitators of, and barriers to, SDoH screening in routine practice. Despite variation in SDoH screening implementation across both institutions, common facilitators and barriers to SDoH screening were observed. Key facilitators included having leadership and physician buy-in to address health equity for patients, ensuring necessary resources for developing and implementing the screening tool to minimize the burden on the clinic, using a brief screening tool with a completion time of five minutes or less, tailoring the screening tool and implementation process to meet clinic needs, and making screening data available before patient review (Figure 2A) to connect patients with relevant resources. Commonly reported barriers to SDoH screening included limited resources to address the identified social risks and provide patients with immediate actionable solutions, limited awareness of available resources among clinicians to address social risks, and insufficient time or training among clinicians to address identified risks (Figure 2B).

Clinician and care team perspectives. Clinician acceptance of the SDoH screening was relatively high. Participating rheumatologists and care team members recognized the value and importance of SDoH screening, even when additional effort was required on their part (Figure 3). Most care team members were willing to continue using the screening tool in their clinics, particularly if resources were available to address the identified

(A) Reported facilitators of routine SDoH screening
Having leadership and physician buy-in to address health equity for patients
Ensuring resources needed to develop and implement the screening tool create minimal burden on clinic
Tailoring the screening tool and implementation process to meet their clinic's needs
Having a non-physician care team member administer the screening tool to avoid disruptions to physician workflow
Using a brief screening tool (ideally ≤5 minutes completion time)
Having immediate availability of screening data in the EMR for physicians to review prior to patient visit

(B) Reported barriers of routine SDoH screening
Institutions having limited resources to address social risks
Clinicians having limited awareness of available resources to address social risks
Insufficient clinician time or training to address identified risks
Patients unavailable to take extra time to meet with a lupus navigator or social worker
Technical challenges with information technology/EMR

Figure 2. Summarized themes of reported (A) facilitators and (B) barriers to SDoH screening from qualitative interviews. EMR, electronic medical record; SDoH, social determinants of health.

Clinician acceptance and value of SDoH screening
Institution 1 Rheumatologist 1: "This whole experience has brought up discussions from both the patient and the clinician side that we would not have had before."
Institution 2 Rheumatologist 1: "I think it [screening for SDoH] opens up high potential. Whether or not we can realize that level of potential really depends on system resources." Co-PI: "...implementing any new changes is going to be hard, but the work is really important. And I think particularly with regards to SDoH, I think that this is something that we really need to add to the workflow, but it's not a trivial task."
Care team willingness to continue SDoH screening
Institution 1 Rheumatologist 2: "… But if there's nothing you can do to address it, it's almost like, why are we collecting this if we're not acting on it? So, for me, because I actually have somebody [lupus patient navigator program] that I can ask to call them and connect them with resources, I will continue to use it." Rheumatologist 3: "I definitely would like to do that [continue screening for SDoH], but I would want to have resources. I mean, to be able to provide them." Medical assistant: "I would definitely continue using it. I think we're all set for it."
Institution 2 Social worker: "I mean, it helps to know. Always does. I do think it's important to collect this data, and I hope we continue to do it." Research coordinator 1: "We are continuing. I do think that we've discussed...how to better address this and continue it to make sure that it's efficient and it's effective."

Figure 3. Clinician and care team perspectives. PI, principal investigator; SDoH, social determinants of health.

social risks (Figure 3). Having a lupus navigator and social worker available was viewed as particularly beneficial (Figure 4).

Patient perspectives. The overall response rate for the patient experience survey was 94.2% (65 of 69). The patient experience survey revealed that the majority of patients were "happy to be asked" about their SDoH challenges (75.4% [49 of 65]), whereas 24.6% (16 of 65) were neutral; no patient was unhappy to be asked those questions. Most patients found the SDoH screening tool "very easy" to understand (90.8% [59 of 65]) and were "very comfortable" answering the SDoH questions (80.0% [52 of 65]; Supplementary Table 2). Physical activity and stress were reported as the most discussed SDoH topics (each 50.0% [16 of 32]), followed by finances (37.5% [12 of 32]; Supplementary Table 2).

Participant suggestions for improved SDoH screening and essential components for successful SDoH screening. Streamlining the referral pathway to social support, educating providers on discussing SDoH with patients, and improving explanations about why SDoH are collected were among the key suggested improvements to SDoH screening. Additionally, providing patients with a dedicated space to complete their SDoH forms or allowing patients to complete forms before the clinic visit could ease SDoH screening processes. Improved wording of the screening tool and increased training of front desk staff were also suggested to facilitate routine SDoH screening. Overall, Supplementary Figure 3 highlights the essential system-, provider-, and patient-level components required for successful SDoH screening in a specialty clinic.

DISCUSSION

We found that routine SDoH screening can be successfully implemented in lupus specialty clinics with minimal disruption to clinical workflow and is viewed as valuable by patients and care teams. The large proportion of patients reporting social risks presented in this study further highlights the importance of screening for SDoH in the broader health care setting.

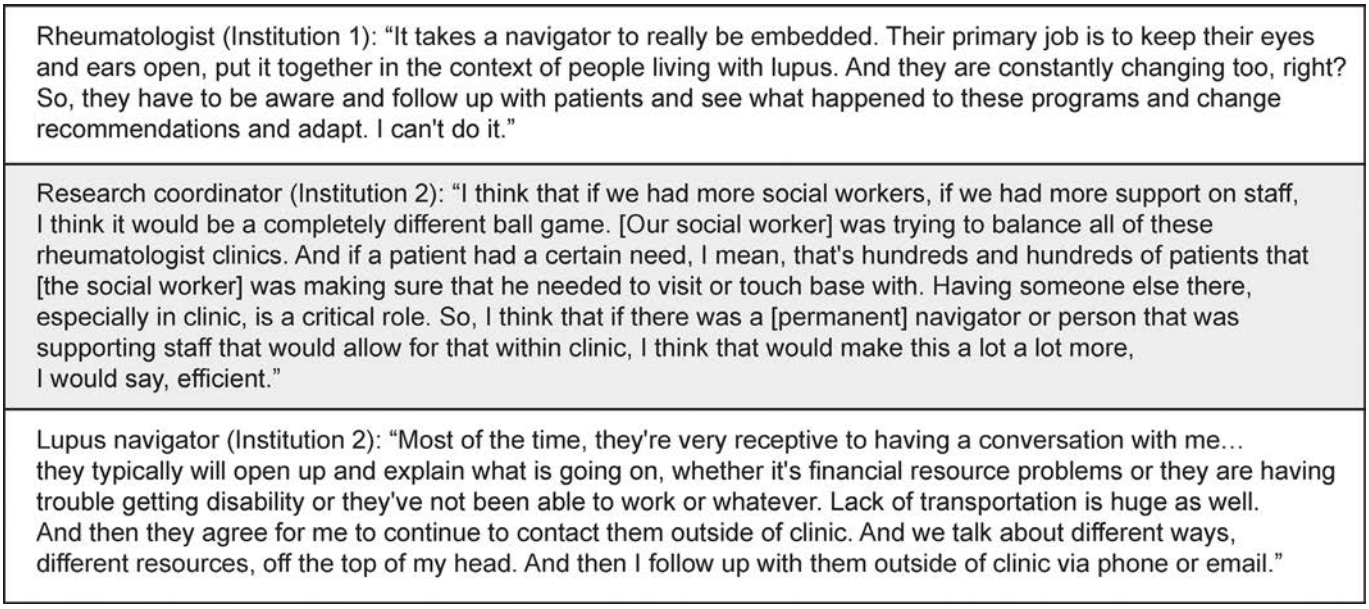


Figure 4. Care team interview responses on the screening benefit of having a lupus navigator or social worker.

Our study has identified a substantial social risk burden in patients with SLE and positive care team experiences with, and support of, SDoH screening in the two participating lupus clinics that are integral to major US health systems. Clinicians and care team members reported that SDoH screening aided their clinical decision-making. Patients reported an appreciation for being asked about their social situation and were comfortable completing the SDoH questionnaire, which they did in a timely manner. These findings are consistent with a previous study in a primary care setting, where a majority of patients agreed that their health system should ask about social needs and help to address them,[13] and a multihospital study reporting that screening and addressing SDoH in rheumatology clinics (not focused on SLE) is feasible.[24]

Leadership buy-in and the use of engaged and willing staff members to conduct the screenings were found to be among the strongest facilitators for SDoH screening implementation. This mirrors previous studies that identified the importance of strong leadership support on the overall success of SDoH screening implementation in hospital systems and community care settings.[17,25] Having leadership recognize the importance of SDoH in all patient populations may enable institutions and their staff to provide support by routinely asking patients about their social risks and SDoH challenges. These findings also serve as a reminder to community physicians and staff members that clinical adherence and response to therapy are often attributable to SDoH, reinforcing the need for consideration of patient-adverse SDoH conditions in everyday clinical practice.

In the current study, investment in social work resources was among the most reported critical components for responding to SDoH screening results. Given the limited awareness of available resources to address social needs among the clinicians reported here, the presence of a lupus navigator and social worker with knowledge of available resources for patients improved their connections to those services. As such, the present study suggests that having access to resources that address social needs is necessary to follow through with screening results to identify and remediate social risks. This has been reflected in another screening study by Schickedanz et al, in which 50% of physicians reported a lack of resources as a major barrier to successful screening.[26] Unfortunately, many lupus clinics do not have access to a full-time lupus navigator; durable funding and professionalization of the lupus navigator position may address this issue.

In our study, institution 1 included questions for four SDoH domains and institution 2 included questions for eight SDoH domains, leading to a degree of heterogeneity of the collected data. Although such heterogeneity is often unavoidable in a real-world clinical setting, standards of care that encourage SDoH screening must be established to effectively incorporate routine SDoH screening for the lupus patient population. Clinical guidelines for rheumatologic conditions should include specific recommendations that encourage routine SDoH screening with validated tools and provide guidance around the frequency and importance of screening. For example, various US institutions have recently released recommendations and call-to-action statements to encourage SDoH screening.[11,12] Creation and standardization of EMR templates to include routine SDoH screening would also aid broader adoption, aligning with national efforts to standardize SDoH screenings in EMRs.[27] Once viable models for SDoH screening in specialty clinics are available, it is recommended that professional societies incorporate training on

SDoH screening and ensuing patient discussions into Continuing Medical Education modules, a process that is already starting to take place.[28]

Organizations such as the American College of Rheumatology and the Lupus Foundation of America could increase private payers and government awareness of the need for SDoH screening and counseling for patients. Increased attention to adverse SDoH conditions will serve to improve patients' quality of life and may lead to improved disease outcomes. Additionally, patient organizations could develop resources to increase patient awareness of the impact of SDoH on health outcomes.[29] The continued movement toward value-based care provides a solid foundation to address SDoH while improving the quality and cost-effectiveness of care.[30] In 2021, the Centers for Medicare and Medicaid Services issued guidance to state health officials to increase the adoption of strategies in addressing SDoH in Medicaid beneficiaries.[31] The incentivization of health care providers to screen and refer patients for SDoH is increasing, with research suggesting that traditional Medicare may pay providers for furnishing SDoH risk assessment services.[11] Ultimately, beyond incentivizing institutions to conduct SDoH screenings, following up with results and providing resources to address social risks are equally important.

Several factors of this study limit its generalizability to other lupus care settings. Existing EMR templates for the SDoH screening tool, together with the appropriate level of education and training to support SDoH screening adoption and implementation, may not be readily available. Furthermore, the presence of dedicated lupus navigators is uncommon at most lupus clinics, and social workers are not present in many care settings. Both institutions were affiliated with safety net hospitals and excluded patients with limited English proficiency, potentially leading to population sample bias. Future studies should focus on larger and more diverse populations of patients with SLE for further data generalizability. As a critical component for addressing identified social risks, other institutions may need to employ a lupus navigator or social worker to successfully influence care management. Separately, the support from institutional leadership and buy-in for SDoH screening may not be generalizable, meaning settings without leadership support may need to highlight the value of SDoH screening to gain support and resources for implementation. It is also important to consider that many patients with lupus may obtain care with local community-based rheumatologists instead of tertiary lupus clinics, where resources and attitudes toward SDoH screening implementation may vary. Additional limitations included variation in the screening tool used across the institutions, the small sample sizes at the institutional level (two participating sites) and at the patient level (69 patients), and the absence of data on patients' education levels, which may be pertinent to their ability to understand and respond to the questions. Although small sample sizes are consistent with qualitative research and pilot studies, extending the observation period to enable increased patient volume could provide further insights into the long-term sustainability of the newly introduced workflows incorporating the SDoH screening tool.

In conclusion, this study demonstrated that with an appropriately resourced and trained care team, successful routine SDoH screening in specialty lupus clinics is feasible, valuable to clinicians and care teams, and effective for connecting patients to needed social resources. Study findings offer the opportunity for wider adoption of SDoH screening at other specialty lupus clinics and reinforce the evidence that SDoH data can inform patient–provider interactions, serving as a first step toward addressing identified social risks. Further research is needed to examine the connection between SDoH data collection; patient outcomes, experience, and care satisfaction; and the impact on holistic care management and treatment.

ACKNOWLEDGMENTS

We thank all the participating institutions, care team members, and individuals impacted by lupus who participated in this implementation study.

AUTHOR CONTRIBUTIONS

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

ROLE OF THE STUDY SPONSOR

GSK was involved in designing the study; contributed to the collection, analysis, and interpretation of the data; supported the authors in the development of the manuscript; and funded the medical writing assistance provided by Madeline Thomas, MRes (Fishawack Indicia Ltd, UK, part of Avalere Health). All authors, including those employed by GSK, approved the content of the submitted manuscript and were involved in the decision to submit the manuscript for publication. Publication of this article was not contingent upon approval by GSK. Medical writing support was provided by Madeline Thomas, MRes (Fishawack Indicia Ltd, UK, part of Avalere Health), and was funded by GSK.

REFERENCES

1. Izmirly PM, Parton H, Wang L, et al. Prevalence of systemic lupus erythematosus in the United States: estimates from a meta-analysis of the centers for disease control and prevention national lupus registries. Arthritis Rheumatol 2021;73(6):991–996.
2. Fanouriakis A, Tziolos N, Bertsias G, et al. Update on the diagnosis and management of systemic lupus erythematosus. Ann Rheum Dis 2021;80(1):14–25.

3. Buie J, McMillan E, Kirby J, et al. Disparities in lupus and the role of social determinants of health: current state of knowledge and directions for future research. ACR Open Rheumatol 2023;5(9):454–464.
4. Carter EE, Barr SG, Clarke AE. The global burden of SLE: prevalence, health disparities and socioeconomic impact. Nat Rev Rheumatol 2016;12(10):605–620.
5. Hasan B, Fike A, Hasni S. Health disparities in systemic lupus erythematosus-a narrative review. Clin Rheumatol 2022;41(11): 3299–3311.
6. Walunas TL, Jackson KL, Chung AH, et al. Disease outcomes and care fragmentation among patients with systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2017;69(9):1369–1376.
7. Maningding E, Dall'Era M, Trupin L, et al. Racial and ethnic differences in the prevalence and time to onset of manifestations of systemic lupus erythematosus: the California Lupus Surveillance Project. Arthritis Care Res (Hoboken) 2020;72(5):622–629.
8. Gómez-Puerta JA, Barbhaiya M, Guan H, et al. Racial/ethnic variation in all-cause mortality among United States Medicaid recipients with systemic lupus erythematosus: a Hispanic and Asian paradox. Arthritis Rheumatol 2015;67(3):752–760.
9. Williams JN, Drenkard C, Lim SS. The impact of social determinants of health on the presentation, management and outcomes of systemic lupus erythematosus. Rheumatology (Oxford) 2023;62(suppl 1): i10–i14.
10. World Health Organization. Social determinants of health. 2024. Accessed July 12, 2024. https://www.who.int/health-topics/social-determinants-of-health#tab=tab_1
11. US Department of Health and Human Services. HHS call to action: addressing health-related social needs in communities across the nation. November 16, 2023. Accessed July 12, 2024. https://aspe.hhs.gov/reports/hhs-call-action#:~:text=This%20HHS%20Call%20to%20Action,the%20goal%20of%20improving%20the
12. National Quality Forum. Connecting patients with community resources to improve health: NQF releases recommendations for using SDOH data. March 5, 2024. Accessed July 12, 2024. https://www.qualityforum.org/News_And_Resources/Press_Releases/2024/Connecting_Patients_With_Community_Resources_to_Improve_Health__NQF_Releases_Recommendations_for_Using_SDOH_Data.aspx
13. Rogers AJ, Hamity C, Sharp AL, et al. Patients' attitudes and perceptions regarding social needs screening and navigation: multi-site survey in a large integrated health system. J Gen Intern Med 2020; 35(5):1389–1395.
14. Fraze TK, Brewster AL, Lewis VA, et al. Prevalence of screening for food insecurity, housing instability, utility needs, transportation needs, and interpersonal violence by US physician practices and hospitals. JAMA Netw Open 2019;2(9):e1911514.
15. Kim RG, Ballantyne A, Conroy MB, et al. Screening for social determinants of health among populations at risk for MASLD: a scoping review. Front Public Health 2024;12:1332870.
16. Bradywood A, Leming-Lee TS, Watters R, et al. Implementing screening for social determinants of health using the Core 5 screening tool. BMJ Open Qual 2021;10(3):e001362.
17. Browne J, Mccurley JL, Fung V, et al. Addressing social determinants of health identified by systematic screening in a Medicaid accountable care organization: a qualitative study. J Prim Care Community Health 2021;12:2150132721993651.
18. Taber KA, Williams JN, Huang W, et al. Use of an integrated care management program to uncover and address social determinants of health for individuals with lupus. ACR Open Rheumatol 2021;3(5): 305–311.
19. Regenstein M, Nolan L, Wilson M, et al. Walking a Tightrope: The State of the Safety Net in Ten U.S. *Communities*. The George Washington University Medical Center; 2004. Accessed July 12, 2024. https://hsrc.himmelfarb.gwu.edu/cgi/viewcontent.cgi?article=1491&context=sphhs_policy_facpubs
20. Grady Health System. Community health needs assessment. 2022. Accessed July 12, 2024. https://www.gradyhealth.org/wp-content/uploads/2022-Community-Health-Needs-Assessment.pdf
21. Barnes-Jewish Hospital. Community health needs assessment and implementation plan. 2022. Accessed July 11, 2024. https://www.barnesjewish.org/About-Us/Community-Benefit/Community-Health-Needs-Assessment
22. Eder M, Henninger M, Durbin S, et al. Screening and interventions for social risk factors: technical brief to support the US preventive services task force. JAMA 2021;326(14):1416–1428.
23. Centers for Medicare & Medicaid Services. The accountable health communities health-related social needs screening tool. 2017. Accessed February 11, 2025. https://www.cms.gov/priorities/innovation/files/worksheets/ahcm-screeningtool.pdf
24. Feldman C, Summit R, Retzel K, et al. Implementation of a rheumatology-based social determinants of health screening program to uncover and address needs in a multihospital health care system [abstract]. Arthritis Rheumatol 2023;75(suppl 9). https://acrabstracts.org/abstract/implementation-of-a-rheumatology-based-social-determinants-of-health-screening-program-to-uncover-and-address-needs-in-a-multihospital-health-care-system/
25. Gruß I, Bunce A, Davis J, et al. Initiating and implementing social determinants of health data collection in community health centers. Popul Health Manag 2021;24(1):52–58.
26. Schickedanz A, Hamity C, Rogers A, et al. Clinician experiences and attitudes regarding screening for social determinants of health in a large integrated health system. Med Care 2019;57(6 suppl 2):S197–S201.
27. The White House. FACT SHEET: The Biden-Harris Administration announces more than $8 billion in new commitments as part of call to action for White House Conference on Hunger, Nutrition, and Health. September 28, 2022. Accessed July 12, 2024. https://bidenwhitehouse.archives.gov/briefing-room/statements-releases/2022/09/28/fact-sheet-the-biden-harris-administration-announces-more-than-8-billion-in-new-commitments-as-part-of-call-to-action-for-white-house-conference-on-hunger-nutrition-and-health/
28. Robeznieks A. Groundbreaking CME series details health inequities' root courses. American Medical Association. September 2, 2021. Accessed July 12, 2024. https://www.ama-assn.org/education/online-learning-cme/groundbreaking-cme-series-details-health-inequities-root-causes
29. Lupus Foundation of America. Health disparities in lupus. 2024. Accessed July 12, 2024. https://www.lupus.org/health-disparities.
30. Crook HL, Zheng J, Bleser WK, et al. How are payment reforms addressing social determinants of health? Policy implications and next steps. 2021. Accessed July 12, 2024. https://www.milbank.org/wp-content/uploads/2021/02/Duke-SDOH-and-VBP-Issue-Brief_v3-1.pdf
31. Center for Medicare & Medicaid Services. CMS issues new roadmap for states to address the social determinants of health to improve outcomes, lower costs, support state value-based care strategies. January 7, 2021. Accessed July 12, 2024. https://www.cms.gov/newsroom/press-releases/cms-issues-new-roadmap-states-address-social-determinants-health-improve-outcomes-lower-costs

Arthritis Care & Research
Vol. 77, No. 11, November 2025, pp 1323–1331
DOI 10.1002/acr.25576

American College of Rheumatology
Empowering Rheumatology Professionals

Systematic Review of Inequitable Population Representation in Systemic Lupus Erythematosus Clinical Trials

Seth Sims,[1] Kaylyn Rowsey,[1] Christian Hemmerich,[1] Haley Howard,[1] Jay Babek,[1] Garrett Jones,[1] Simran Demla,[2] Alicia Ito Ford,[1] and Matt Vassar[1]

Objective. This systematic review and meta-analysis aimed to evaluate the participation of historically marginalized populations in systemic lupus erythematosus (SLE) clinical trials conducted in the United States.

Methods. SLE, a complex autoimmune disease characterized by a dysregulated immune response leading to inflammation and tissue damage in multiple organ systems, exhibits a mortality rate four times higher in historically marginalized populations compared to the general population. It is essential for clinical trials to accurately represent the disease population to effectively evaluate treatment modalities. However, the current trial design lacks appropriate representation of historically marginalized populations, limiting the generalizability of results. Our study addresses this research gap by evaluating the participant demographics in SLE clinical trials. Relevant clinical trials were obtained in a comprehensive search of MEDLINE (PubMed) and Embase (Elsevier) in May 2024. Included trials were published in the United States between January 1, 2018, and December 31, 2023. Two reviewers independently performed screening and data extraction via a standardized Google Form.

Results. Having met our inclusion criteria, 18 US SLE clinical trials were evaluated for participant sex, age, racial, and ethnic data. Analysis of sex and gender revealed that the included population accurately represented the disease population. Regarding race and ethnicity participation, 11 of 18 studies (61.1%) received an overall poor rating, and none received a good rating. Analysis revealed that 14 of 18 studies (77.8%) demonstrated statistically insignificant underrepresentation of Black, Asian, and Hispanic populations. No studies reported the inclusion of older adults in their sample, suggesting a significant need for better age representation.

Conclusion. The results of this study reveal disparities in the representation of the SLE disease population in clinical trials, emphasizing insufficient inclusion of Black, Asian, and Hispanic and Latino participants and the disproportionate overrepresentation of White participants. Our study highlights the need for the initiation of effective strategies to engage historically marginalized populations in SLE clinical trials. Addressing these gaps is necessary to prioritize the participation of inequitable populations, increase standardization of SLE treatments, and improve the relevance of SLE research.

INTRODUCTION

Observed differences in incidence, prevalence, and disease complications regarding systemic lupus erythematosus (SLE) have led to increased awareness that racial and ethnic minorities tend to have a worse prognosis among affected patients.[1] SLE is a complex, chronic autoimmune disease defined by an interrupted immune response that leads to inflammation and tissue damage in various organ systems.[2] Although SLE affects people of all ages, genders, and races and ethnicities, with a prevalence that varies geographically, recent epidemiologic studies demonstrate higher rates of lupus among women, racial and ethnic minorities, and those of reproductive age.[3–5] Despite advancements in diagnosis and treatment, SLE remains a

[1]Seth Sims, BS, Kaylyn Rowsey, BS, Christian Hemmerich, BS, Haley Howard, BS, Jay Babek, BS, Garrett Jones, BS, Alicia Ito Ford, PhD, Matt Vassar, PhD: Oklahoma State University Center for Health Sciences, Tulsa; [2]Simran Demla, DO: Baylor Scott and White All Saints Medical Center, Fort Worth, Texas.

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.25576).

Author disclosures and graphical abstract are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25576.

Address correspondence via email to Seth Sims, BS, at seth.sims.research@gmail.com.

Submitted for publication October 21, 2024; accepted in revised form May 20, 2025.

SIGNIFICANCE & INNOVATIONS

- Our systematic review examines the representation of historically marginalized populations with systemic lupus erythematosus (SLE) in clinical trials developed in the United States. We found that high-risk populations—such as Black, Hispanic, and Asian individuals—were significantly underrepresented, whereas clinical trials predominantly enrolled White participants. This inequitable representation compromises the generalizability of interventions and outcomes among marginalized populations, further perpetuating health disparities and inequities in the United States.
- Using Agboola and Wright's diversity evaluation framework, we conducted a structured analysis of demographic factors that have been historically overlooked, quantifying disparities with the participation-to-disease representation ratio. This approach not only highlights existing inequities but also sets a precedent for future SLE research, encouraging targeted recruitment strategies to foster inclusivity.
- Our work extends beyond academic findings, urging researchers to prioritize diversity in study design to improve evidence-based treatment protocols that meet the needs of all patients. It adds to the growing call for reform in clinical research practices, advocating for more equitable trials and better health outcomes for marginalized populations.
- Our review exposes critical gaps in clinical trials representation, particularly among high-risk racial and ethnic minorities and older individuals, thus providing a framework to enhance diversity and equity in future SLE studies.

significant burden, with a considerable influence on patients' quality of life and mortality rates.

A range of structural and social determinants contribute to the variable health outcomes of patients with lupus, including lack of provider knowledge about SLE diagnosis and management, poor health care access associated with lower socioeconomic status, and environmental exposures.[6] Although participant diversity in clinical research is an important established consideration, many studies do not enroll and retain an appropriate sample of demographically representative participants.[7] This lack of representation compromises clinical knowledge, promoting false generalization, hindering innovation, and aggravating existing health inequalities.[8]

Given these challenges, this study specifically examines the representation of historically marginalized populations in SLE clinical trials conducted in the United States. Although global disparities exist, US-based research faces unique structural, historical, and policy-related barriers that impact recruitment and retention.[5,9–11] Despite calls for more equitable inclusion, no systematic review to date has comprehensively assessed this issue within the US context. Our study addresses this gap using a structured diversity rating framework to evaluate demographic representation in recent SLE trials.

Enhancing diversity in SLE research has far-reaching implications for clinical practice. Not only does it create better-informed treatment protocols and management strategies, but it can inform public health policies and resource allocation. For example, improved recruitment of historically marginalized populations in clinical trials has exposed how SLE affects these underresourced populations at a higher rate, causing the EULAR to update its treatment protocol in 2019 to include detailed instructions for treating SLE in different populations.[12] As a result of enhanced diversity in SLE research, efforts to combat the disease are effective and inclusive. Thus, promoting diversity in SLE research will drive more precise medical advancements, benefiting all patient populations affected by this complex condition.

Building on the importance of inclusive research, this article aims to investigate the diversity in SLE clinical trials conducted in the United States using the structured diversity evaluation framework developed by Agboola and Wright.[13] This framework offers a systematic approach to assess and enhance diversity in SLE clinical research by providing clear guidelines for evaluating the representation of various demographic groups and identifying gaps in study populations.

MATERIALS AND METHODS

Study design. To assess the diversity and representation of clinical research participants, we conducted a systematic search of the literature for SLE clinical trials and used the Clinical Trial Diversity Rating framework previously developed by Agboola and Wright[13] to evaluate the representation of research participants across different demographics. We then conducted a meta-analysis to characterize the diversity of SLE trials overall. Studies were selected and evaluated in a standardized manner, following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses 2020 checklist.[14] Our protocol, raw data, and findings are publicly available on Open Science Framework (OSF) to ensure reproducibility.[15] This study was determined to not include human participants by the institutional review board and therefore not subject to further research oversight according to the United States Code of Federal Regulations.

Search strategy. A detailed search string was created using the Cochrane Database to identify search terms from relevant review articles. We then identified relevant medical subject headings terms ("lupus erythematosus, cutaneous" and "lupus erythematosus, systemic") in PubMed, which we compiled into our search string.[15] With this search string, we searched major medical literature databases, MEDLINE (PubMed) and Embase (Elsevier), for all SLE interventional clinical trials. All returns were compiled and imported into Rayyan, a systematic review

screening platform (http://rayyan.qcri.org). Masked, two authors (SS and KR) screened the title and abstract while following the inclusion criteria discussed in the following section.

Eligibility criteria. Eligibility criteria for our meta-analysis included clinical studies that were (1) published between January 1, 2018, and December 31, 2023; (2) assessed an SLE intervention (behavioral, pharmaceutical, or otherwise); and (3) had trial site(s) only in the United States. We excluded publications outside the intended date range, trials without SLE as the focus, secondary database analyses, erratum, corrigendum, trial updates, and trial site(s) outside the United States. Two authors (SS and KR) reconciled their title and abstract screening decisions after the initial masked, duplicate screen. If a consensus could not be reached, a third investigator was available to arbitrate. While applying the same eligibility criteria, a final masked, duplicate full-text screening was performed. Two authors (SS and KR) reconciled their decisions again with a third investigator available to adjudicate.

Data collection. Two authors extracted data (SS and KR) using a standardized Google Form. General characteristics extracted were study title, year of publication, funding source, intervention type (pharmaceutical, behavioral, supplemental or holistic, and surgical or procedural), sample size, clinical trial phase (1, 2, 3, 4, or not reported), and trial site type (single center or multicenter).

Demographic data for race and ethnicity, sex, and age were first evaluated as "yes" or "no" if they were reported by the study. As referenced by Agboola and Wright,[13] these demographics were chosen due to them being the most common metrics used to report and evaluate participant diversity in previous research policies and publications. For demographic categories that were reported, specific data were then extracted as follows: race (proportions of White, Black, Asian, Pacific Islander and Hawaiian, and Native American and Alaskan Native), ethnicity (proportion of Hispanic and Latino), sex (proportion of male and female), and age ranges and means with SDs (overall and for participant subgroups).

Data analysis: diversity rating and descriptive statistics. To evaluate diversity, we applied the Clinical Trial Diversity Rating framework developed by Agboola and Wright.[13] This method uses the participation-to-disease representation ratio (PDRR) to compare trial participant demographics with known disease prevalence across race and ethnicity, sex, and age categories. Disease-specific prevalence rates for SLE were obtained from the Centers for Disease Control and Prevention (CDC) National Lupus Registry.[4] The PDRR is calculated as:

$$PDRR = \frac{Proportion\ of\ trial\ participants\ in\ a\ demographic\ group}{Disease\ prevalence\ in\ that\ group}$$

For example, if 40% of patients with SLE are male, but only 20% of a trial's participants are male, the PDRR for male participants in that trial would be 0.50 (0.20/0.40).

Based on previous literature, we used the following thresholds to interpret PDRR values: a score of ≥0.8 was classified as adequate representation, >1.2 as overrepresentation, and <0.8 as underrepresentation.[13,16,17] PDRR values were then converted into representation scores as follows:

- Three points for PDRR ≥0.8
- Two points for PDRR ≥0.5 and <0.8
- One point for PDRR >0 and <0.5
- Zero points for PDRR = 0

Scores for each demographic category (eg, Black, Hispanic, and White for race; male and female for sex) were summed to calculate a total representation score. Final ratings of "good," "fair," or "poor" were then assigned based on the percentage of the maximum possible score (eg, for sex, a maximum of 6 points if both male and female scores were 3). A worked example and formula breakdown are provided in the Supplementary Table 1.

Age-specific prevalence estimates were not incorporated into the final analysis due to inconsistent age band reporting across included trials. Although CDC-supported lupus registries offer age-stratified data across four US regions and the Indian Health Service, aligning trial-reported age data with these standards was not feasible. This limitation is discussed further in the article's Discussion section. For race and sex data, which are consistently reported in US trials, missing values were scored as zero to ensure uniform application of the framework.

Meta-analysis. Additionally, the PDRRs serve as the point estimate for our meta-analysis, capturing the relative representation of demographic groups across studies. The meta-analysis involved pooling representation scores across studies to obtain an overall diversity estimate and assess heterogeneity between studies. The SE was calculated using the individual study sample size as n. The upper and lower confidence intervals on the log scale were calculated using the natural log-transformed PDRR and the SE. The log transformation was necessary to allow for the calculation of confidence intervals of PDRR values because ratios often have skewed distributions. The confidence intervals were then back-transformed by taking the exponential of the upper and lower confidence intervals. The final PDRR and upper and lower 95% confidence intervals were used to create forest plots to visually illustrate the results of the meta-analysis. All original data, final reconciled data, and statistical analysis methods were uploaded to OSF for transparency and reproducibility.[15]

RESULTS

Trial inclusion and exclusion. Our comprehensive search returned 747 records. After deduplication, 597 records remained and underwent title and abstract screening. Following this initial screening, 132 records were selected for full-text review. Ultimately, 18 studies met our inclusion criteria for data extraction. Our exclusion reasons are detailed in Figure 1, and the full list of included studies is in Supplementary Table 1.

Trial characteristics. All analyzed trials were evenly distributed across the time frame of 2018 to 2023. Of these, most were funded by industry (12 of 18 studies, 66.7%) and used a pharmaceutical intervention (12 of 18 studies, 66.7%). Most studies were double-masked (9 of 18 studies, 50.0%), with six studies being open-label (6 of 18 studies, 33.3%). Additionally, 10 of the included studies had a sample size of less than 50 (10 of 18 studies, 55.6%). Additional characteristics can be found in Table 1.

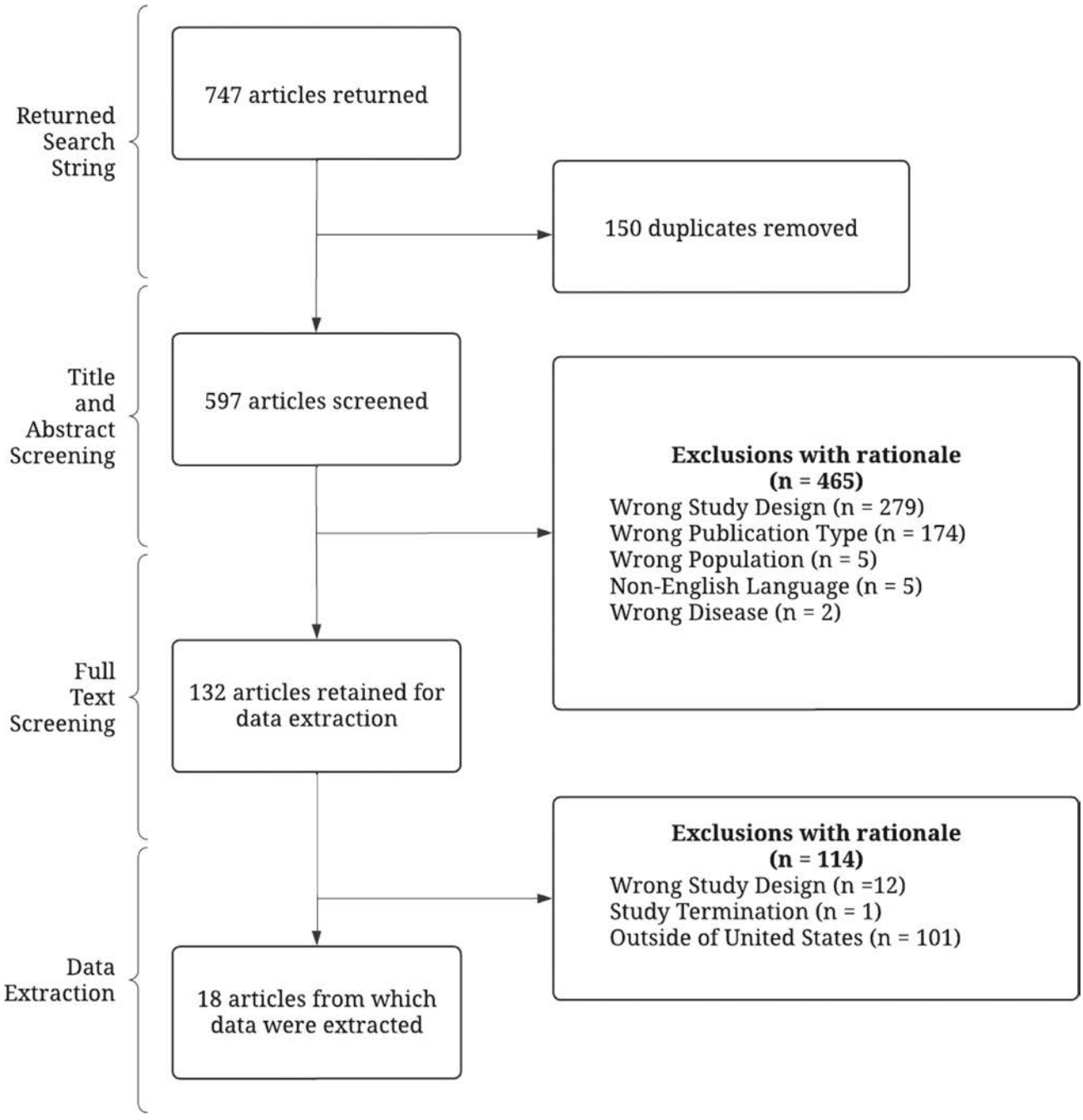


Figure 1. Flow diagram of study selection.

Table 1. Summary of characteristics of included studies (n = 18)*

Review characteristics	Studies, n/N (%)
Sample size	
<50	10/18 (55.6)
50–100	4/18 (22.2)
101–150	2/18 (11.1)
151–200	0
201–250	0
251–300	2/18 (11.1)
No. of centers conducting clinical trial	
Single center	9/18 (50)
Multicenter	8/18 (44.4)
Funding source	
No funding statement provided	0
Private	1/18 (5.6)
Industry	12/18 (66.7)
No funding received	0
Government	3/18 (16.7)
Hospital/university	1/18 (5.6)
Self-funded	0
Type of intervention	
Pharmaceutical	12/18 (66.7)
Behavioral	4/18 (22.2)
Supplemental/OTC	0
Surgical/procedural	2/18 (11.1)
Other	0
Clinical trial phase	
Phase 1	8/18 (44.4)
Phase 2	4/18 (22.2)
Phase 3	2/18 (11.1)
Phase 4	0
Not reported	4/18 (22.2)
Masking	
Single	2/18 (11.1)
Double	9/18 (50)
Triple	0
None	6/18 (33.3)

* Total N does not sum to 18 for all variables due to missing data in some trial publications. Categories with zero-value entries are retained for transparency. OTC, over the counter.

Diversity score. Supplementary Table 1 illustrates diversity ratings that relate to a comparison of SLE prevalence in the United States and trial representation of prevalent populations. Male and female representation was rated poor in 12 of 18 studies (66.7%), all due to underrepresentation of men, and 1 of 18 studies (5.5%) received a good rating. Eleven of 18 studies (61.1%) received a poor rating regarding race and ethnicity inclusion, and the remainder (7 of 18 studies, 38.9%) received a fair rating. No studies received a good rating regarding race and ethnicity. No studies reported age band data to be able to assess their inclusion of older adults.

Forest plot and meta-analysis. The forest plot showing racial and ethnic prevalence ratio by study (Figure 2) showcases trends in SLE trial representation. Of the 18 studies, 2 demonstrated statistically significant underrepresentation of Black, Asian, or Hispanic populations (2 of 18 studies, 11.1%), whereas 14 of 18 studies (77.8%) demonstrated underrepresentation of the same populations that was not statistically significant. The distribution of the White population demonstrated more than half (10 of 18 studies, 55.6%) with statistically significant overrepresentation, and 4 of 18 studies (22.2%) had overrepresentation that was not statistically significant. The included studies reported a distribution of male and female participants that mirrors the prevalence of SLE in the general population (Figure 3), with 2 of the total 18 trials (11.1%) falling outside of the 95% confidence interval range.

DISCUSSION

Our results suggest that the SLE disease population is not adequately represented in SLE clinical trials in the United States. Overall, groups from ethnic and racial minorities are underrepresented, whereas White groups were overrepresented. Accurate representation in clinical trials is crucial for the application of SLE intervention modalities to the general population, yet few trials have incorporated a diversity protocol to accommodate this gap in research. With this being said, we found that not a single trial in our sample of studies qualified to receive an overall good diversity rating for race and ethnicity. Of the 18 studies, 15 studies reported the race and ethnicity of their participants. Regarding Black participants, only 5 studies received a PDRR rating above 0.8, indicating that at least 10 studies reported Black participation below the disease prevalence in the general population. Conversely, regarding White participants, the majority (14 of 15 studies, 93.3%) received a PDRR score above 1.2, indicating overrepresentation of White participants, with 10 of these being statistically significant in the meta-analysis. Regarding sex, our results demonstrated an appropriate distribution of male and female participants relative to the disease population. Ultimately, our findings suggest that the disease prevalence of historically marginalized racial and ethnic groups in the general population was not adequately represented in US-based clinical trials for SLE interventions.

Our results echo previously reported trends regarding the underrepresentation of historically marginalized populations in clinical trial research. Hamel et al[18] noted that in cancer research, low enrollment of historically marginalized populations remains a problem despite National Institutes of Health requirements to include historically marginalized participants. Kwiatkowski et al[19] take this idea further, reporting that between 2001 and 2010—despite revised National Institutes of Health requirements—only 6.2% of clinical trial participants were Black, 3.3% of participants were Asian, and 2.2% of participants were Hispanic. In a narrative review of randomized controlled trials published between 1997 and 2017, Falasinnu et al[20] reported that although White individuals accounted for only 33% of SLE prevalence, they represented 51% of clinical trial participants; conversely, Black individuals made up 43% of disease prevalence but only 14% of participants. Our study found similar disparities, demonstrating that underrepresentation remains a persistent issue in recent years despite

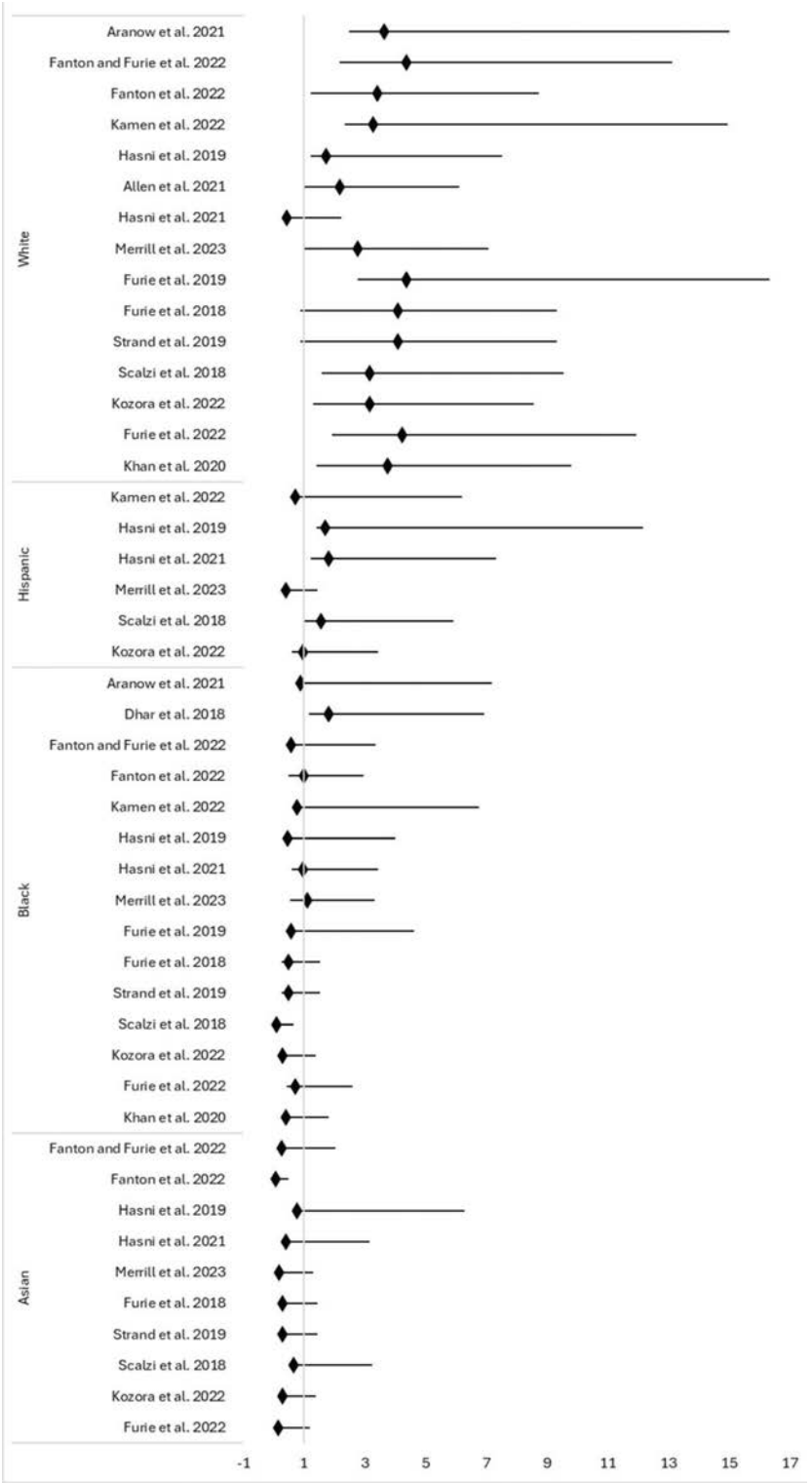


Figure 2. Forest plot of participation-to-disease representation ratios, with 95% confidence intervals (vertical line at one). Values <1 indicate underrepresentation; values >1 indicate overrepresentation.

growing attention to diversity. We intentionally began our analysis in 2018 to build on the findings by Falasinnu et al[20] and assess whether current studies improved demographic inclusivity. Our study found that female participants were accurately represented in SLE trials, but there was significant misrepresentations of Black, Asian, and Hispanic populations. Although the examples

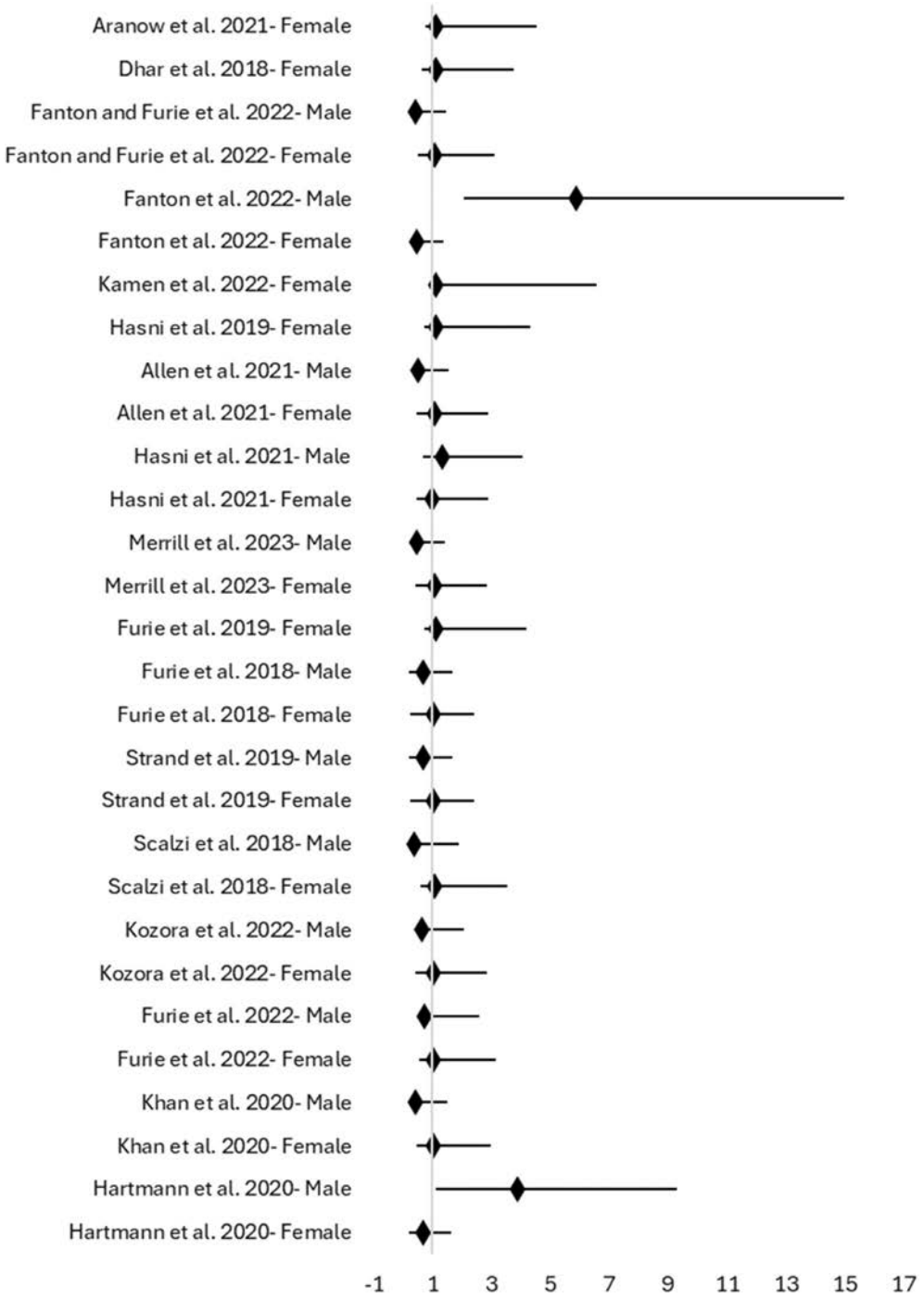


Figure 3. Forest plot of participation-to-disease representation ratios, with 95% confidence intervals (vertical line at one). Values <1 indicate underrepresentation; values >1 indicate overrepresentation.

previously mentioned are not all directly related to SLE, they underscore the broader issue of racial, ethnic, and sexual misrepresentation in clinical trials as a whole.

The underrepresentation of specific demographic groups in our sample of SLE clinical trials likely stems from a combination of structural, social, and historical factors. A systematic review by Ford et al[21] identified access-related barriers—such as limited information, distance from trial sites, inadequate health insurance, and low awareness of research opportunities—as the most commonly reported obstacles to clinical trial participation among marginalized populations. Religious and spiritual beliefs may also contribute; for instance, some individuals from historically marginalized backgrounds may attribute healing to divine intervention rather than medical treatment, which can influence their willingness to participate.[22] However, if these barriers can be breached, Wendler et al[23] found that individuals from historically

marginalized populations are just as likely to consent to research as their White counterparts.

Beyond these contemporary challenges, historically rooted injustices such as the Tuskegee Syphilis Study and redlining have created lasting mistrust in the health care system among Black communities and restricted access to academic medical centers.[9–11,24] These factors compound existing socioeconomic barriers and may partially explain persistent underrepresentation of racial and ethnic minorities in US-based SLE clinical trials. These historical influences are not only relevant to trial enrollment but also play a broader role in shaping the health outcomes of patients with lupus by perpetuating systemic inequities in diagnosis, treatment, and long-term care.

Considering our findings align with the broader literature on challenges in the involvement of historically marginalized populations in clinical trials, it is crucial to contemplate actionable steps for improvement. Studies have shown that providing clinical trials at outreach clinics would increase access to clinical trials for lower-income patients residing in rural communities, crossing the opportunity barrier to reach these populations.[25–27] Furthermore, studies have shown that efforts to increase community-wide education regarding clinical trials—via physicians, church groups, and patient advocacy—have proportionately increased trial participation, as opposed to an approach focused solely on education.[28–32] Another study by Advani et al shows that using nonwritten consent forms and informative aids has increased the recruitment of historically marginalized populations.[22] By adopting a modified recruitment framework, researchers can develop tailored strategies to improve participant diversity, ultimately leading to more robust and generalizable findings that benefit all patients.

Based on our findings, the aim of future research needs redirection to focus more on diversifying clinical trial cohorts, enhancing representation of the disease population. Evaluating different approaches taken to overcome misrepresentation in clinical trials will be valuable moving forward. These efforts could include patient education, outreach clinics, community groups, and patient advocacy. After effective methodologies and propositions have been identified, an archive of these efforts should be assembled and made available to clinical trialists to establish education and encourage the usage of these successful strategies. With their development and public availability, a step toward improvement of the PDRR ratio across all clinical trials will be made, benefiting both SLE research and the disease population.

Our study contains many strengths. The methodology is rigorous and reproducible because of our adherence to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines, which allows for transparency and credibility. We also used OSF as a data-sharing tool to further enhance the transparency of our study. Materials uploaded to OSF were done a priori, and any later changes were uploaded as amendments.[15] Another strength of our study is the comprehensive search strategy of major medical databases employed to include high-standard peer-reviewed publications. Our screening and extraction techniques were performed in a masked, duplicate manner by two independent authors who reconciled disagreements by discussion to avoid bias, which is the gold standard for meta-analysis research. Despite our strengths, this study presents several limitations, as well. We excluded studies performed outside the United States as well as studies not in English, which limits the application of our results in international countries and non-English speaking populations. By focusing exclusively on clinical trials conducted in the United States, our study allows for interpretation of underrepresentation trends within the framework of unique sociopolitical and historical influences on the US health care system. The study period was limited to 2018 to 2023 to reflect contemporary trial reporting standards and align with recent policy developments, including the 2022 US Food and Drug Administration mandate requiring diversity action plans in clinical development programs. Older trials may not have been subject to the same expectations for demographic transparency or diversity accountability, potentially skewing comparisons with present-day practices. Although expanding the date range could have increased the sample size, our focus was on capturing the most current representation practices.

Additionally, given the limited number of eligible trials within this time frame, stratified analyses by funding source, intervention type, or trial phase were not feasible. However, future investigations with broader temporal scopes or larger data sets may be better positioned to explore these dimensions and identify patterns relevant to industry-specific practices or trial characteristics. Although age-specific prevalence estimates for SLE are available through the CDC-supported lupus registries and the Indian Health Service, their integration into our analysis was limited by inconsistent reporting of age data across trials. Many studies did not provide participant age bands or reported only mean values without adequate distribution details. As a result, we were unable to generate reliable age-specific PDRR values. Future studies may be able to incorporate this metric if standardized age reporting becomes more common across clinical trials. Despite our careful search and extraction methods, it is possible that some relevant studies were not captured in our search or that data errors occurred. Future studies should address these limitations by including international studies and studies on non-English speaking populations. By including these more diverse sources, a more detailed examination of diversity scores could be obtained.

The findings of this study shed light on how the SLE disease population is inaccurately represented in clinical trials conducted in the US, highlighting the inadequate inclusion of people from Black, Asian, and Hispanic and Latino backgrounds and the disproportionate overrepresentation of White participants. Without accurate participant representation of the disease population,

these trials are not generalizable to the general SLE populace. Future studies should employ novel and intentional improvements to protocol and methods to ensure more inclusive results, which would, in turn, provide improved interventional strategies for SLE.

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, Mr Sims 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. Hasan B, Fike A, Hasni S. Health disparities in systemic lupus erythematosus-a narrative review. Clin Rheumatol 2022;41(11):3299–3311.
2. Siegel CH, Sammaritano LR. Systemic lupus erythematosus: a review. JAMA 2024;331(17):1480–1491.
3. Izmirly PM, Ferucci ED, Somers EC, et al. Incidence rates of systemic lupus erythematosus in the USA: estimates from a meta-analysis of the Centers for Disease Control and Prevention national lupus registries. Lupus Sci Med 2021;8(1):e000614.
4. Izmirly PM, Parton H, Wang L, et al. Prevalence of systemic lupus erythematosus in the United States: estimates from a meta-analysis of the Centers for Disease Control and Prevention national lupus registries. Arthritis Rheumatol 2021;73(6):991–996.
5. Fanouriakis A, Tziolos N, Bertsias G, et al. Update on the diagnosis and management of systemic lupus erythematosus. Ann Rheum Dis 2021;80(1):14–25.
6. Peschken CA. Health disparities in systemic lupus erythematosus. Rheum Dis Clin North Am 2020;46(4):673–683.
7. Turner BE, Steinberg JR, Weeks BT, et al. Race/ethnicity reporting and representation in US clinical trials: a cohort study. Lancet Reg Health Am 2022;11:100252.
8. National Academies of Sciences, Engineering, and Medicine; Policy and Global Affairs; Committee on Women; Bibbins-Domingo K, Helman A. Why Diverse Representation in Clinical Research Matters and the Current State of Representation Within the Clinical Research Ecosystem. National Academies Press (US); 2022.
9. Scharff DP, Mathews KJ, Jackson P, et al. More than Tuskegee: understanding mistrust about research participation. J Health Care Poor Underserved 2010;21(3):879–897.
10. Sistrunk C, Tolbert N, Sanchez-Pino MD, et al. Impact of federal, state, and local housing policies on disparities in cardiovascular disease in Black/African American men and women: from policy to pathways to biology. Front Cardiovasc Med 2022;9:756734.
11. National Academies of Sciences, Engineering, and Medicine; Health and Medicine Division; Policy and Global Affairs; Roundtable on Black Men and Black Women. Historical Roots of Medical Mistrust. In: Laurencin CT, Morgan RC Jr, Bright CM, eds. Leveraging Trust to Advance Science, Engineering, and Medicine in the Black Community: Proceedings of a Workshop. National Academies Press (US); 2025.
12. Fanouriakis A, Kostopoulou M, Alunno A, et al. 2019 update of the EULAR recommendations for the management of systemic lupus erythematosus. Ann Rheum Dis 2019;78(6):736–745.
13. Agboola F, Wright AC. A framework for evaluating the diversity of clinical trials. J Clin Epidemiol 2024;169:111299.
14. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Rev Esp Cardiol (Engl Ed) 2021;74(9):790–799.
15. Ghebrehiwet M, Clark P, Autaubo J, et al. Diversity. June 20, 2024. Accessed June 20, 2024. https://osf.io/89y6f/?view_only=d95fe828e09d45278735dbcf68f80bc8
16. Poon R, Khanijow K, Umarjee S, et al. Participation of women and sex analyses in late-phase clinical trials of new molecular entity drugs and biologics approved by the FDA in 2007-2009. J Womens Health (Larchmt) 2013;22(7):604–616.
17. Eshera N, Itana H, Zhang L, et al. Demographics of clinical trials participants in pivotal clinical trials for new molecular entity drugs and biologics approved by FDA From 2010 to 2012. Am J Ther 2015;22(6):435–455.
18. Hamel LM, Penner LA, Albrecht TL, et al. Barriers to clinical trial enrollment in racial and ethnic minority patients with cancer. Cancer Control 2016;23(4):327–337.
19. Kwiatkowski K, Coe K, Bailar JC, et al. Inclusion of minorities and women in cancer clinical trials, a decade later: have we improved? Cancer 2013;119(16):2956–2963.
20. Falasinnu T, Chaichian Y, Bass MB, et al. The representation of gender and race/ethnic groups in randomized clinical trials of individuals with systemic lupus erythematosus. Curr Rheumatol Rep 2018;20(4):20.
21. Ford JG, Howerton MW, Lai GY, et al. Barriers to recruiting underrepresented populations to cancer clinical trials: a systematic review. Cancer 2008;112(2):228–242.
22. Advani AS, Atkeson B, Brown CL, et al. Barriers to the participation of African-American patients with cancer in clinical trials: a pilot study. Cancer 2003;97(6):1499–1506.
23. Wendler D, Kington R, Madans J, et al. Are racial and ethnic minorities less willing to participate in health research? PLoS Med 2006;3(2):e19.
24. Corbie-Smith G. The continuing legacy of the Tuskegee Syphilis Study: considerations for clinical investigation. Am J Med Sci 1999;317(1):5–8.
25. Yeatts DE, Crow T, Folts E. Service use among low-income minority elderly: strategies for overcoming barriers. Gerontologist 1992;32(1):24–32.
26. McCabe MS, Varricchio CG, Padberg RM. Efforts to recruit the economically disadvantaged to national clinical trials. Semin Oncol Nurs 1994;10(2):123–129.
27. Lannin DR, Mathews HF, Mitchell J, et al. Influence of socioeconomic and cultural factors on racial differences in late-stage presentation of breast cancer. JAMA 1998;279(22):1801–1807.
28. Hughes GH, Cutter G, Donahue R, et al. Recruitment in the coronary artery disease risk development in young adults (Cardia) Study. Control Clin Trials 1987;8(4) (suppl):68S–73S.
29. Schoenberger JA. Recruitment experience in the Aspirin Myocardial Infarction Study. Control Clin Trials 1987;8(4) (suppl):74S–78S.
30. Borhani NO, Tonascia J, Schlundt DG, et al; Hypertension Prevention Trial Research Group. Recruitment in the hypertension prevention trial. Control Clin Trials 1989;10(3) (suppl):30S–39S.
31. Bjornson-Benson WM, Stibolt TB, Manske KA, et al. Monitoring recruitment effectiveness and cost in a clinical trial. Control Clin Trials 1993;14(2) (suppl):52S–67S.
32. Sheikh SZ, Wanty NI, Stephens J, et al. The state of lupus clinical trials: minority participation needed. J Clin Med 2019;8(8):1245.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 77, No. 11, November 2025, pp 1332–1339
DOI 10.1002/acr.25566

Carpal Tunnel Syndrome as an Early Underrecognized Feature of Rheumatoid Arthritis: A Population-Based Study of Carpal Tunnel Syndrome Occurrence Before and After Rheumatoid Arthritis Incidence

Roslin Jose George, Noah Frechette, Manuela Oviedo, Iqra Javed, Sara J. Achenbach, Ryan J. Lennon, Bradley Alex Kimbrough, Elena K. Joerns, Vanessa L. Kronzer, Anne Gingery, John M. Davis III , Cynthia S. Crowson, and Elena Myasoedova

Objective. We aimed to assess the occurrence of carpal tunnel syndrome (CTS) before and after rheumatoid arthritis (RA) incidence and by serologic status.

Methods. This population-based study included residents of a geographically defined area meeting the 1987 American College of Rheumatology classification criteria for RA in 1980 to 2019 matched 1:1 with individuals without RA. At least two diagnosis codes ≥30 days apart were used to identify CTS. Cumulative incidence of CTS adjusting for competing risk of death was assessed. Logistic regression and Cox proportional hazard models were used, adjusting for age, sex, calendar year, smoking, obesity, diabetes mellitus, and hypothyroidism.

Results. We included 1,335 patients with RA and 1,331 individuals without RA. The overall prevalence of CTS before or on RA incidence or index was 179 patients with RA (13%) and 85 individuals without RA (6%), respectively (odds ratio [OR] 2.23; 95% confidence interval [CI] 1.69–2.94). Most previous incidences of CTS occurred ≥2 years before the index date (112 events in patients with RA and 75 events in individuals without RA, respectively). Following RA incidence or index, individuals with RA (vs those without RA) had ∼80%-higher risk of CTS (hazard ratio [HR] 1.78; 95% CI 1.38–2.30). The risk estimates of CTS in patients with seronegative (vs seropositive) RA were OR 1.33 (95% CI 0.96–1.84) before RA incidence and HR 1.37 (95% CI 0.99–1.88) after RA incidence. In RA, obesity (HR 1.42, 95% CI 1.02–1.99) and seronegative cyclic citrullinated peptide antibody status (HR 1.79, 95% CI 1.07–2.99), but not other risk factors, were associated with increased CTS risk.

Conclusion. We found a more than two-fold increase in risk of CTS in the years preceding RA and a 1.8-fold increased risk of incident CTS following RA onset.

INTRODUCTION

Carpal tunnel syndrome (CTS) is the most common peripheral nerve entrapment syndrome, with multifactorial etiology and an estimated prevalence of approximately[1] 1% to 5%. Characterized by progressive fibrosis of the subsynovial connective tissue, resulting in compression or entrapment of the median nerve, CTS is more prevalent among the middle aged to older adult population and impacts twice as many women as men.[2–4] CTS is associated with several chronic conditions (eg, hypothyroidism, diabetes mellitus [DM]).[5–7] Among the various risk factors, inflammatory processes have also been associated with the pathology and severity of CTS.[8,9]

The content of this article is solely the responsibility of the authors and does not represent the official views of the NIH or the Mayo Clinic.

Supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases, NIH (grants R01-AR-46849, R01-AR-076347, and R01-AR-084831) and the National Institute on Aging, NIH (grants R01-AG-068192 and K24-AG-078179). This study used the resources of the Rochester Epidemiology Project medical records-linkage system, which is supported by the National Institute on Aging, NIH (AG-058738), by the Mayo Clinic Research Committee, and by fees paid annually by Rochester Epidemiology Project users.

Roslin Jose George, MBBS, MPH, Noah Frechette, BA, Manuela Oviedo, MD, Iqra Javed, MBBS, Sara J. Achenbach, MS, Ryan J. Lennon, MS, Bradley Alex Kimbrough, MD, RhMSUS, Elena K. Joerns, MD, MPH, Vanessa L. Kronzer, MD, MSCI, Anne Gingery, PhD, John M. Davis III, MD, MS, RhMSUS, Cynthia S. Crowson, PhD, Elena Myasoedova, MD, PhD: Mayo Clinic, Rochester, Minnesota.

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.25566).

Author disclosures and graphical abstract are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25566.

Address correspondence via email to Elena Myasoedova, MD, PhD, at Myasoedova.elena@mayo.edu.

Submitted for publication December 9, 2024; accepted in revised form April 3, 2025.

SIGNIFICANCE & INNOVATIONS

- In this study, patients with rheumatoid arthritis (RA) (vs those without RA) were two times more likely to develop carpal tunnel syndrome (CTS) in the years preceding the RA index date.
- Obesity and the absence of cyclic citrullinated peptide antibodies increased the risk of CTS in RA.
- Our findings highlight the potential utility of CTS as an early marker of "at risk of RA," including for seronegative RA, in which diagnosis has been historically delayed.

Because systemic inflammation is the hallmark of rheumatoid arthritis (RA), prolonged inflammatory processes affecting the anatomy of the carpal tunnel have been associated with the development of CTS in RA.[10,11] In addition to CTS, RA is associated with several other compression neuropathies such as tarsal tunnel syndrome, cubital tunnel syndrome, and Guyon canal syndrome; however, CTS is the most prevalent neuropathy in RA.[12] Previous small cross-sectional observational studies have identified various risk factors for CTS in patients with RA. These include longer disease duration, symptom severity, history of DM, and positive serologies of rheumatoid factor (RF) and anti–cyclic citrullinated peptide (CCP).[13,14]

Although there are several cross-sectional studies on the incidence of CTS in patients with RA,[11,13,14] CTS occurrence before RA diagnosis and as a potential sign of early RA is not well understood. This is because longitudinal clinical data before RA diagnosis are rarely available through registry and billing claims data sets. In addition, serologic status in RA and its potential connection to CTS has not been explored in population-based longitudinal cohorts. We aimed to assess the occurrence of CTS before and after RA incidence and by serologic status compared to individuals without RA. We hypothesized that CTS is more common both before and after RA disease onset, compared to the general population, and the risk is higher in seropositive patients with RA.

MATERIALS AND METHODS

Study design and population. This population-based study included residents of Olmsted County, Minnesota, using the Rochester Epidemiology Project (REP). REP is a records-linkage system with more than half a century of detailed longitudinal information on individuals from Olmsted and surrounding counties in Minnesota, USA.[15] Started in 1966, the REP links data from different institutions using various electronic health records, providing comprehensive demographic, social, and medical information of more than half a million unique individuals.[16]

Measures and outcome definitions. This study included individuals 18 years and older who met the 1987 American College of Rheumatology (ACR) classification criteria for RA between 1980 and 2019.[17] Patients with RA were matched 1:1 with individuals without RA by age, sex, and RA incidence or index year (further referred to as the "index date"). All individuals were observed until death, migration, or December 31, 2023. Data were abstracted by trained nurse abstractors and research trainees or fellows (RJG, NF, and IJ) who were masked to the study protocol and hypothesis. Sociodemographic information was collected for individuals with and without RA. DM and hypothyroidism were both defined by having at least two *International Classification of Diseases, Ninth Revision* (IDC-9) or *Tenth Revision* (ICD-10) codes for the diagnosis of interest at least 30 days apart from an electronic data pull.

The Charlson Comorbidity Index (CCI) was defined at the RA index date using diagnostic codes during the five years of previous medical history. The rheumatologic component of the CCI was excluded to ensure comparability between the cohorts.

For patients with RA, seropositivity was defined as positivity for RF and/or anti-CCP antibodies up to one year following the date RA criteria was met; patients without RF and CCP antibodies were considered seronegative. If the information on serostatus was not available, it was reported as missing data. We additionally collected data on the following RA disease characteristics: erosions or destructive joint changes on radiographs of hands and feet, large joint swelling, joint surgeries, and extraarticular manifestations, as previously described.[18] Information on antirheumatic medications, that is, conventional and synthetic disease-modifying antirheumatic drugs (DMARDs), biologic DMARDs, and glucocorticoids, was also gathered for individuals with RA.

CTS ascertainment: Incident CTS was defined as the presence of at least two ICD-9 or ICD-10 codes for CTS at least 30 days apart. Carpal tunnel release surgery was defined as the presence of one electronic procedure code. Carpal tunnel injections were defined as the presence of one Current Procedural Terminology (CPT) code (CPT = 20526): "Injection, Therapeutic (eg, local anesthetic, corticosteroid), Carpal Tunnel."

The study adhered to the Strengthening the Reporting of Observational Studies in Epidemiology guidelines for observational studies[19] and assents to the Declaration of Helinski. The analysis was approved by the institutional review boards (IRBs) of the Mayo Clinic (IRB no. 17-002593) and Olmsted Medical Center (IRB no. 017-OMC-17). Per Minnesota law, patients who declined research authorization for their medical records were not included in the analysis. The need for informed consent was waived.

Statistical analysis. The data were summarized using descriptive statistics. Logistic regression models were used to examine the association with CTS before RA diagnosis using a

case-control study design with (1) RA and (2) seronegativity in patients with RA as outcome measures. Cox proportional hazard models were used to compare the occurrence of CTS after RA the index date in a cohort study design among (1) patients with RA and without RA and (2) patients with RA who are seronegative versus seropositive. The models were adjusted for age, sex, calendar year, smoking, obesity, DM, and hypothyroidism. Cumulative incidence of CTS adjusting for competing risk of death was estimated. Individuals with CTS before index date were excluded from Cox models and cumulative incidence analyses. Analyses were performed using SAS version 9.4 (SAS Institute) and R version 4.2.2 (R Foundation for Statistical Computing).

RESULTS

We included 1,335 patients with RA (mean age 56.1 years, 69% female, 63% with RF or CCP seropositivity) and 1,331 individuals without RA. The mean medical history before the index date was 31.8 years in both cohorts. Descriptive statistics of the two groups are summarized in Table 1.

The overall prevalence of CTS at any time before or on the index date was 179 patients with RA (13%) compared to 85 individuals without RA (6%) (odds ratio [OR] 2.23; 95% confidence interval [CI] 1.69–2.94). Most previous incidences of CTS occurred at least two years before the index date (112 events in patients with RA and 75 events in individuals without RA). The plots in Figure 1 show that until approximately five years before the index date, the RA and non-RA cohorts had a similar known prevalence of CTS. Within five years of the index date, the difference in prevalence widened as the prevalence increased more sharply in the RA cohort (Figure 1). As seen from Supplementary Figure 1, medical and surgical data were available in >50% of individuals between 30 and 20 years preceding the index date, and in ≥75% of patients over the 20 years preceding the index date.

Among patients with RA, CTS before RA incidence was somewhat more common in RF- or CCP-negative (76 patients, 15%) than CCP-positive (103 patients, 12%), but this difference was not statistically significant (OR 1.33; 95% CI 0.96–1.84). Before or on the index date, 126 patients with RA (9%) had undergone carpal tunnel release surgery or carpal tunnel injection, compared to 50 individuals without RA (4%) ($P < 0.001$). Of them, 109 patients (8%) with RA had undergone carpal tunnel release surgery, compared to 47 individuals without RA (4%) ($P < 0.001$), and 38 patients with RA (3%) had undergone carpal

Table 1. Characteristics of patients with RA and individuals without RA*

Characteristics	RA (n = 1,335)	Non-RA (n = 1,331)
Age, mean (SD), y	56.1 (15.6)	56.1 (15.6)
Female, n/N (%)	917 (69)	914 (69)
Race, n/N (%)		
White (not Hispanic)	1,206/1,324 (91)	1,198/1,305 (92)
Not White	118/1,324 (9)	107/1,305 (8)
BMI		
Baseline BMI, mean (SD)	28.7 (6.7)	28.4 (6.5)
Obesity (BMI >30), n/N (%)	467/1,335 (35)	411/1,310 (31)
Smoking status, n/N (%)		
Current	264/1,335 (20)	209/1,330 (16)
Former	435/1,335 (33)	384/1,330 (29)
Never	636/1,335 (48)	737/1,330 (55)
Comorbidities, n/N (%)		
Diabetes mellitus	134/1,335 (10)	128/1,331 (10)[a]
Hypothyroidism	198/1,335 (15)	162/1,331 (12)
RA disease characteristics at RA incidence, n/N (%)		
RF positivity	749/1,318 (57)	–
Anti-CCP positivity	333/668 (50)	–
RF/anti-CCP positive	833/1,329 (63)	–
Erosions/destructive joint changes on radiographs, n/N (%)	346/1,335 (26)	–
Large joint swelling, n/N (%)	798/1,335 (60)	–
Severe Extraarticular manifestations, n/N (%)	67/1,335 (5)	–
ESR, median (IQR), mm/hr		
ESR at RA incidence	19.0 (8–35)	–
Highest ESR within the first year of meeting criteria	26.0 (12–43)	–
CTS,[b] n/N (%)	179/1,335 (13)	85/1,331 (6)
Carpal tunnel release surgery,[b] n/N (%)	109/1,335 (8)	47/1,331 (4)
Carpal tunnel injection,[b] n/N (%)	38/1,335 (3)	6/1,331 (<1)
Years of previous medical history, mean (SD)	31.8 (17.3)	31.8 (16.7)

* anti-CCP, anti–cyclic citrullinated antibody; BMI, body mass index; CTS, carpal tunnel syndrome; ESR, erythrocyte sedimentation rate; IQR, interquartile range; RA, rheumatoid arthritis; RF, rheumatoid factor.
[a] The comparators matched to more recent patients with RA did not have chart review done.
[b] CTS, CTS release surgery, and carpal tunnel injection before the index date.

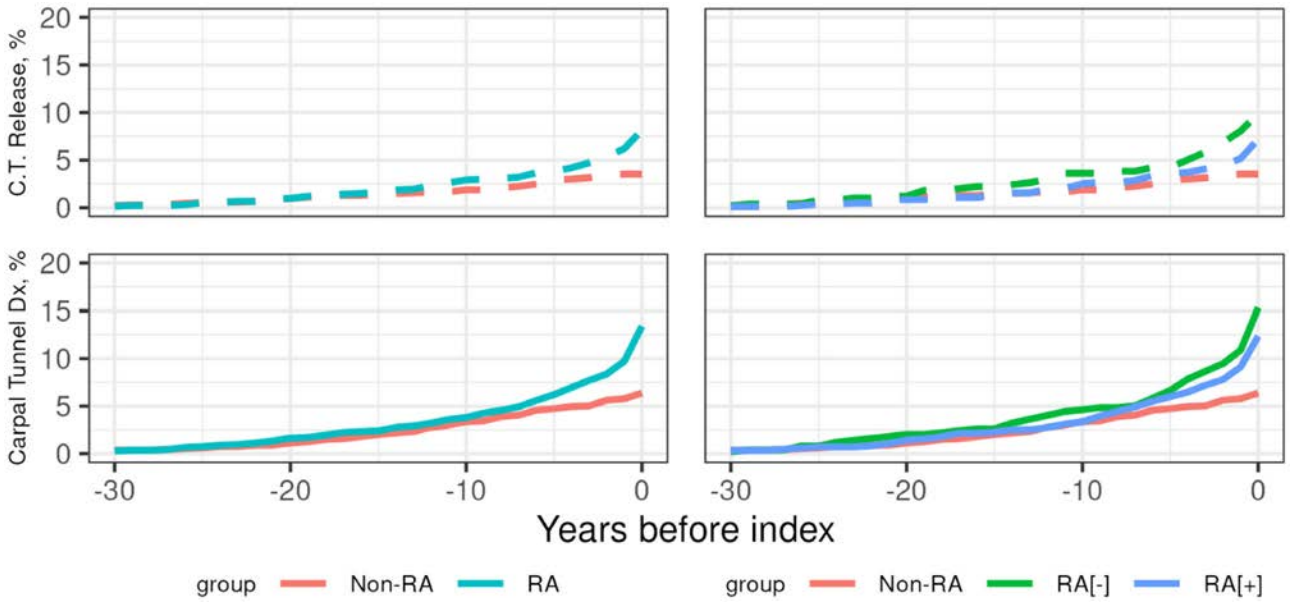


Figure 1. Summary statistics of timing of carpal tunnel syndrome (CTS) and CTS release surgery before the index date in individuals with RA and individuals without RA. The mean medical history before the index date was 31.8 years in the RA cohort and 31.8 years in the non-RA cohort. The plots show the percent of the cohort at yearly intervals before the index date who were known to have two previous diagnoses of CTS (solid lines) or carpal tunnel release (dashed lines). C.T., carpal tunnel; Dx, diagnosis; RA, rheumatoid arthritis.

tunnel injection, compared to 6 individuals without RA (<1%) ($P < 0.001$).

During median follow-ups of 12.8 and 13.8 years in RA and non-RA groups, respectively, 154 patients with RA and 102 individuals without RA developed CTS among those who did not have CTS before the index date. The cumulative incidence of CTS was 9.6% (95% CI 7.9–11.5%) at 10 years after the index date among patients with RA compared to 4.5% (95% CI 3.4–5.8%) among those without RA (Figure 2). Individuals with RA were 78% more likely to develop CTS, when compared to those without RA (hazard ratio [HR] 1.78, 95% CI 1.38–2.30).

We found a 36% increase in CTS after RA incidence in patients with seronegative versus seropositive RA, although the difference did not reach statistical significance (HR 1.36, 95% CI 0.99–1.88). We further assessed for risk factors associated with CTS in patients with RA adjusting for important confounders (Table 2). We found an increase in CTS in people with RA who were obese (HR 1.42, 95% CI 1.02–1.99) and seronegative for

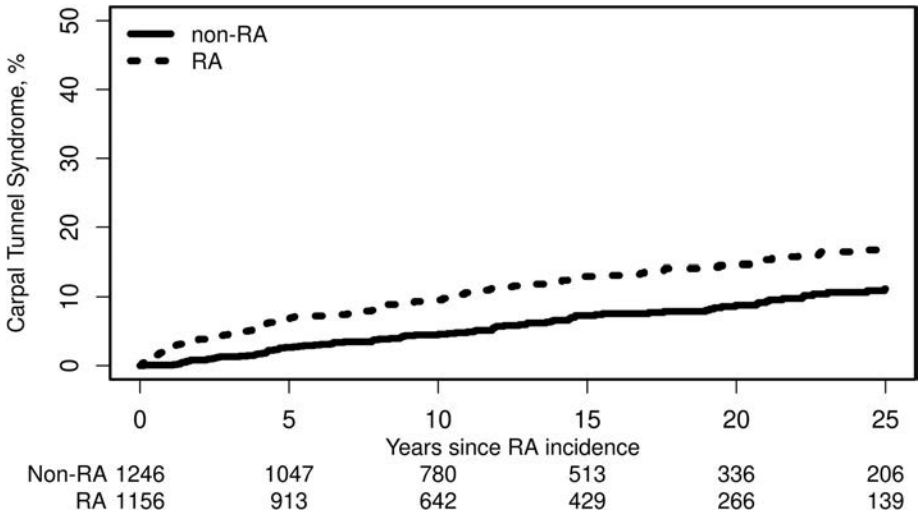


Figure 2. Cumulative incidence of carpal tunnel syndrome (CTS) in patients with RA adjusted for competing risk of death. The cumulative incidence of CTS was 9.6% (95% confidence interval [CI] 7.9–11.5%) at 10 years after index date among patients with RA compared to 4.5% (95% CI 3.4–5.8%) among those without RA. RA, rheumatoid arthritis.

Table 2. Predictors of carpal tunnel syndrome in patients with RA*

Parameter	Total (event)	HR (95% CI)[a]
Age at index per 10-y increase, y	1,156 (154)	0.99 (0.89–1.11)
Female	1,156 (154)	1.44 (0.99–2.10)
Year RA criteria was met per 5-y increase	1,156 (154)	1.01 (0.93–1.11)
Race (White)	1,145 (152)	**2.36 (1.03–5.38)**
More than a high school education	1,099 (148)	1.13 (0.79–1.61)
Former smoker	1,156 (154)	1.25 (0.88–1.80)
Current smoker	1,156 (154)	0.90 (0.57–1.43)
BMI per 5-kg/m^2 increase at index date	1,156 (154)	**1.22 (1.09–1.38)**
Obese (BMI ≥30) at baseline	1,156 (154)	**1.42 (1.02–1.99)**
RF negativity	1,141 (154)	1.37 (1.00–1.89)
Anti-CCP negativity	558 (66)	**1.79 (1.07–2.99)**
RF/anti-CCP negativity	1,150 (154)	1.36 (0.99–1.88)
Baseline ESR per 5-unit increase, mm/hr	1,125 (148)	0.96 (0.91–1.00)
Highest ESR per 5-unit increase,[b] mm/hr	1,108 (147)	0.97 (0.94–1.01)
Charlson comorbidity index	1,156 (154)	1.00 (0.88–1.12)
Radiographic erosions or destructive changes[c]	1,156 (138)	1.36 (0.95–1.95)
Rheumatoid nodules[c]	1,156 (138)	1.23 (0.81–1.88)
Large joint swelling[c]	1,156 (138)	1.12 (0.76–1.64)
Severe extra-articular manifestations of RA[c]	1,156 (138)	1.36 (0.75–2.47)
Count of different DMARD/biologic RA medications[c]	1,156 (138)	1.10 (0.97–1.26)

* Statistically significant associations are shown in bold. anti-CCP, anti–cyclic citrullinated peptide antibody; BMI, body mass index; CI, confidence interval; DMARD, disease-modifying antirheumatic drug; ESR, erythrocyte sedimentation rate; HR, hazard ratio; RA, rheumatoid arthritis; RF, rheumatoid factor.
[a] Adjusted for age, sex, calendar year, smoking, obesity, diabetes mellitus, and hypothyroidism.
[b] Within the first year of meeting criteria.
[c] Time-dependent variables.

CCP antibodies (HR 1.79, 95% CI 1.07–2.99). The associations with other risk factors were not statistically significant.

Following index date, patients with RA were two times more likely to undergo carpal tunnel release surgery or carpal tunnel injection compared to individuals without RA (HR 2.23, 95% CI 1.66–3.02). Among patients with RA, we found a statistically significant 50% increase in the incidence of carpal tunnel release surgery or carpal tunnel injection in patients with seronegative RA versus seropositive RA (HR 1.50, 95% CI 1.06–2.13).

DISCUSSION

This population-based study, leveraging more than 30 years of medical and surgical history before RA index date, evaluated occurrence of CTS and carpal tunnel release surgeries before and after the index date in individuals with versus without RA and by serostatus. We found that patients with RA were approximately twice as likely to be diagnosed with CTS any time before or on the index date compared to individuals without RA. Specifically, within the five years preceding the index date, CTS was more common in patients with RA than in individuals without RA, with a prevalence gap increasing toward the date of RA incidence. We also found increased rates of CTS-related interventions (ie, carpal tunnel release surgery or carpal tunnel injection) in patients with RA before the index date.

These major and significant findings of increased occurrence and surgical treatment of CTS years before the onset of RA indicate that CTS could be an early manifestation of RA or a component of "at risk of RA" state, rather than only a potential complication following RA onset, as previously suggested.[20–22] During a comprehensive literature search, we identified 13 previously published studies of CTS in RA (Supplementary Table 1). Among these previous studies, only four studies evaluated CTS before RA onset, and there was some inconsistency in findings. A case-control study from the UK Clinical Practice Research Datalink found that patients with RA had significantly increased medical consultations for CTS in the two years before RA onset compared to controls.[23] Concordantly, a population-based study with a database of 500 general practices from The Netherlands reported increased utilization of primary care services for musculoskeletal symptoms and CTS in the 1.5 years before RA diagnosis.[24] However, another study from The Netherlands retrospectively evaluating a group of consequent patients with CTS from two hospital-based cohorts between 1990 and 2000 found that CTS was rarely a presenting symptom or predictor of connective tissue diseases such as RA, which contradicts our finding.[25] Albeit, this study had a smaller sample size, shorter duration of follow-up, and no comparison group (Supplementary Table 1: de Rijk et al[25]). Another recent single-center study from Japan reported that patients with a previous diagnosis of CTS at the time of RA diagnosis had distinct characteristics (old age, female sex, and negative anti-CCP antibodies) compared to the non-CTS group and underscored the need for increased vigilance for RA in patients with CTS; however, there were no men in the CTS group, and the median age of both groups were in the older age category (73 years for the CTS group and 64 years for the non-CTS group).[26]

In this study, carpal tunnel release surgery and carpal tunnel injection were significantly more common among individuals with RA before or on the index date. Carpal tunnel release surgery, in which the transverse carpal ligament is cut to relieve pressure on the median nerve, is performed in more severe and persistent CTS cases.[5] The risk of carpal tunnel release surgery in years preceding RA incidence has not been previously studied in population-based cohorts. A previous study found that, among those who underwent carpal tunnel release surgery, patients with diagnosed RA had an increased risk of undergoing additional revision carpal tunnel release.[27] However, another study from Olmsted County reported decreased risk of reintervention and surgery in patients with RA who had undergone previous glucocorticoid injection treatment for CTS.[28] Although these previous

studies focused on carpal tunnel release surgery after RA diagnosis, our findings of the likelihood of carpal tunnel release surgery or carpal tunnel injection preceding index date extend these results and suggest a more severe and persistent presentation of CTS not only in patients with diagnosed RA but also in the years before the index date as compared to the general population.

Secondly, concordant with previous small (<100 patients with RA), cross-sectional and retrospective, hospital-based studies, we found CTS to be more likely to occur in individuals with RA following the index date.[6,25,29] A recent two-sample Mendelian randomization study using data from genome-wide association studies of CTS and RA established that a genetic susceptibility to RA could increase the likelihood of CTS; however, they found no causal relationship between CTS and increased risk of RA.[30] Nevertheless, not all studies agree with our findings; a retrospective cohort study found no difference in the prevalence of CTS between patients with RA and the general population.[31] However, the study had 176 patients from a referral center, with only approximately 55% of patients participating in the long-term follow-up and no comparison group from the same population (Supplementary Table 1: Lee et al[31]).

A third important finding of this study was that CTS tended to occur more likely before and after the onset of seronegative RA compared with seropositive RA, although these associations were not statistically significant. Patients with seronegative RA were found to be significantly more likely to have had previous carpal tunnel release surgery or carpal tunnel injection, suggesting potentially greater severity of CTS in patients with seronegative versus seropositive RA. Furthermore, in our study, being negative for CCP antibodies was associated with an increased risk of CTS in RA, adjusting for age, sex, calendar year, smoking, obesity, DM, and hypothyroidism. Similar to our findings, a recent small study from Japan found that patients with RA with CTS versus those without CTS at the time of diagnosis were characterized by negative anti-CCP antibodies; however, there were only a few cases of CTS in this study (n = 12), and all of them occurred in older adult women, thus the generalizability of their findings is limited and the findings of this study and our study cannot be directly compared.[26]

Nonetheless, several previous studies have found that RF and anti-CCP positivity was more prevalent in patients with RA with CTS when compared to patients with RA without CTS,[11,14,29] possibly aligning with the general impression of a higher disease activity and severity in seropositive versus seronegative RA. This contrasts with our finding that patients with seronegative RA had slightly higher CTS prevalence, which persisted even after adjusting for age, sex, smoking, obesity, DM, and hypothyroidism and contrasts with our hypothesis that seropositive RA is more likely to associate with CTS than seronegative RA. Possible explanations for the discrepancy include considerably smaller sample sizes over shorter time periods in the previous studies. In addition, the gestalt of lower RA disease activity in seronegative versus seropositive RA has been recently debated, and some studies have shown that patients with seronegative RA have higher disease activity at the time of RA onset and a lower chance for remission attainment after RA onset, in part due to the delayed diagnosis and delayed initiation of DMARDs.[32,33] Moreover, the ACR/EULAR 2010 classification criteria lack sensitivity in identifying seronegative RA,[34,35] which can potentially lead to underreporting of seronegative RA in the studies and delays in identification of seronegative RA in the clinical practice. Our finding of slightly higher risk of CTS and carpal tunnel release surgery or carpal tunnel injection in seronegative RA further underscores that seronegative RA status does not always imply a milder disease and warrants early diagnosis and management, with CTS being a potential early "at risk for RA" clinical feature.

Among the predictors for CTS in patients with RA, apart from seronegative status for CCP antibodies, we found that a higher body mass index and obesity were associated with an increased risk of CTS (adjusted for age, sex, calendar year, smoking, obesity, DM, and hypothyroidism). Obesity and metabolic syndrome have previously been associated with increased risk and severity of CTS.[36–38] When considering other factors such as the relationship between smoking and CTS, previous studies have found conflicting evidence on the risk of CTS among smokers.[39,40] Our study did not find an increased risk of CTS among patients with RA who are current or former smokers. Additional factors such as RA severity and RA medication use were not statistically significant risk factors for CTS among patients with RA in our study.

Our study also found that the prevalence of CTS was highest during the 2000 to 2009 decade. This may be attributed to improved diagnosis and identification methods during this decade compared to the previous two decades. Although serostatus was not a significant risk factor for CTS, the occurrence of CTS tended to be higher in seronegative patients. This could have played a role in the increase in CTS cases in recent years, as the incidence of seronegative RA has also increased in recent years.[41,42]

This study's findings have multiple practical implications, including that individuals diagnosed with CTS should have increased awareness for RA risk. Identifying possible warning signs of RA is important because early diagnosis and treatment improves the chances of remission and can reduce the risk of joint damage and disability.[43] Thus, increased vigilance is needed in individuals with CTS regarding the risk of developing RA. Referral to a rheumatologist can be considered, particularly in persistent and severe forms of CTS (eg, patients requiring carpal tunnel release surgery), that are not otherwise explained by traditional risk factors (eg, obesity, hypothyroidism, DM) and regardless of serostatus. Recognizing CTS as potential early feature of RA can be particularly helpful in facilitating earlier evaluation and care in individuals with seronegative RA who are at a higher

risk of delayed diagnosis and treatment compared to those with seropositive RA.[33]

Strengths of this study include the availability of complete records of patients with RA and CTS with data before and after index dates. The availability of more than 30 years of comprehensive health care data before the RA index date through the REP provided information on all clinical encounters with any provider for each patient, thus uniquely strengthening our study. The population-based cohort study design with long and complete follow-up of patients with RA and comparators without RA across four decades (1980–2019) further strengthens the study and provides real-world evidence on occurrence of CTS.

This study has several limitations. The data used to conduct this study were from Olmsted County and surrounding counties in Minnesota, USA, that are predominantly White. Therefore, the findings may not be representative of larger, more diverse populations, and hence limit the generalizability of our study. In this study, we used code-based definitions for CTS ascertainment. As patients with milder CTS may not be seeking medical care, there is potential for health care utilization bias. Additionally, we did not collect information on unilateral versus bilateral CTS during record review and information on swollen and tender joint counts, which may have provided a better understanding of CTS severity and its association with RA (ie, in cases with bilateral involvement, higher joint count score). Future studies involving a larger, more diverse population would help our understanding of the relationship between CTS and RA and could inform risk stratification and earlier treatment for patients.

Our study shows that CTS was two-fold more common in individuals with RA (vs individuals without RA) two years or more before the index date, suggesting that it can be an early underrecognized clinical feature of RA disease continuum, in other words, an "at risk of RA" state. The incidence of CTS after index date is significantly higher in patients with RA versus controls without RA. Obesity and seronegative CCP antibodies status increased the risk of CTS in RA. Carpal tunnel release surgery or carpal tunnel injection were more likely in individuals with RA (vs individuals without RA) both before and after the index date. These findings suggest the need for increased awareness of heightened risk of RA in patients with persistent CTS and consideration for rheumatology referral in patients at risk of RA with CTS that is not explained by traditional risk factors.

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 Myasoedova 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. Padua L, Cuccagna C, Giovannini S, et al. Carpal tunnel syndrome: updated evidence and new questions. Lancet Neurol 2023;22(3):255–267.
2. Werner RA, Andary M. Electrodiagnostic evaluation of carpal tunnel syndrome. Muscle Nerve 2011;44(4):597–607.
3. Gelfman R, Melton LJ III, Yawn BP, et al. Long-term trends in carpal tunnel syndrome. Neurology 2009;72(1):33–41.
4. Chikenji T, Gingery A, Zhao C, et al. Transforming growth factor-β (TGF-β) expression is increased in the subsynovial connective tissues of patients with idiopathic carpal tunnel syndrome. J Orthop Res 2014;32(1):116–122.
5. Osiak K, Elnazir P, Walocha JA, et al. Carpal tunnel syndrome: state-of-the-art review. Folia Morphol (Warsz) 2022;81(4):851–862.
6. Kaya Subaşı P, Güler T, Yurdakul FG, et al. Carpal tunnel syndrome in patients with rheumatoid arthritis and psoriatic arthritis: an electrophysiological and ultrasonographic study. Rheumatol Int 2021;41(2):361–368.
7. Donnelly JP, Hanna M, Sperry BW, et al. Carpal tunnel syndrome: a potential early, red-flag sign of amyloidosis. J Hand Surg Am 2019;44(10):868–876.
8. Omole AE, Awosika A, Khan A, et al. An integrated review of carpal tunnel syndrome: new insights to an old problem. Cureus 2023;15(6):e40145.
9. Arshad MS, Mattoo B, Alam I. Exploring pathogenic pathways in carpal tunnel syndrome: sterile inflammation and oxidative stress. J Basic Clin Physiol Pharmacol 2024;35(3):189–198.
10. Deane KD, Holers VM. The natural history of rheumatoid arthritis. Clin Ther 2019;41(7):1256–1269.
11. Smerilli G, Di Matteo A, Cipolletta E, et al. Ultrasound assessment of carpal tunnel in rheumatoid arthritis and idiopathic carpal tunnel syndrome. Clin Rheumatol 2021;40(3):1085–1092.
12. DeQuattro K, Imboden JB. Neurologic manifestations of rheumatoid arthritis. Rheum Dis Clin North Am 2017;43(4):561–571.
13. Karadag O, Kalyoncu U, Akdogan A, et al. Sonographic assessment of carpal tunnel syndrome in rheumatoid arthritis: prevalence and correlation with disease activity. Rheumatol Int 2012;32(8):2313–2319.
14. Dede BT, Oğuz M, Bulut B, et al. Carpal tunnel syndrome evaluation with ultrasound in rheumatoid arthritis patients. ARP Rheumatol 2023;2(4):330–337.
15. St Sauver JL, Grossardt BR, Yawn BP, et al. Use of a medical records linkage system to enumerate a dynamic population over time: the Rochester Epidemiology Project. Am J Epidemiol 2011;173(9):1059–1068.
16. Rocca WA, Grossardt BR, Brue SM, et al. Data resource profile: expansion of the Rochester Epidemiology Project medical records-linkage system (E-REP). Int J Epidemiol 2018;47(2):368–368j.
17. Arnett FC, Edworthy SM, Bloch DA, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum. 1988;31(3):315–324.
18. Kimbrough BA, Crowson CS, Davis JM III, et al. Decline in incidence of extra-articular manifestations of rheumatoid arthritis: a population-based cohort study. Arthritis Care Res (Hoboken) 2024;76(4):454–462.
19. von Elm E, Altman DG, Egger M, et al; STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol 2008;61(4):344–349.

20. Greenblatt HK, Kim HA, Bettner LF, et al. Preclinical rheumatoid arthritis and rheumatoid arthritis prevention. Curr Opin Rheumatol 2020; 32(3):289–296.
21. Young A, Koduri G. Extra-articular manifestations and complications of rheumatoid arthritis. Best Pract Res Clin Rheumatol 2007;21(5): 907–927.
22. Deane KD, Holers VM, Emery P, et al. Therapeutic interception in individuals at risk of rheumatoid arthritis to prevent clinically impactful disease. Ann Rheum Dis 2025;84(1):14–28.
23. Muller S, Hider S, Machin A, et al. Searching for a prodrome for rheumatoid arthritis in the primary care record: a case-control study in the clinical practice research datalink. Sem Arthritis Rheum 2019;48(5): 815–820.
24. Beers-Tas MV, Nielen MM, Twisk JWR, et al. Increased primary care use for musculoskeletal symptoms, infections and comorbidities in the years before the diagnosis of inflammatory arthritis. RMD Open 2020;6(2):e001163.
25. de Rijk MC, Vermeij FH, Suntjens M, et al. Does a carpal tunnel syndrome predict an underlying disease? J Neurol Neurosurg Psychiatry 2007;78(6):635–637.
26. Nakamura T, Nagira K, Nakagawa N, et al. Characteristics of patients presenting with concomitant carpal tunnel syndrome at the initial diagnosis with rheumatoid arthritis. Mod Rheumatol 2024;34(5):960–965.
27. Westenberg RF, Oflazoglu K, de Planque CA, et al. Revision carpal tunnel release: risk factors and rate of secondary surgery. Plast Reconstr Surg 2020;145(5):1204–1214.
28. Evers S, Bryan AJ, Sanders TL, et al. Corticosteroid injections for carpal tunnel syndrome: long-term follow-up in a population-based cohort. Plast Reconstr Surg 2017;140(2):338–347.
29. Mahmoud W, El-Naby MMH, Awad AA. Carpal tunnel syndrome in rheumatoid arthritis patients: the role of combined ultrasonographic and electrophysiological assessment. Egypt Rheumatol Rehabil 2022;49(1):62.
30. Gong C, Zhao D, Wen X, et al. Causal relationship between rheumatoid arthritis and carpal tunnel syndrome: a bidirectional two-sample Mendelian randomization study. J Orthop Surg Res 2024;19(1):613.
31. Lee KH, Lee CH, Lee BG, et al. The incidence of carpal tunnel syndrome in patients with rheumatoid arthritis. Int J Rheum Dis 2015; 18(1):52–57.
32. Nordberg LB, Lillegraven S, Lie E, et al; and the ARCTIC working group. Patients with seronegative RA have more inflammatory activity compared with patients with seropositive RA in an inception cohort of DMARD-naïve patients classified according to the 2010 ACR/EULAR criteria. Ann Rheum Dis 2017;76(2):341–345.
33. Coffey CM, Crowson CS, Myasoedova E, et al. Evidence of diagnostic and treatment delay in seronegative rheumatoid arthritis: missing the window of opportunity. Mayo Clin Proc 2019;94(11):2241–2248.
34. Kaneko Y, Kuwana M, Kameda H, et al. Sensitivity and specificity of 2010 rheumatoid arthritis classification criteria. Rheumatology (Oxford) 2011;50(7):1268–1274.
35. Kasturi S, Goldstein BL, Malspeis S, et al. Comparison of the 1987 American College of Rheumatology and the 2010 American College of Rheumatology/European League Against Rheumatism criteria for classification of rheumatoid arthritis in the Nurses' Health Study cohorts. Rheumatol Int 2014;34(3):407–411.
36. Balci K, Utku U. Carpal tunnel syndrome and metabolic syndrome. Acta Neurol Scand 2007;116(2):113–117.
37. Salamh F, Habib SS, AlRouq F, et al. Relationship of obesity indices with clinical severity and nerve conduction studies in females presenting with median nerve compression neuropathy at the wrist. J Family Med Prim Care 2024;13(4):1291–1295.
38. Mi J, Liu Z. Obesity, type 2 diabetes, and the risk of carpal tunnel syndrome: a two-sample Mendelian randomization study. Front Genet 2021;12:688849.
39. Pourmemari MH, Viikari-Juntura E, Shiri R. Smoking and carpal tunnel syndrome: a meta-analysis. Muscle Nerve 2014;49(3):345–350.
40. Lampainen K, Hulkkonen S, Ryhänen J, et al. Is smoking associated with carpal tunnel syndrome? A meta-analysis. Healthcare (Basel) 2022;10(10):1988.
41. Myasoedova E, Davis J, Matteson EL, et al. Is the epidemiology of rheumatoid arthritis changing? Results from a population-based incidence study, 1985-2014. Ann Rheum Dis 2020;79(4):440–444.
42. Matthijssen XME, Huizinga TWJ, van der Helm-van Mil AHM. Increasing incidence of autoantibody-negative RA is replicated and is partly explained by an aging population. Ann Rheum Dis 2022;81(5):e69.
43. Demoruelle MK, Deane KD. Treatment strategies in early rheumatoid arthritis and prevention of rheumatoid arthritis. Curr Rheumatol Rep 2012;14(5):472–480.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 77, No. 11, November 2025, pp 1340–1349
DOI 10.1002/acr.25583

Predicting the Impact of Air Quality Index on Rheumatoid Arthritis Disease Activity

Ahmad Alsaber,[1] Adeeba Al-Herz,[2] Maryam Al-Mazedi,[3] Iyad Abu Doush,[4] Afsah Abdullah,[1] Dhary Alkandary,[5] Ahmad T. Al-Sultan,[6] Parul Setiya,[7] and Jiazhu Pan[8]

Objective. This study explored the relationship between air pollution and rheumatoid arthritis (RA), focusing on how specific air quality components affect RA disease activity as measured by the Disease Activity Score in 28 joints (DAS28).

Methods. The research involved data that were obtained from six air-monitoring stations across Kuwait, and information on patients with RA was obtained from the Kuwait Registry for Rheumatic Diseases. This study analyzed the impact of pollutants such as sulfur dioxide, nitrogen dioxide (NO_2), ozone (O_3), particulate matter with a diameter of 10 micrometers or less, nitrogen oxide, and carbon monoxide on the DAS28.

Results. The results highlighted that NO_2 and O_3 were the most significant Air Quality Index components influencing DAS28 scores. NO_2 had a lag effect of two months ($P < 0.01$, effect score = 0.43), whereas O_3 exhibited a lag effect of three months ($P < 0.05$, effect score = 0.31), both correlating with increased RA disease activity. The study used a Vector Error Correction Model and cointegration analysis to examine short- and long-term associations between predicted and actual DAS28 scores was adjusted over the following year using air quality index, indicating that strong long-term cointegration with the error correction term was negative and significant (–0.54, $P < 0.001$).

Conclusion. These findings emphasize the importance of air quality management in mitigating the impact of environmental factors on RA, suggesting that exposure to elevated levels of NO_2 and O_3 beyond regulatory standards could exacerbate RA symptoms. This study provides a foundation for future public health interventions aimed at improving disease prognosis in patients with RA by addressing environmental factors, such as air pollution.

INTRODUCTION

Numerous researchers have extensively examined the effects of ambient air pollution on human health.[1,2] Studies have shown the impact of air pollution on the development of several ailments including respiratory disorders, cardiovascular diseases, and cancer.[3] Air pollution, even at low levels, may aggravate or even cause a host of respiratory and other ailments such as diabetes, asthma, bronchitis, rheumatoid arthritis (RA), and chronic obstructive pulmonary disease. Public health has been concerned about the health effects of air pollution exposure for approximately 700 years.[4] The majority of studies on pollutants in the air and case studies related to health research have concentrated on the effects of short-term, brief exposure as opposed to long-term, prolonged exposure.

RA is a chronic autoimmune illness characterized by continuous joint inflammation that may lead to pain, swelling, and joint damage. Numerous environmental, genetic, and immunologic factors are involved in the complex pathophysiology of RA.[5] Air pollution is one element that could affect how severe and when RA develops.[6] Research has shown that exposure to air pollutants such as nitric oxide (NO), particulate matter with a diameter

Partially supported by the Kuwait Foundation for the Advancement of Sciences (KFAS) under the project titled "KFAS Project: CN22-14SL-1653." The American University of Kuwait funded the open-access publication fees for this research.

[1]Ahmad Alsaber, PhD, Afsah Abdullah, BSE: American University of Kuwait, Salmiya, Kuwait; [2]Adeeba Al-Herz, PhD: Al-Amiri Hospital, Kuwait City, Kuwait; [3]Maryam Al-Mazedi, PhD: Public Authority for Applied Education and Training, Shuwaikh, Kuwait; [4]Iyad Abu Doush, PhD: American University of Kuwait, Salmiya, Kuwait, and Yarmouk University, Irbid, Jordan; [5]Dhary Alkandary, PhD: Kuwait University, Al-Shadadiya, Kuwait; [6]Ahmad T. Al-Sultan, PhD: Kuwait University, Jabriya, Kuwait; [7]Parul Setiya, PhD: University of Agriculture and Technology, Udham Singh Nagar, India; [8]Jiazhu Pan, PhD: University of Strathclyde, Glasgow, 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.25583).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25583.

Address correspondence via email to Ahmad Alsaber, PhD, at aalsaber@auk.edu.kw.

Submitted for publication October 15, 2024; accepted in revised form May 28, 2025.

SIGNIFICANCE & INNOVATIONS

- The significance of this study lies in its novel exploration of the relationship between air quality and rheumatoid arthritis (RA) disease activity in Kuwait. By leveraging data from air quality monitoring stations and the Kuwait Registry for Rheumatic Diseases, the study provides critical insights into how specific air pollutants such as nitrogen dioxide and ozone influence the Disease Activity Score in patients with RA.
- The innovative use of multivariate time-series models, particularly the Vector Error Correction Model, enables the examination of both the short- and long-term effects of air pollution on RA disease activity.
- This research fills a significant gap in understanding the environmental triggers of RA, offering valuable public health implications for treating patients with RA in regions with high pollution levels.
- These findings highlight the importance of regulating air pollution to improve disease outcomes and provide a foundation for future research and public health interventions aimed at mitigating the impact of environmental factors on RA.

of 2.5 micrometers or less ($PM_{2.5}$), and nitrogen dioxide (NO_2) is positively associated with an elevated risk of RA.[6] Additionally, studies have shown a possible link between the elevated risk of RA and certain air pollutants connected to transportation and home heating sources.[7] Additionally, cohort studies that included the general public have shown that air pollution might exacerbate autoimmune disorders, such as RA.[8] Several epidemiologic studies have examined the association between air pollution and RA. These include the Nurses' Health Study in the United States, British Columbian Study in Canada, and Swedish Epidemiological Investigation of Rheumatoid Arthritis.[9] There may be a connection between RA and air pollution, and this research has helped us understand it better.

One common way to measure air pollution is via the Air Quality Index (AQI). The list of dangerous pollutants it contains includes ozone (O_3), sulfur dioxide (SO_2), NO_2, PM_{10}, carbon monoxide (CO), and NO. It is well recognized that these pollutants have several harmful effects on human health, especially on the cardiovascular and pulmonary systems. According to recent studies, autoimmune diseases such as RA may potentially be affected by air pollution. Increased systemic inflammation has been linked to prolonged exposure to high levels of air pollutants, which may initiate or exacerbate autoimmune reactions, such as those observed in RA. However, how air pollution affects RA remains unclear.

Significant morbidity, early death, and symmetrical inflammatory polyarthritis are the outcomes of RA, a chronic multisystem disease of unclear origin.[10] In Kuwait, the frequency of RA is approximately 1%, similar to that reported in other nations.[11] However, Kuwait and other countries in the Middle East have limited descriptive data on patients with RA.[12]

According to several studies, exposure to air pollution increases the risk of developing RA.[13] Additionally, epidemiologic data indicate a robust correlation between exposure to environmental pollution components, such as noise, dioxin, cigarette smoke, and traffic-related air pollution, and the risk of RA.[14–16] Air pollutants are believed to aggravate the disease and affect RA symptoms.

This study aimed to assess the impact of key air pollutants—NO_2, SO_2, O_3, NO, PM_{10}, and CO—on Disease Activity Score in 28 joints (DAS28) scores among patients with RA in Kuwait, using both short- and long-term perspectives. By integrating data from the Kuwait Registry for Rheumatic Diseases (KRRD) with air quality data from six monitoring stations, this study provides a comprehensive analysis of environmental influences on RA. As the first study of its kind in Kuwait, it applied advanced multivariate time-series methods, including cointegration analysis and Vector Error Correction Model (VECM), offering novel insights into how air pollution affects RA disease activity over time.

MATERIALS AND METHODS

AQI measurement. The AQI changes from one nation to another, and from one pollutant to another. The recommended AQI incorporates five pollutants (CO, NO_2, SO_2, PM_{10}, and O_3) for which short-term (24-hour average period limit) National Ambient Air Quality Standards have been set. To account for the known environmental concentrations, applicable limits, and expected health impacts, we divided each of these pollutants into its own subindex. In this study, the AQI for the State of Kuwait was used, following the definition provided by Al-Shayji et al.[17]

Air quality data set. This study used air quality data from six long-standing monitoring stations established by the Kuwait Environmental Public Authority (K-EPA), each strategically located across different regions of the country, to ensure comprehensive environmental coverage. The selected stations included Al-Fahaheel, Al-Mansouriya, Al-Jahra, Al-Rumaithiya, Al-Ahmadi, and Al-Salam (see Supplementary Figure 1). These permanent stations, part of a larger national network initiated in 1984, were equipped to automatically record and store hourly data on various pollutants, which were aggregated into daily averages for the purpose of this research. Data were sourced from the Environmental Monitoring Information System of Kuwait and spanned from January 2012 to December 2023. Positioned mainly in residential areas, these stations monitor both anthropogenic and natural sources of pollution, thereby providing a robust foundation for assessing long-term environmental exposure and its potential health impacts.

RA data collection. The KRRD is a comprehensive national registry that collects clinical data on patients with rheumatic diseases across Kuwait. KRRD supplied the data on patients with RA. Patients diagnosed with RA who meet the specific requirements established by the American College of Rheumatology have their medical information included in the KRRD.[18] This study investigated data from 4,459 follow-up visits and 1,526 patients with RA who were enrolled in the KRRD between January 1, 2012, and December 30, 2023.

DAS28 calculation formula. DAS28 is a clinical measure used to assess RA disease activity by evaluating 28 joints for tenderness and swelling (see Supplementary Figure 2), along with markers of inflammation and patient-reported health. It combines four components as markers of inflammation: tender joint count, swollen joint count, erythrocyte sedimentation rate (ESR), and C-Reactive Protein (CRP), and the Patient Global Health Assessment rated on a visual analog scale. The DAS28 score was calculated using established formulas that vary depending on whether ESR or CRP is used. It serves as a standard tool in both clinical practice and research for tracking RA progression and assessing treatment responses.

Linkage between AQI and RA data. To examine the relationship between air pollution and RA disease activity, AQI data were merged with patient records from the KRRD based on date and residential location, aligning each patient with the nearest monitoring station (Supplementary Figure 3). This spatial-temporal matching is essential for accurately assessing environmental exposure. The integration and analysis of these large data sets were conducted using R and Python, with R Studio (version 1.1.463) used for statistical analysis, and Python for data integration and geospatial processing. This approach ensured a precise link between air quality and RA disease activity data.

Time series and cointegration procedures. To examine the dynamic relationship between the air pollution components and RA disease activity, a structured time-series econometric procedure was implemented. This multistep process includes unit root testing, lag length selection, cointegration analysis, and eventual modeling using the VECM. The analysis began by assessing the stationarity of each variable through standard unit root tests, such as the Augmented Dickey-Fuller (ADF), modified ADF–generalized least squares (GLS), and Kwiatkowski–Phillips–Schmidt–Shin (KPSS) tests. These tests are essential to identify whether the mean and variance of each time series remain constant over time. The ADF and ADF-GLS tests for the presence of a unit root,[19] whereas KPSS serves as a complementary approach by testing the null hypothesis of stationarity.[20] The integration of these tests increases robustness, particularly in data sets susceptible to structural changes or external shocks, which are common in environmental and health-related data.[21]

Following the confirmation of stationarity properties, a lag length selection was conducted to determine the appropriate number of lag periods to be included in the models. This decision significantly influences both the predictive accuracy and interpretation of time-dependent relationships. Established model selection criteria, such as the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Hannan-Quinn Criterion (HQC), were employed.[22] Although AIC is known to accommodate more complex models, BIC favors parsimony and HQC provides a balanced approach.[23] Given the environmental context in which pollutant effects may be delayed, choosing an optimal lag structure is particularly important to avoid misinterpretation.[24]

The next phase involves Johansen cointegration analysis, which evaluates whether a long-term equilibrium relationship exists between nonstationary variables that may otherwise fluctuate independently in the short term. This method identifies cointegrating vectors that combine pollutant concentrations and RA disease activity scores (DAS28). The presence of cointegration indicates that although short-term deviations occur, these variables are bound by a stable long-term path.[25,26] The main cointegration equation is expressed as $\beta' X_t = u_t$, in which "X_t" is the vector of nonstationary variables (eg, DAS28 and AQI components), "β'" is the cointegrating vector, and "u_t" is the stationary residual. Once cointegration has been established, a VECM is applied.[27] This model incorporates both long- and short-run dynamics by including an error correction term that quantifies how the system returns to equilibrium after a deviation. Through this integrated framework, this study captures how air pollution affects RA disease activity immediately and over extended periods, offering nuanced insights for both clinical and environmental health research.[28]

Imputing air pollution missing data. Before performing the time-series analysis, missing values in the air quality data set were imputed to ensure data completeness and prevent bias in evaluating the relationship between air pollution and RA disease activity. Missing data in environmental data sets often arise from equipment failures, sampling issues, and human error, which can compromise the reliability of statistical analyses.[29] To address this, several advanced machine-learning-based imputation methods have been employed, including k-nearest neighbor, predictive mean matching (PMM), random forest, and missForest, all of which are tailored to the specific characteristics of environmental data. This study builds on the previous work by Alsaber et al[29] (2021), which highlighted the importance of accurate data imputation in air quality research. The developed imputation framework, based on K-EPA data sets, enhances data integrity for pollutants such as NO_2, SO_2, CO, PM_{10}, and O_3, thereby supporting robust and precise environmental health research.[29]

Statistical procedures. Time-series graphs were generated for each pollutant to visualize trends and fluctuations over time using aggregated means and SDs from multiple monitoring stations. A univariate time-series analysis was used to assess the temporal behavior of each pollutant. For data from patients with RA, descriptive statistics summarized demographics, disease duration, and DAS28 scores, with medians, means, interquartile ranges, and SDs for continuous variables and frequencies for categorical data. Pearson's correlation was used to assess linear relationships between the DAS28 and AQI components. To control for confounding factors, regression models included individual-level factors such as age, sex, and comorbidity status.

Ethical consideration and data availability statement. This study was conducted in compliance with the ethical guidelines and principles outlined by relevant institutional review boards. Ethical approval for the study was issued by the Ministry of Health in Kuwait under the ethical reference number (2022/2194), which covered all major hospitals in Kuwait. Additionally, Kuwait University approved the establishment of KRRD under reference number (2016/477). The data supporting this study are available from the corresponding author upon reasonable request, due to privacy concerns.

RESULTS

Patients characteristics. Table 1 presents the demographic characteristics of the study sample, which consisted of 1,526 patients. The data indicated that female participants constituted the majority of the sample (77.7%), with a male proportion of 22.3%. Of the cohort, 53.7% of participants were Kuwaiti citizens, whereas the remaining 46.3% were of other nationalities. Participants' ages ranged from 17 to 95 years, with an average age of 55.3 years (SD = 13.1 years). Patients' disease durations ranged greatly from those with a recently diagnosed illness to those with a 58-year disease history, with a mean of 11.4 years (SD = 7.8 years). The body mass index ranges from 10.5 to 56.5, with an average of 29.2 (SD = 6.0). The majority of patients (88.3%) did not smoke, and only 11.4% of them now reported smoking, with the majority of patients (75.9%) having no family history of RA. Of the patients, 5.7% had deformities and 4.3% had rheumatoid nodules, which were unusual. A total of 68.6% of the patients who tested positive for anti–cyclic citrullinated peptide antibodies, 42.2% of patients tested positive for antinuclear antibodies, and 76.9% of patients tested positive for rheumatoid factor. A total of 23.4% of patients report having Sjögren's International Collaborative Clinical Alliance symptoms in terms of clinical symptoms. The cohort's average height was 159.6 cm (SD = 10.9 cm), and the average weight was 75.5 kg (SD = 23.6 kg). The DAS28 indicates a range of activity of the disease from remission to high levels of activity, with a mean of 2.8 (SD = 1.3).

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

Characteristic	Overall (*N* = 1,526)
Sex, n (%)	
Female	1,185 (77.7)
Male	341 (22.3)
Nationality, n (%)	
Kuwait	820 (53.7)
Other nationalities	706 (46.3)
Age	
Mean (SD)	55.3 (13.1)
Range	17.0–95.0
Disease duration	
Mean (SD)	11.4 (7.8)
Range	0.0–58.0
BMI	
Mean (SD)	29.2 (6.0)
Range	10.5–56.5
Family history, n (%)	
Negative	859 (75.9)
Positive	273 (24.1)
Smoking, n (%)	
Ex-smoker	4 (0.3)
No	1,015 (88.3)
Yes	131 (11.4)
Rheumatoid nodules, n (%)	
No	1,255 (95.7)
Yes	57 (4.3)
RF, n (%)	
Negative	316 (23.1)
Positive	1,052 (76.9)
Deformities, n (%)	
No	1,205 (94.3)
Yes	73 (5.7)
SICCA symptoms, n (%)	
No	1,010 (76.6)
Yes	308 (23.4)
Anti-CCP, n (%)	
Negative	366 (31.4)
Positive	799 (68.6)
ANA, n (%)	
Negative	606 (57.8)
Positive	443 (42.2)
Weight	
Mean (SD)	75.5 (23.6)
Range	0.0–660.0
Height	
Mean (SD)	159.6 (10.9)
Range	4.9–254.0
DAS28	
Mean (SD)	2.8 (1.3)
Range	0.0–7.8

* ANA, antinuclear antibody; anti-CCP, anti–cyclic citrullinated peptide antibody; BMI, body mass index; DAS28, Disease Activity Score in 28 joints; RF, rheumatoid factor; SICCA, Sjögren's International Collaborative Clinical Alliance.

Dealing with missing data using a mixed approach of PMM and missForest. The imputation of missing data preserves data set integrity by substituting estimated values for missing or incomplete data. Frequently used techniques to guarantee that the imputed values closely match the original data distribution include missForest and PMM. This procedure reduces any biases

Table 2. Descriptives statistics for the AQI of pollutant variables and Disease Activity Score in 28 joints*

	Mean	Median	IQR	SD	Minimum	Maximum
CO_AQI	21.247	19.6	8.82	10.02	5.94	276.74
NO AQI	45.840	38.6	36.40	32.41	5.01	400.00
NO_2_AQI	65.239	62.2	37.32	23.96	11.57	128.07
O_3 AQI	39.792	35.2	14.82	27.48	8.34	319.63
PM_{10} AQI	201.313	145.5	107.65	137.12	40.61	599.94
SO_2_AQI	35.819	29.3	22.84	23.34	6.52	209.18

* AQI, Air Quality Index; CO, carbon monoxide; IQR, interquartile range; NO, nitric oxide; NO_2, nitrogen dioxide; O_3, Ozone; PM_{10}, particulate matter with a diameter of 10 micrometers or less; SO_2, sulfur dioxide.

caused by missing data and helps preserve the accuracy of the statistical assessments. This approach offers a stronger basis for examining air quality trends and helps solve any biases resulting from missing data. The visualization likely depicts a comparison of the pre- and postimputation distributions, stressing how the technique reduces the distortion of the features of the original data (see Supplementary Figure 4).

Kuwait air quality assessment from 2012 to 2023. An analysis of Kuwait's AQI data from 2012 to 2023 revealed notable pollution patterns, particularly concerning PM_{10}, which had the highest average AQI value (201.313), indicating substantial particulate matter in the air and raising public health concerns. Most pollutants had median values lower than their means, suggesting the presence of extreme pollution events that skew the data, especially for PM_{10}, with a peak AQI of 599.94. NO_2, combined with NO, showed moderate pollution levels, whereas CO had relatively low concentrations. SO_2 and O_3 are also present in manageable amounts, although they still require attention. The variability in pollution levels was significant for NO and PM_{10}, as indicated by their high SDs (see Table 2). Furthermore, the relationship between AQI and RA disease activity (DAS28) was positively correlated with NO, NO_2, and CO levels, implying a potential link to worsening health outcomes. Conversely, PM_{10} and O_3 were negatively correlated with DAS28, suggesting more complex health impacts. These findings highlight the need for ongoing environmental monitoring and policy measures to reduce pollution and protect public health (Table 3).

Unit root testing and stationary testing. Unit root testing was conducted on six air pollutant variables—CO, NO, NO_2, O_3, PM_{10}, and SO_2—using the ADF, ADF-GLS, and KPSS methods to assess their stationarity. The results indicated that all variables were nonstationary at their original levels but became stationary after the first differencing, as evidenced by significant *P* values and test statistics. This pattern suggests the presence of a unit root in the level series, which is a common characteristic of environmental time-series data. Consistency across all three tests validates the use of cointegration techniques in subsequent analyses to explore the long-term equilibrium relationships among these variables. Supplementary Table 1 presents the results of the unit root tests (ADF, ADF-GLS, and KPSS) applied to all air pollutant variables, both at the level and first difference, with and without trends. Across all tests, the variables were found to be nonstationary at level but became stationary after the first differencing, as indicated by the statistically significant ADF and ADF-GLS results and nonsignificant KPSS statistics after differencing. This justifies the use of cointegration analysis and supports the appropriateness of the VECM.

Lag selection criteria. The lag-order selection for the model, incorporating a constant, was evaluated using a range of information criteria, including log-likelihood, likelihood ratio *P* values, AIC, HQC, and BIC, across lag lengths from 1 to 20. Among these, the HQC was prioritized because of its balance between model fit and complexity. Unlike BIC, which heavily penalizes additional parameters, or AIC, which may overfit by favoring longer lag structures, HQC offers a moderate trade-off

Table 3. Pearson correlation analysis between AQI of pollutant variables and DAS28*

	CO	NO	NO_2	O_3	PM_{10}	SO_2	DAS28
CO	–	–	–	–	–	–	–
NO	0.069[a]	–	–	–	–	–	–
NO_2	0.104[a]	0.324[a]	–	–	–	–	–
O_3	−0.062[a]	−0.084[a]	−0.050[b]	–	–	–	–
PM_{10}	−0.025	−0.003	−0.044[b]	0.056[a]	–	–	–
SO_2	−0.082	0.053[a]	0.178[a]	0.227[a]	0.052[a]	–	–
DAS28	0.038	0.032[c]	0.040[b]	−0.061[a]	−0.046[b]	−0.055[a]	–

* AQI, Air Quality Index; CO, carbon monoxide; DAS28, Disease Activity Score in 28 joints; NO, nitric oxide; NO_2, nitrogen dioxide; O_3, ozone; PM_{10}, particulate matter with a diameter of 10 micrometers or less; SO_2, sulfur dioxide.
[a] $P < 0.001$.
[b] $P < 0.05$.
[c] $P < 0.01$.

Table 4. Cointegration regression examining the long-term association between DAS28 and ambient air pollutants, adjusted for age, sex, and comorbidities*

	Coefficient (R^2 = 0.010)	SE	t-ratio (adjusted R^2 = 0.009)	*P* value
const	0.552	0.084	6.535	<0.001[a]
CO_AQI	0.003	0.002	1.637	0.102
NO_AQI	0.001	0.001	1.067	0.286
NO_2_AQI	0.002	0.001	2.252	0.024[b]
O_3_AQI	−0.002	0.001	−2.764	0.006[a]
PM_{10}_AQI	−0.000	0.000	−2.584	0.010[a]
SO_2_AQI	−0.003	0.001	−3.041	0.002[a]

* AQI, Air Quality Index; CO, carbon monoxide; const, constant; DAS28, Disease Activity Score in 28 joints; NO, nitric oxide; NO_2, nitrogen dioxide; O_3, ozone; PM_{10}, particulate matter with a diameter of 10 micrometers or less; SO_2, sulfur dioxide.
[a] $P < 0.001$.
[b] $P < 0.05$.

ideal for medium-sized data sets. Selecting the lag based on HQC ensures that the model remains both flexible and robust, thus supporting sound inference and reliable forecasting. Supplementary Table 2 shows the lag-order selection criteria, in which the HQC identified lag 7 as optimal, balancing model complexity and fit. This selection guided the construction of the VECM to examine the short- and long-term dynamics between the air pollution components and RA disease activity.

Multivariate time series: Johansen cointegration test and VECM. The cointegration test results, summarized in Supplementary Table 3, revealed multiple long-term relationships between air pollutants and RA disease activity indicators. Using both trace and maximum eigenvalue (Lmax) statistics, all ranks showed highly significant outcomes ($P < 0.001$), rejecting the null hypothesis of no cointegration and indicating that the variables tend to move together over the long term despite short-term fluctuations. The regression results shown in Table 4 further confirm these associations. The constant term (β = 0.5519, $P < 0.001$) established a significant baseline, whereas several AQI variables—PM_{10}_AQI ($P = 0.0098$), NO_2_AQI ($P = 0.0024$), O_3_AQI ($P = 0.0057$), and SO_2_AQI ($P = 0.0024$)—were significantly linked with DAS28. NO_2_AQI showed a positive effect, whereas PM_{10}_AQI, O_3_AQI, and SO_2_AQI had negative coefficients. In contrast, the NO_AQI and CO_AQI were not statistically significant. The negative relationship between O_3 and DAS28 may reflect seasonal patterns, in which lower ozone levels in colder months coincide with increased RA activity.[30] Similarly, the inverse correlations of PM_{10} and SO_2 with DAS28 may have been influenced by seasonal human activity and exposure patterns. Lag length selection was based on AIC, BIC, and HQC, with HQC chosen for its balanced approach. The Johansen cointegration test confirmed persistent long-term links, especially for NO_2 and O_3, whereas the VECM results demonstrated a gradual correction of deviations from equilibrium and significant short-term effects for specific pollutants.

The VECM analysis (2012–2023), as presented in Table 5, indicates that the DAS28 is significantly affected by its own lagged values and specific AQI components. Notably, the NO_2_AQI and O_3_AQI exhibited consistent and statistically significant short-term effects across several lags, highlighting their strong influence on RA disease activity. In contrast, CO_AQI, SO_2_AQI, and NO_AQI showed limited or isolated significance, whereas PM_{10}_AQI showed no significant effect at any lag. The error correction term was highly significant and negative (β = –0.5425, $P < 0.01$), confirming that deviations from long-term equilibrium were corrected over time. The models' $R^2 = 0.4554$ and adjusted $R^2 = 0.4494$ suggest a moderate level of explanatory power. As illustrated in Figure 1 and detailed in Supplementary Figure 5, the forecast results show that most of the variance in DAS28 is explained by its own history, with meaningful contributions from AQI components—particularly NO_2_AQI and O_3_AQI—supporting their role in influencing RA activity levels.

DISCUSSION

Recent research has suggested a possible relationship between air pollution and air pollution. Elevated RA disease activity and flares have been linked to pollution including[31–33] NO_2, SO_2, PM_{10}, and $PM_{2.5}$. According to certain studies, air pollution may be involved in the development of juvenile idiopathic arthritis and systemic autoimmune rheumatic disorders.[34] Potentially exacerbating and causing RA, air pollution may increase processes associated with inflammation and autoantibody formation.[35] The results are not completely consistent across studies; however, some studies have demonstrated no obvious correlation between RA risk and air pollution.[7] Further research is needed on the complicated interplay of environmental and genetic elements in RA pathogenesis, including air pollution.[9] Pollutants such as $PM_{2.5}$, PM_{10}, and NO_2 might cause systemic inflammation that can aggravate RA symptoms and raise flare-up severity and frequency. Kuwait presents a significant public health problem owing to its unique industrial and environmental landscape, necessitating stringent air quality management measures and public health interventions.

Table 5. Vector Error Correction Model estimating short- and long-term associations between ambient air pollutants and DAS28, adjusted for age, sex, and comorbidities*

	Coefficient (R^2 = 0.455)	SE	Z score (adjusted R^2 = 0.449)	*P* value
const	0.2547	0.1065	2.3910	0.017[a]
d_DAS28_1	−0.3152	0.0305	−10.34	<0.001[b]
d_DAS28_2	−0.2703	0.0282	−9.602	<0.001[b]
d_DAS28_3	−0.2916	0.0252	−11.57	<0.001[b]
d_DAS28_4	−0.3325	0.0223	−14.91	<0.001[b]
d_DAS28_5	−0.2888	0.0193	−14.99	<0.001[b]
d_DAS28_6	−0.2161	0.0148	−14.56	<0.001[b]
d_CO_AQI_1	−0.0013	0.0033	−0.4282	0.669
d_CO_AQI_2	−0.0028	0.0032	−0.8775	0.380
d_CO_AQI_3	−0.0045	0.0030	−1.485	0.138
d_CO_AQI_4	−0.0057	0.0028	−2.028	0.043[a]
d_CO_AQI_5	−0.0007	0.0026	−0.3012	0.763
d_CO_AQI_6	−0.0020	0.0022	−0.9502	0.342
d_NO_2_AQI_1	−0.0032	0.0011	−2.865	0.004[b]
d_NO_2_AQI_2	−0.0048	0.0012	−3.943	<0.001[b]
d_NO_2_AQI_3	−0.0039	0.0013	−3.111	0.002[b]
d_NO_2_AQI_4	−0.0028	0.0013	−2.256	0.024[a]
d_NO_2_AQI_5	−0.0018	0.0012	−1.566	0.117
d_NO_2_AQI_6	0.0001	0.0010	0.1271	0.899
d_O_3_AQI_1	0.0034	0.0012	2.7880	0.005[b]
d_O_3_AQI_2	0.0031	0.0012	2.5970	0.009[b]
d_O_3_AQI_3	0.0015	0.0011	1.3360	0.182
d_O_3_AQI_4	0.0027	0.0011	2.4890	0.013[a]
d_O_3_AQI_5	0.0021	0.0010	2.1170	0.034[a]
d_O_3_AQI_6	0.0011	0.0008	1.3100	0.190
d_PM_{10}_AQI_1	0.0001	0.0003	0.5221	0.602
d_PM_{10}_AQI_2	0.0001	0.0002	0.5490	0.583
d_PM_{10}_AQI_3	0.0001	0.0002	0.5031	0.615
d_PM_{10}_AQI_4	0.0001	0.0002	0.2599	0.795
d_PM_{10}_AQI_5	0.0001	0.0002	0.3444	0.731
d_PM_{10}_AQI_6	0.0001	0.0001	0.5898	0.555
d_SO_2_AQI_1	0.0015	0.0015	0.9825	0.326
d_SO_2_AQI_2	0.0004	0.0015	0.2428	0.808
d_SO_2_AQI_3	0.0001	0.0014	0.0707	0.944
d_SO_2_AQI_4	0.0006	0.0013	0.4527	0.651
d_SO_2_AQI_5	−0.0017	0.0012	−1.484	0.138
d_SO_2_AQI_6	−0.0022	0.0010	−2.368	0.018[a]
EC1	−0.5425	0.0319	−17.03	<0.001[b]
EC2	0.0006	0.0031	0.1885	0.851
EC3	0.0016	0.0012	1.3230	0.186
EC4	0.0034	0.0007	5.0790	<0.001[b]
EC5	−0.0026	0.0012	−2.314	0.021[a]
EC6	−0.0003	0.0003	−1.361	0.174

* AQI, Air Quality Index; CO, carbon monoxide; const, constant; DAS28, Disease Activity Score in 28 joints; EC, error correction term; NO, nitric oxide; NO_2, nitrogen dioxide; O_3, ozone; PM_{10}, particulate matter with a diameter of 10 micrometers or less; SO_2, sulfur dioxide.
[a] $P < 0.05$.
[b] $P < 0.001$.

These results indicate a complicated relationship between RA and air pollution. The findings show that DAS28, in both short- and long-term exposures, is significantly influenced by NO_2 and O_3. Studies using Granger causality have shown that O_3 and NO_2 are very significant indicators of both the DAS28 and the Clinical Disease Activity Index (CDAI), suggesting that increasing concentrations of these pollutants may lead to increased disease activity. Likewise, changes in CDAI and DAS28 were also shown to be induced by CO, SO_2, and PM_{10}, but the effects of these pollutants varied less and were not as stable across time.

Short-term contact with air pollutants, particularly O_3 and NO_2, has been linked to increased RA disease activity, as evaluated by CDAI and DAS28 scores.[31,36] On the other hand, a number of investigations have shown conflicting or minimal effects of brief exposure to air pollution on RA activity.[37] There is evidence linking a higher frequency of RA to long-term exposure to several air pollutants including[31] O_3, CO, SO_2, and NO_2.

This study investigated the long-term relationship between DAS28 scores and several air pollutants—CO, SO_2, PM_{10}, NO_2, and O_3—among patients with RA in Kuwait using cointegration analysis. This approach allowed for the assessment of whether

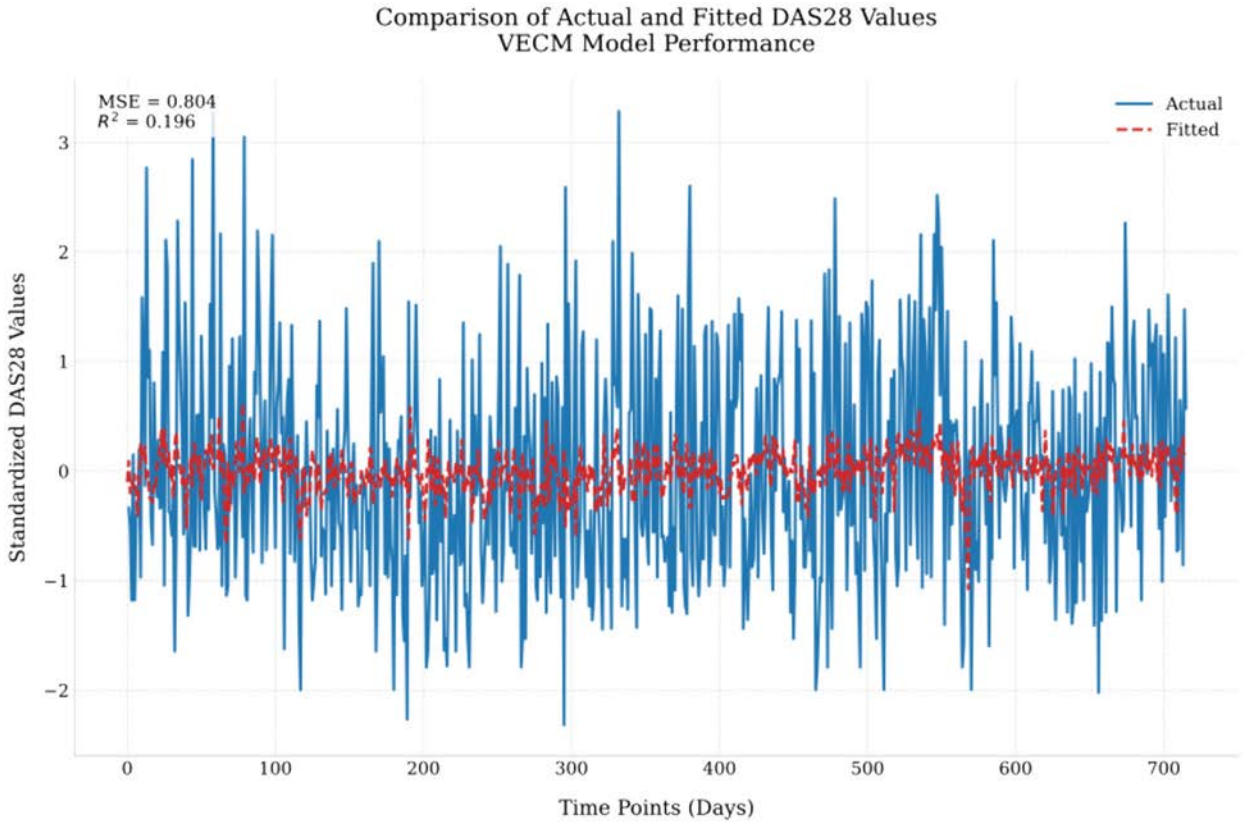


Figure 1. Comparison of actual and fitted standardized DAS28 values over time using a VECM model performance. DAS28, Disease Activity Score in 28 joints; MSE, Mean Squared Error; VECM, Vector Error Correction Model.

DAS28 scores and pollutant levels moved together over time, even amid short-term fluctuations. The presence of cointegration suggests a stable equilibrium relationship, implying that changes in pollutant levels can lead to adjustments in disease activity in RA. These findings highlight the potential of long-term environmental exposure to contribute to chronic inflammation and exacerbate RA symptoms.

Supporting evidence from international studies further emphasizes this link. For example, a longitudinal study in Italy associated higher pollutant concentrations with increased RA flare-ups,[33] and a Taiwanese cohort found strong associations between RA risk and exposure[38] to SO_2, NO_2, PM_{10}, CO, and O_3. Our cointegration analysis identified O_3 and NO_2 as significantly related to long-term increases in the DAS28 scores. This aligns with the theory that prolonged exposure to these pollutants causes systemic inflammation. Nonetheless, some research has shown inconsistent findings, especially concerning[34,37] PM_{10} and O_3, suggesting the need for continued investigation into environmental influences on RA.

The VECM is a useful tool for investigating the long-term equilibrium connections between variables, as well as short-term dynamics. It combines cointegration analysis and error correction methods to identify both short- and long-term changes.[39] In terms of explanatory capacity and predictive capability, it often performs better than other models such as Vector Autoregression.[40] By adding an error correction term, we can measure how quickly the variables return to equilibrium after a shock, allowing us to distinguish between short-term shifts and long-term trends.[41]

The analysis revealed a complex relationship between RA disease activity and air quality indicators, with O_3 and NO_2 showing the strongest influence on the DAS28 and CDAI scores. Granger causality within the VECM framework confirmed both short- and long-term links between the higher levels of these pollutants and increased RA activity. This may be due to the inflammatory effects of NO_2 and the role of O_3 in oxidative stress. The negative and significant error correction term indicated that any short-term deviation in disease activity caused by pollution eventually returned to long-term equilibrium. Additionally, impulse response functions showed that shocks in NO_2 and O_3 levels led to sustained increases in DAS28 scores, whereas the effects of CO, PM_{10}, and SO_2 were more variable and often short-lived.

Impulse response functions and VECMs have identified the short- and long-term effects of air pollution on RA activity.[36] These effects may change depending on the particular pollutant, the length of exposure, and patient characteristics. Moreover, treatment with ongoing disease-modifying antirheumatic medications may influence the relationship between air pollution and RA

disease activity.[37] These findings suggest that air pollutants could be a risk factor for RA, necessitating further investigation and careful application in clinical settings.

The analytical strategy employed in this study enabled the investigation of both the long- and short-term associations between air pollution and RA activity. Unit root testing ensured that stationarity requirements were met for the cointegration analysis, which identified enduring relationships between DAS28 and specific pollutants. The VECM approach accounts for short-run variations while incorporating an error correction mechanism, offering insight into how fluctuations in air quality may influence disease progression over time. This dual framework contributes to a comprehensive understanding of the environmental determinants of RA.

This study faced several limitations that need to be acknowledged. First, the observational nature of the research limits its capacity to conclusively prove causality. Although the Granger causality and VECM tests provide some insight into possible causal relationships, they cannot completely rule out the impact of unobserved confounding factors. Furthermore, fixed monitoring sites provided data on air pollution, which would not entirely reflect individual-level exposure fluctuations given that Kuwait has varied microclimates.

This study is based on aggregated data from air-monitoring stations and population-level RA activity measures and is thus subject to the potential for ecological fallacy. The observed associations represent group-level trends and may not necessarily reflect individual-level causal relationships. Although the analysis provides valuable insights into the potential environmental influences on RA activity at the population scale, caution is warranted in interpreting these findings at the individual level.

Future research should further investigate the complex relationship between air pollution and RA, with a particular focus on long-term exposure to key pollutants, such as SO_2 and NO_2, which have been associated with increased RA risk and heightened disease activity. Special attention should be paid to gene-environment interactions, as emerging evidence suggests that individuals with a strong genetic predisposition may experience nearly double the incidence of RA when exposed to elevated air pollutant levels. Additionally, future studies should adopt longitudinal designs with larger and more diverse cohorts, integrate advanced machine learning algorithms for predictive modeling, and incorporate biologic pathway analyses to elucidate the mechanistic links between environmental exposure and RA pathogenesis.

This pioneering study was the first to explore the relationship between air pollution and RA activity in Kuwait, offering important epidemiologic insights from a region with distinct environmental conditions. Using VECM to capture both short- and long-term effects, the study found significant associations between DAS28 scores and pollutants, particularly NO_2 and O_3, which were strong predictors of increased disease activity. These findings emphasize the impact of air quality on RA progression and highlight the need for targeted public health interventions and clinical strategies to reduce environmental risks and improve patient outcomes.

ACKNOWLEDGMENTS

The authors extend their heartfelt appreciation to the KRRD for providing the RA data used in this study. Additionally, we acknowledge the K-EPA and its Air-Monitoring Stations, operating under the Environmental Monitoring Information System of Kuwait and the Environmental Data Management Department, for their invaluable contributions.

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 Alsaber 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. Tobón GJ, Youinou P, Saraux A. The environment, geo-epidemiology, and autoimmune disease: rheumatoid arthritis. Autoimmun Rev 2010; 9(5):A288–A292.
2. Bernatsky S, Smargiassi A, Barnabe C, et al. Fine particulate air pollution and systemic autoimmune rheumatic disease in two Canadian provinces. Environ Res 2016;146:85–91.
3. McEachan RRC, Rashid R, Santorelli G, et al. Study protocol. Evaluating the life-course health impact of a city-wide system approach to improve air quality in Bradford, UK: a quasi-experimental study with implementation and process evaluation. Environ Health 2022; 21(1):122.
4. Ward F, Lowther-Payne HJ, Halliday EC, et al. Engaging communities in addressing air quality: a scoping review. Environ Health 2022; 21(1):89.
5. Viatte S, Plant D, Raychaudhuri S. Genetics and epigenetics of rheumatoid arthritis. Nat Rev Rheumatol 2013;9(3):141–153.
6. Liu B, Sun G, Liu Y, et al. Observational studies: ambient air pollution and hospitalization for RA-ILD in a heavily polluted city in China. Medicine (Baltimore) 2022;101(19):e29309.
7. Hart JE, Källberg H, Laden F, et al. Ambient air pollution exposures and risk of rheumatoid arthritis: results from the Swedish EIRA case-control study. Ann Rheum Dis 2013;72(6):888–894.
8. Adami G, Pontalti M, Cattani G, et al. Association between long-term exposure to air pollution and immune-mediated diseases: a population-based cohort study. RMD Open 2022;8(1):e002055.
9. Essouma M, Noubiap JJN. Is air pollution a risk factor for rheumatoid arthritis? J Inflamm (Lond) 2015;12:48.
10. Cross M, Smith E, Hoy D, et al. The global burden of rheumatoid arthritis: estimates from the global burden of disease 2010 study. Ann Rheum Dis 2014;73(7):1316–1322.
11. Al-Awadhi AM, Olusi SO, Al-Saeid K, et al. Incidence of musculoskeletal pain in adult Kuwaitis using the validated Arabic version of the WHO-ILAR COPCORD Core Questionnaire. Ann Saudi Med 2005; 25(6):459–462.

12. Cheng T, Zhang G. Worldwide research productivity in the field of rheumatology from 1996 to 2010: a bibliometric analysis. Rheumatology (Oxford) 2013;52(9):1630–1634.
13. Chang KH, Hsu CC, Muo CH, et al. Air pollution exposure increases the risk of rheumatoid arthritis: a longitudinal and nationwide study. Environ Int 2016;94:495–499.
14. De Roos AJ, Koehoorn M, Tamburic L, et al. Proximity to traffic, ambient air pollution, and community noise in relation to incident rheumatoid arthritis. Environ Health Perspect 2014;122(10):1075–1080.
15. Hart JE, Laden F, Puett RC, et al. Exposure to traffic pollution and increased risk of rheumatoid arthritis. Environ Health Perspect 2009; 117(7):1065–1069.
16. Cowan KN, Wyatt LH, Luben TJ, et al. Effect measure modification of the association between short-term exposures to $PM_{2.5}$ and hospitalizations by longs-term $PM_{2.5}$ exposure among a cohort of people with chronic obstructive pulmonary disease (COPD) in North Carolina, 2002-2015. Environ Health 2023;22(1):49.
17. Al-Shayji K, Lababidi H, Al-Rushoud D, et al. Development of a fuzzy air quality performance indicator. Kuwait J Sci Eng 2008;35:101–126.
18. Aletaha D, Neogi T, Silman AJ, et al. Rheumatoid arthritis classification criteria: An American College of Rheumatology/European League against Rheumatism collaborative initiative. Arthritis & Rheumatism 2010;62(9):2569–2581. https://doi.org/10.1002/art.27584.
19. Hailu SM, Vural G. Testing the weak form market efficiency of Borsa Istanbul: an empirical evidence from Turkish banking sector stocks. Pressacademia 2020;7(3):236–249.
20. Shaari MS, Razak NAA, Basri BH. The effects of electricity consumption and economic growth on Carbon Dioxide emission. Int J Energy Econ Policy 2017;7:287–290.
21. Humbatova S. The impact of oil prices on state budget income and expenses: case of Azerbaijan. Int J Energy Econ Policy 2023;13(1): 189–212.
22. Ing CK, Wei CZ. Order selection for same-realization predictions in autoregressive processes. Ann Stat 2005;33(5):2423–2474.
23. Kamila K. General Hannan and Quinn Criterion for Common Time Series. 2021. https://doi.org/10.48550/arXiv.2101.04210.
24. Bardet JM, Kamila K, Kengne W. Consistent model selection criteria and goodness-of-fit test for common time series models. Electron J Stat 2020;14(1):2009–2052.
25. Kim Y, Radoias V. Severe air pollution exposure and long-term health outcomes. Int J Environ Res Public Health 2022;19(21):14019.
26. Stieb DM, Burnett RT, Smith-Doiron M, et al. A new multipollutant, no-threshold air quality health index based on short-term associations observed in daily time-series analyses. J Air Waste Manag Assoc 2008;58(3):435–450.
27. Johansen S. Likelihood-Based Inference in Cointegrated Vector Autoregressive Models. Oxford University Press; 1995.
28. Kermani M, Dowlati M, Ja'fari AJ, et al. Effect of air pollution on the emergency admissions of cardiovascular and respiratory patients, using the air quality model: a study in Tehran, 2005-2014. Health Emerg Disasters Quarterly 2016;1(3):137–146.
29. Alsaber AR, Pan J, Al-Hurban A. Handling complex missing data using random forest approach for an air quality monitoring dataset: a case study of Kuwait environmental data (2012 to 2018). Int J Environ Res Public Health 2021;18(3):1333.
30. Zhao D, Cheng J, Bao P, et al. Effect of ambient temperature on outpatient admission for osteoarthritis and rheumatoid arthritis in a subtropical Chinese city. BMC Public Health 2022;22(1):172.
31. Alsaber A, Pan J, Al-Herz A, et al; On Behalf Of The Krrd Group. Influence of ambient air pollution on rheumatoid arthritis disease activity score Index. Int J Environ Res Public Health 2020;17(2):416.
32. Alsaber AR. Estimating Air Pollution and Its Long and Short Run Causality with Chronic Diseases Using Vector Error Correction Model (VECM). PhD Thesis, University of Strathclyde; 2022.
33. Adami G, Viapiana O, Rossini M, et al. Association between environmental air pollution and rheumatoid arthritis flares. Rheumatology (Oxford) 2021;60(10):4591–4597.
34. Sun G, Hazlewood G, Bernatsky S, et al. Association between air pollution and the development of rheumatic disease: a systematic review. Int J Rheumatol 2016;2016:5356307.
35. Gumtorntip W, Kasitanon N, Louthrenoo W, et al. Potential roles of air pollutants on the induction and aggravation of rheumatoid arthritis: from cell to bedside studies. Environ Pollut 2023;334:122181.
36. Alsaber AR, Setiya P, Al-Sultan AT, et al. Exploring the impact of air pollution on COVID-19 admitted cases: evidence from Vector Error Correction Model (VECM) approach in explaining the relationship between air pollutants towards COVID-19 cases in Kuwait. Jpn J Stat Data Sci 2022;5(1):379–406.
37. Ingegnoli F, Ubiali T, Schioppo T, et al (2020) AB0229 short-term exposure to outdoor air pollutants and rheumatoid arthritis activity in metropolitan areas in the north of Italy. Ann Rheum Dis 2020;79(Suppl 1):1415.
38. Jung CR, Hsieh HY, Hwang BF. Air pollution as a potential determinant of rheumatoid arthritis: a population-based cohort study in Taiwan. Epidemiology 2017;28(Suppl 1):S54–S59.
39. Kong L, Chang Y. The impulse response of domestic and foreign interest rate in output, price, exchange rate model, a deconstructed derivation and economic calibration of Vector Error Correction Model. J Math Finance 2022;12:646–660.
40. Choi CS. A study on the long-term equilibrium relationship between the housing market and macroeconomic variables: focused on VECM and VAR Models. Korea Real Estate Society 2023;70:319–341.
41. Zardad A, Mohsin A, Zaman K. Estimating long-run equilibrium real exchange rates: short-lived shocks with long-lived impacts on Pakistan. Springerplus 2013;2(1):292.

Arthritis Care & Research
Vol. 77, No. 11, November 2025, pp 1350–1358
DOI 10.1002/acr.25580

American College of Rheumatology
Empowering Rheumatology Professionals

Agreement of Administrative Pharmacy Dispensing With Self-Reported Use of Oral Prednisone in US Veterans With Rheumatoid Arthritis

Beth I. Wallace,[1] Bryant R. England,[2] Joshua F. Baker,[3] Michael D. George,[4] Brian C. Sauer,[5] Jorge Rojas,[6] Punyasha Roul,[7] Katherine D. Wysham,[8] Hannah Brubeck,[8] Isaac Smith,[9] Liron Caplan,[10] Paul A. Monach,[11] Gail S. Kerr,[12] Gary Kunkel,[5] Tawnie Braaten,[5] Ted R. Mikuls,[2] and Grant W. Cannon[5]

Objective. Administrative claims are used to evaluate oral glucocorticoid use in rheumatoid arthritis (RA), despite limited evidence to support accuracy. We aimed to evaluate the performance of claims-based algorithms for glucocorticoid use compared to self-report in an RA population.

Methods. Participants with RA enrolled at seven Veterans Affairs Rheumatoid Arthritis (VARA) Registry sites were asked six questions as part of clinical care assessing current prednisone use and dose, recent use, "stockpiling," and receiving prednisone outside the Department of Veterans Affairs (VA). Algorithms using VA prescription claims operationalized current use (active prescription on date of self-report assessment), current dose (that prescription's mean dose), and recent use (active course overlapping the prior 30 or 90 days). We assessed performance characteristics and agreement, benchmarked on self-report.

Results. Of 284 participants, 13% reported current prednisone use and 20% reported 90-day use. Sensitivity, specificity, positive predictive value, and negative predictive value were 0.70, 0.98, 0.84, and 0.96, respectively, for current use and 0.71, 0.92, 0.72, and 0.92, respectively, for 90-day use. Cohen's κ was 0.68 for current use and 0.63 for 90-day use. Among participants reporting ≤5 mg/day, agreement for dose was high (weighted κ 0.67). One in four participants reported a stockpile, and one in four reported receiving prednisone from a non-VA provider.

Conclusion. Algorithms derived from VA claims detecting prednisone prescriptions have high validity compared to patient self-report. The modest sensitivity of these algorithms may reflect stockpiling and non-VA prescriptions. These findings form a basis for contextualizing real-world studies of glucocorticoid use in RA and improve clinical estimation of glucocorticoid use not captured in claims.

Dr Wallace's work was supported by the VA (grant IK2-CX-002430). Dr England's work was supported by the VA (grant IK2-CX-002203) and the Rheumatology Research Foundation. Dr Baker's work was supported by the VA (grants I01-RX-003644 and I01-RX-004770l). Dr Wysham's work was supported by VA Clinical Science Research & Development (grant ICX-002351). Dr Smith's work was supported by a Career Development Award from the Duke Center for Research to Advance Healthcare Equity (grant 3U54MD012530-05S2). Dr Mikuls' work was supported by the VA (grant BX-003635), the NIH (grant U54-GM-115458), and the Department of Defense (grant PR200793).

[1]Beth I. Wallace, MD, MSc: VA Ann Arbor Healthcare System and University of Michigan, Ann Arbor, Michigan; [2]Bryant R. England, MD, PhD, Ted R. Mikuls, MD, MSPH: VA Nebraska-Western Iowa Healthcare System and University of Nebraska Medical Center, Omaha, Nebraska; [3]Joshua F. Baker, MD, MSCE: Hospital of the University of Pennsylvania and VA Philadelphia Healthcare System, Philadelphia, Pennsylvania; [4]Michael D. George, MD, MSCE: Hospital of the University of Pennsylvania, Philadelphia, Pennsylvania; [5]Brian C. Sauer, PhD, Gary Kunkel, MD, Tawnie Braaten, MD, Grant W. Cannon, MD: VA Salt Lake City Healthcare System and University of Utah, Salt Lake City, Utah; [6]Jorge Rojas, MS: VA Salt Lake City Healthcare System, Salt Lake City, Utah, and VA Puget Sound Health Care System, Seattle, Washington; [7]Punyasha Roul, BA: University of Nebraska Medical Center, Omaha, Nebraska; [8]Katherine D. Wysham, MD, Hannah Brubeck, BA: VA Puget Sound Health Care System and University of Washington, Seattle; [9]Isaac Smith, MD: Duke University and VA Durham Health Care System, Durham, North Carolina; [10]Liron Caplan, MD, PhD: Rocky Mountain Regional VA Medical Center, Aurora, and University of Colorado, Denver, Colorado; [11]Paul A. Monach, MD, PhD: VA Boston Healthcare System and Harvard Medical School, Boston, Massachusetts; [12]Gail S. Kerr, MD: VA Washington DC Healthcare System, Georgetown-Howard University, and Howard University, Washington, DC.

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.25580).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25580.

Address correspondence via email to Beth I. Wallace, MD, MSc, at brennerb@umich.edu.

Submitted for publication February 24, 2025; accepted in revised form May 23, 2025.

SIGNIFICANCE & INNOVATIONS

- In a cohort of participants with rheumatoid arthritis (RA) enrolled in a national Department of Veterans Affairs (VA)–based registry, algorithms derived from pharmacy claims can acceptably rule in current prednisone use and use within 30 and 90 days, compared to a reference standard of self-report.
- In an exploratory analysis of participants reporting current prednisone use, we saw substantial agreement between self-reported and claims-based assessments of daily dose when limited to participants reporting a dose ≤5 mg/day, corresponding to a claims-based mean dose threshold of ≤7.5 mg/day. We observed reduced agreement for higher doses, perhaps due to dose variability not adequately captured by mean daily dose (eg, a prescribed taper, use of stockpiled prednisone).
- Using patient self-report as a reference standard, our algorithm underestimated prednisone dose, contextualized by our finding of high self-reported stockpiling and prescribing by non-VA providers.
- These findings may be used to estimate the performance of claims-based algorithms for glucocorticoid use in research and clinical contexts and to improve clinical estimation of glucocorticoid use not captured in claims.

INTRODUCTION

Up to 90% of patients with rheumatoid arthritis (RA) use glucocorticoids, with 25% to 50% taking them long-term.[1,2] Although glucocorticoids are effective in treating RA, there is concern that their dose-dependent toxicity outweighs their benefits.[3] Clinical trials cannot assess real-world usage patterns and are underpowered to assess toxicity, especially at lower doses.[4] Observational studies using electronic health records or administrative algorithms provide the opportunity to evaluate real-world glucocorticoid use and long-term toxicities, as well as to identify high-risk glucocorticoid use to inform clinical practice or quality improvement initiatives. However, algorithms that estimate glucocorticoid exposure from administrative claims may not adequately account for use that is different than prescribed or that changes over time (eg, a taper or burst). There is also no clear reference standard against which to benchmark these algorithms, and assessments of previously proposed glucocorticoid algorithms are limited and have substantial methodologic variability.[5,6] Self-report is likely to be the most accurate reference standard for glucocorticoid use in RA because patient-directed changes in use (eg, dose increases or reductions, "stockpiling") and prescribing by providers other than the patient's rheumatologist are common. Despite this, prior studies benchmarking on self-report are small and may not be generalizable to the majority of patients with RA.[6,7]

This study aimed to evaluate the accuracy of claims-based algorithms for current prednisone use in a population of veterans with RA using patient-reported use as a reference standard. Although self-reported medication use has limitations, primarily related to recall and desirability bias, it shows good agreement with use derived from other sources (eg, medical records, prescription data), is easier to obtain than provider report, and is more sensitive than provider report in capturing use occurring without that provider's knowledge (eg, patient-initiated changes, prescribing by another provider). Thus, demonstrating comparable performance between self-reported and provider-reported glucocorticoid use would help facilitate future research in this area.

PATIENTS AND METHODS

Participants. As part of clinical care, 284 participants enrolled in the multicenter prospective Veterans Affairs Rheumatoid Arthritis (VARA) Registry[8] were asked semistructured questions about their glucocorticoid use at the time of a routine in-person rheumatology clinic visit between October 2023 and August 2024 (Supplemental Table 1). Data reported were obtained from the following VARA sites: Omaha Department of Veterans Affairs (VA) Medical Center, Rocky Mountain Regional VA Medical Center, VA Ann Arbor Healthcare System, VA Durham Health Care System, VA Philadelphia Healthcare System, VA Puget Sound Health Care System, and VA Salt Lake City Healthcare System. Each site obtained institutional review board approval, and participants provided informed consent to participate in the VARA Registry.

Reference standard: Self-reported glucocorticoid use. We limited our analysis to oral prednisone-dispensing episodes to enable accurate dose estimation by participants. In a prior evaluation of prescription data, we found that prescriptions for prednisone represented 95.9% of oral glucocorticoids prescribed to VARA-enrolled patients (Supplemental Table 2). To assess self-reported current prednisone use, participants were asked, "Did you take prednisone yesterday?" with response options "yes" or "no." This served as the reference standard for current use. Participants who responded "yes" were asked to select the dose they took from the following categories: <5 mg, 5 mg, 10 mg, 15 mg, >15 mg, or a different dose. Participants were also asked about "recent use" in the past 30 and 90 days using the following categories: none, less than half the days (≤14 or ≤44 days in a 30- or 90-day period, respectively), more than half the days (≥15 or ≥45 days in a 30- or 90-day period, respectively), or "I took prednisone in the past 30 [or 90] days, but don't remember how often." Two barriers to accurate claims-based evaluation of prednisone use were also assessed:

prednisone "stockpiling" and obtaining prednisone from non-VA providers.

Administrative algorithms. We generated three administrative algorithms to evaluate current prednisone use using only prescription information captured directly in the VA Corporate Data Warehouse (CDW). These included the following: (1) an algorithm for current use, defined as an active prednisone-dispensing episode overlapping the self-report assessment date; (2) a 30-day algorithm for recent use, defined as an active prednisone course overlapping the 30 days before the self-report assessment date; and (3) a 90-day algorithm for recent use, defined as an active prednisone course overlapping the 90 days before the self-report assessment.[9]

For the current use algorithm (algorithm 1), we recorded the dispensing episode start date, the number of tablets dispensed, the days' supply of medication, and the unit dose for each prednisone-dispensing episode. We calculated the dispensing episode end date by adding the days' supply to the dispensing episode start date, and we calculated the cumulative prednisone dose by multiplying the number of tablets dispensed by the unit dose for the dispensing episode. To meet algorithm 1, the day before the self-report assessment was required to fall within an active prednisone-dispensing episode.[9]

For the recent use algorithms (algorithms 2 and 3), each prednisone course was defined as a group of all consecutive dispensing episodes that had a ≤90-day gap between the end date of one dispensing episode and the start date for the next dispensing episode. We defined course duration as the period from the start date of the first dispensing episode in the course through the end date of the last dispensing episode of the course (ie, the end of the last dose before a ≥90-day gap). We calculated the cumulative prednisone dose for each course by summing the total doses for each dispensing episode in that course, and we calculated the average daily dose by dividing the cumulative dose of prednisone for a course by the course duration (course end date minus course start date). To meet algorithm 2, a patient was required to have any part of an active prednisone course overlapping the period between the self-report assessment date and 30 days before the self-report assessment date. To meet algorithm 3, a patient was required to have any part of an active prednisone course overlapping the period between the self-report assessment date and 90 days before the self-report assessment date[9] (Figure 1). When assessing agreement for recent use, self-reported prednisone use during the past 30 and 90 days was dichotomized as any use versus no use.

Performance of algorithms against the reference standard. We calculated the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of each administrative algorithm using prednisone-dispensing data from the VA CDW against the reference standard of self-reported current use. Because there is no validated benchmark for assessing glucocorticoid use, we also calculated Cohen's kappa to assess agreement between administrative algorithms and self-reported use. Agreement based on Cohen's κ was classified as near perfect (0.8–1.0), substantial (0.6–0.8), moderate (0.4–0.6), fair (0.2–0.4), or slight (0.0–0.2).[10] We used Youden's J, equivalent to sensitivity + specificity – 1, to summarize algorithm performance; it ranges from –1 to 1, with 0 being equivalent to random chance and 1 representing a perfect test.[11]

For participants reporting current use, we performed an exploratory analysis assessing agreement between the following: (1) self-reported current use and an active dispensing episode in claims (Figure 2) and (2) self-reported daily dose (<5 mg, 5 mg, 10 mg, 15 mg, >15 mg, or other) and mean observed daily dose in claims for any active dispensing episode (Figure 3). To allow for error in our claims-based assessment of prednisone dose (eg, the fact that prescriptions intended to be taken as a burst and taper will be represented in our analysis as a single mean daily dose), we categorized claims-based mean prednisone dose thresholds as follows: no active course, ≤4 mg, 4.1 to 7.5 mg, 7.6 to 12.5 mg, 12.6 to 17.5 mg, and ≥17.6 mg. To assess agreement, we calculated the weighted Cohen's kappa for (1) all participants reporting current use, (2) participants with both self-reported current use and an active dispensing episode, and (3) participants with self-reported current use ≤5 mg and an active dispensing episode. For participants reporting never use, we identified the year of their last prescription-dispensing episode by examining all prednisone-dispensing episodes in the CDW since October 1, 1998 (Figure 4). Analyses were completed using Stata version 18. Data are reported in accordance with the policies of the VA Office of Research and Development.

RESULTS

Two hundred eighty-four participants provided data on prednisone use. Demographics reflect those of prior studies in the VARA Registry, with a mean age of 69.2 years, 84.2% of participants being male, 83.4% being White, and 68.7% being current or former smokers (Table 1). Eighty-seven percent of those surveyed were taking disease-modifying antirheumatic drugs (DMARDs), with 43.7% taking biologic or targeted synthetic DMARDs. The mean Clinical Disease Activity Assessment score in this population was 9.33, representing low-moderate RA disease activity.

Thirty-seven participants (13%) reported current prednisone use, 52 (18%) reported use in the past 30 days, and 59 (20%) reported use in the past 90 days. Of those reporting use in the past 30 and 90 days, 23 (44%) and 29 (49%) reported use on less than half the days and 29 (56%) and 29 (49%) reported use on more than half the days, respectively.

Table 2 shows the performance of each claims-based algorithm for current prednisone use. Algorithm 1, assessing current

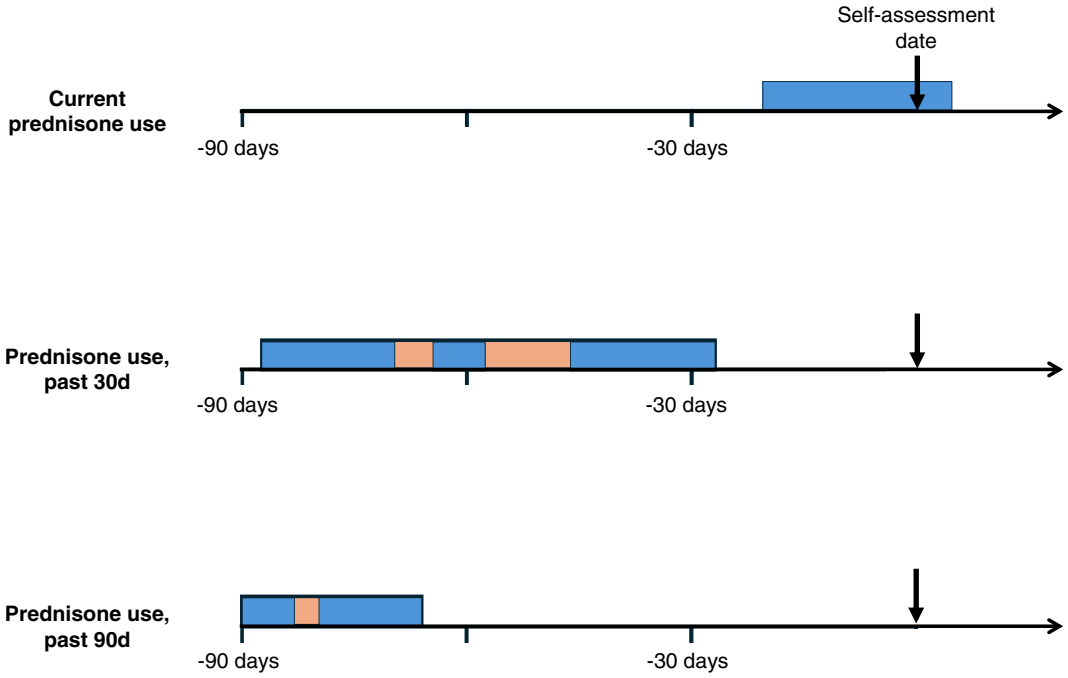


Figure 1. Schematic representation of claims-based algorithms for assessing current and recent prednisone use. Blue bars represent prednisone-dispensing episodes. Orange bars represent prednisone courses, defined as a group of all consecutive dispensing episodes with a ≤90-day gap between the end date of one dispensing episode and the start date for the next dispensing episode. Course duration was defined as the period from the start date of the first dispensing episode in the course through the end date of the last dispensing episode of the course (ie, the end of the last dose before a ≥90-day gap).

use, had a sensitivity of 0.70, a specificity of 0.98, a PPV of 0.84, and an NPV of 0.96. Algorithms 2 and 3, estimating use in the past 30 and 90 days, had sensitivities of 0.63 and 0.71, specificities of 0.96 and 0.92, PPVs of 0.79 and 0.72, and NPVs of 0.92 and 0.92, respectively. Substantial agreement was observed between administrative algorithms and self-reported

		Self-report			
		On GC	**Not on GC**	**Predictive value**	
CDW Prescripion	**+ Active GC prescription**	**26**	**5**	PPV: 0.84	**Total + prescription: 31**
	- Active GC prescription	**11**	**242**	NPV: 0.96	**Total - prescription: 253**
	Sensitivity & Specificity	Sensitivity 0.70	Specificity: 0.98	**Agreement:** Kappa 0.73 Accuracy 0.94	
		Total on GC 37	**Total not on GC 247**		

Figure 2. Error matrix: current prednisone use. CDW, Corporate Data Warehouse; GC, glucocorticoid; NPV, negative predictive value; PPV, positive predictive value. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25580/abstract.

	Self-report					
From CDW	**<5mg**	**5mg**	**10mg**	**15mg**	**>15mg**	**Total**
No active dispensing episode	4	3	1	0	1	9
≤4mg	8	1	1	0	0	10
4.1 to 7.5mg	0	10	2	0	0	12
7.6 to 12.5mg	1	1	0	0	1	3
12.5 to 17.5mg	0	0	0	1	0	1
≥17.6mg	0	0	0	0	0	0
N reporting a dose*	13	15	4	1	2	35

Figure 3. Daily dose from claims for participants reporting current prednisone use. Weighted κ for 26 participants with an active dispensing episode: 0.55 (95% CI 0.43–0.67, $P < 0.0001$). Weighted κ for 21 participants with an active dispensing episode reporting ≤5 mg/day: 0.67 (95% CI 0.49–0.85, $P = 0.0001$). * Two participants reported current use but did not report a verifiable dose. CDW, Corporate Data Warehouse; CI, confidence interval. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25580/abstract.

use. Cohen's κ and Youden's J for each algorithm were 0.68 and 0.68 for current use, 0.64 and 0.59 for use in the past 30 days, and 0.63 and 0.57 for use in the past 90 days, respectively. Percent agreement was 0.94 for current use, 0.90 for use in the past 30 days, and 0.88 for use in the past 90 days.

Figure 2 shows the error matrix for participants reporting current use compared to the claims-based algorithm. Of 37 participants reporting current use, 26 (70%) had a concurrent overlapping prednisone prescription. Of 247 participants reporting no current prednisone use, 242 (98%) had no concurrent overlapping prednisone prescription. The most common discordance was self-reported use without an overlapping prednisone prescription course (11 of 37 self-reporting use).

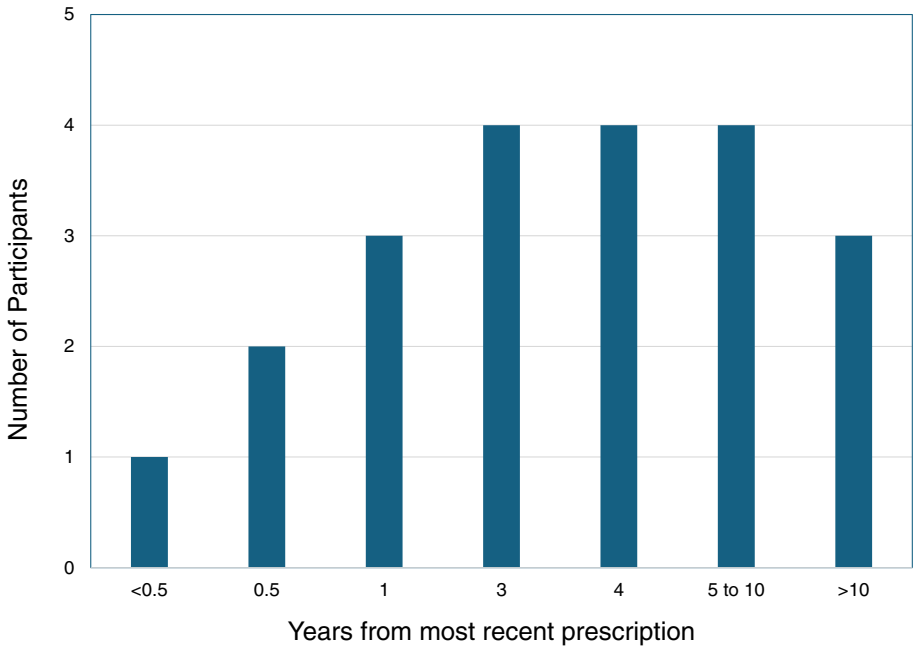


Figure 4. Time (years) from most recent prednisone prescription to self-report assessment for participants who report never using prednisone but have a pharmacy claim for it (n = 21). Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25580/abstract.

Table 1. Characteristics of patients with RA asked about glucocorticoid use*

Variable	Participants (N = 284)
Age, mean (SD), y	69.2 (10.9)
Male, n (%)	239 (84.2)
Site, n (%)	
Rocky Mountain Regional VAMC	15 (5.3)
Omaha VAMC	70 (24.6)
VA Philadelphia Healthcare System	23 (8.1)
VA Salt Lake City Healthcare System	104 (36.6)
VA Puget Sound Health Care System	53 (18.7)
Other sites[a]	19 (6.7)
Race, n (%)	
White	237 (83.4)
Black	27 (9.5)
Other or not reported	20 (7.0)
Ethnicity, n (%)	
Non-Hispanic	267 (94.01)
Hispanic or not reported	17 (6.0)
Smoking status, n (%)	
Current	46 (16.20)
Former	149 (52.46)
Never	82 (28.87)
Unknown	7 (2.46)
RA duration, mean (SD), y	16.1 (12.84)
Seropositivity, n (%)	
Anti-CCP	185 (65.14)
RF	173 (60.92)
Active MTX course at time of survey, n (%)	115 (40.49)
Active csDMARD course at time of survey, n (%)	202 (71.13)
Active btDMARD course at time of survey, n (%)	124 (43.66)
Active btDMARD or csDMARD course at time of survey, n (%)	248 (87.32)
Most recent CDAI at time of survey, mean (SD)	9.33 (9.71)

* btDMARD, biologic or targeted synthetic disease-modifying antirheumatic drug; CCP, cyclic citrullinated peptide; CDAI, Clinical Disease Activity Index; csDMARD, conventional synthetic disease-modifying antirheumatic drug; MTX, methotrexate; RA, rheumatoid arthritis; RF, rheumatoid factor; VA, Department of Veterans Affairs; VAMC, Department of Veterans Affairs Medical Center.
[a] VA Ann Arbor Healthcare System and VA Durham Health Care System.

Figure 3 shows the agreement between self-reported prednisone dose and mean observed daily dose for the 35 participants who reported current use and provided a dose. Of these, 9 (26%) had no prednisone prescription overlapping the date of the self-report assessment. Four (44%) of these nine reported having a prednisone stockpile, and two (22%) of these nine reported receiving a prednisone prescription from a non-VA provider in the past year. We observed moderate agreement between self-reported current dose and mean observed daily dose in this population, with a weighted Cohen's κ of 0.55 and agreement of 89%. Among the subset of participants with an active prednisone prescription who self-reported a current dose ≤5 mg/day (n = 19, 51% of those reporting current use), we observed substantial agreement with a claims-based dose of ≤7.5 mg/day, with a weighted Cohen's κ of 0.67 and agreement of 95%.

Thirty-two participants (11%) reported that they had never taken prednisone. Of these, 21 (66%) had a previous prednisone prescription fill documented in the CDW. Figure 4 shows the year of the most recent prednisone prescription for these 21 participants. Of these, zero had a prednisone prescription filled in the 90 days before the self-report assessment, three (9%) had their most recent prescription filled within the prior year, and four (19%) had their most recent prescription filled ≥10 years prior.

Sixty-six participants (23%) reported having a stockpile of prednisone. Of these, 49 (74%) reported no current use, 40 (60%) reported no use in the past 30 days, and 35 (53%) reported no use in the past 90 days. One hundred forty-five participants (51%) reported that they had never received prednisone from a non-VA provider. Of the remainder, 13 (9%) reported receiving prednisone from a non-VA provider in the past year, and 60 (43%) reported receiving prednisone from a non-VA provider more than a year prior.

DISCUSSION

In a cohort of participants with RA enrolled in the VARA Registry,[8] claims-based algorithms evaluating prednisone use in the past day, the past 30 days, and the past 90 days performed acceptably when benchmarked on self-report. In particular, the NPV of these algorithms was excellent, exceeding 0.9 in all three cases, partly because the prevalence of use was generally low. In a cohort with a higher prevalence of use, we may expect a superior PPV using these algorithms but a somewhat worse NPV.

Among the subset of participants reporting current use, we saw moderate agreement between self-reported and claims-based assessments of current daily dose; agreement became substantial when we limited the analysis to the 51% of participants reporting a daily dose of ≤5 mg/day. For our current use algorithm, misclassification was low and predominantly underrepresented exposure. For our current dose algorithm, misclassification underrepresented exposure in the setting of one in four participants reporting they had a prednisone stockpile and one in four participants reporting they received prednisone from a provider not captured in the claims source used. Overall, our results suggest that the claims-based algorithms we present here perform adequately to (1) rule out prednisone use during the past 90 days in the context of low rates of use (NPV > 90%), (2) rule in current use in a cohort with similar rates of use (PPV > 80%), and (3) estimate current daily dose for participants with claims-based dose estimates ≤7.5 mg/day of the prednisone equivalent. These findings support the use of such algorithms, compared to a benchmark of self-report, for estimating glucocorticoid exposure in both research and clinical contexts.

We observed reduced agreement with higher reported mean daily doses, which may be due to prescribing patterns not adequately represented by a mean daily dose, such as a taper, or to use not adequately captured in claims, such as stockpiled or

Table 2. Performance characteristics for administrative algorithms vs self-reported GC use*

Concept	Question	CDW algorithm	Sensitivity	Specificity	PPV	NPV	κ	Youden's J
Current GC use	Did you take prednisone yesterday?	1. Active prednisone prescription overlapping day of assessment	0.70	0.98	**0.84**	0.96	0.68	0.68
Recent GC use	In the past 30 days, how often did you take prednisone?	2. Active prednisone course[9] overlapping the period from 30 days before self-report assessment to date of assessment	0.63	0.96	**0.79**	0.92	0.64	0.59
	In the past 90 days, how often did you take prednisone?	3. Active prednisone course[9] overlapping the period from 90 days before self-report assessment to date of assessment	0.71	0.92	**0.72**	0.92	0.63	0.63
Barriers to estimating GC use in claims	Do you have a "stockpile" or "reserve supply" of extra prednisone pills saved up?	–	–	–	–	–	–	–
	Have you ever gotten prednisone from a health care provider outside the VA?	–	–	–	–	–	–	–

* Bold values indicate positive predictive value. CDW, Corporate Data Warehouse; GC, glucocorticoid; NPV, negative predictive value; PPV, positive predictive value; VA, Department of Veterans Affairs.

externally prescribed prednisone. We observed some decline in claims-based algorithm performance, as our algorithms assessed longer lookback periods, and reduced agreement for questions assessing ever use of prednisone. This may be a limitation of the reference standard due to inferior recall for remote use and suggests algorithms benchmarked on self-report may not be ideal for validating administrative measures of long-term exposure.

Prior work in this area includes a recent study of 494 patients with RA enrolled in the CorEvitas registry that compared 90-day average daily glucocorticoid dose in linked Medicare claims to a gold standard of physician-reported use at a registry visit during that period.[5] Although agreement benchmarking on physician report was similar (overall κ of 0.61 and percent agreement of 0.9), the prevalence of physician-reported 90-day use was 31% (vs our finding of 20% self-reported use), sensitivity was higher at 0.88, and specificity was lower at 0.79. These differences may be due to different algorithm criteria and a distinct reference standard used.

A small UK-based study of 78 patients with RA reported excellent algorithm performance (sensitivity 0.84, specificity 0.87, PPV 0.87, NPV 0.85) when comparing presence of an active glucocorticoid prescription in the Clinical Practice Research Datalink to self-reported glucocorticoid use via mail-in survey. However, an important distinction from the current study is that only patients with documented glucocorticoid use in the past two years were eligible, and the survey response rate was only 16%, raising some concern for nonparticipant bias.[6] A single-center study of 91 participants in the Brigham Rheumatoid Arthritis Sequential Study registry found a moderate to high correlation between self-reported cumulative glucocorticoid dose and cumulative dose derived from chart review (Pearson's r 0.59) and strong agreement between self-report and chart-derived quartiles of dose (weighted κ 0.67). Thus, claims-based algorithms evaluating current and recent glucocorticoid use appear to have reasonable performance despite differences in population, reference standard, and algorithm criteria. As the largest and most recent evaluation to benchmark glucocorticoid use on self-report as a reference standard, our current study supports its use as a reference standard for validation of claims-based algorithms. This is especially relevant in settings where provider report may underestimate glucocorticoid use (eg high patient sharing, lower provider connectedness).[12]

We chose self-report as our reference standard for several reasons. First, self-reported glucocorticoid use in RA has only been examined as a reference standard in Europe, where glucocorticoid prescribing patterns differ from those in the United States.[13] Second, self-report not only is more likely to capture glucocorticoid use occurring without a provider's knowledge but also allows us to ask about and assess the impact of potential sources of such use, as we do here. Third, although self-reported glucocorticoid use suffers from recall bias, it is not clear that provider report suffers less from such bias. In prior studies, provider-reported use was either derived directly from the medical record or collected during a patient visit when the provider had access to both the medical record and the patient.[5,6,14] Provider-reported glucocorticoid can thus be conceptualized as

the provider's transcription of (1) patient-reported use at a visit and (2) prescribing data and/or previous provider reports from the electronic medical record. Provider-reported glucocorticoid use is highly variable in consistency and detail and does not capture use of glucocorticoids from outside sources (stockpiles, outside prescribers), patient-initiated dose reduction or discontinuation, or nondocumented verbal instructions given by the provider.[15] Fourth, the literature comparing provider report and self-report as benchmarks for glucocorticoid use is limited. Demonstrating comparable performance between benchmarks would help facilitate future research in this area. It is also reassuring that self-reported medication use generally shows good agreement with use derived from other data sources, even for medications that are used episodically (eg, antacids, antibiotics, muscle relaxants).[16–18] Although we designed this study with self-report use as the reference standard, it is likely more appropriate to consider prescription claims-based algorithms and self-reported use as complementary assessments. Each has its own strengths and limitations, and combining these two sources together would be ideal, when feasible.

Our data also provide valuable input into glucocorticoid stockpiling, external glucocorticoid sources, and patient-directed glucocorticoid consumption. Although outside the scope of this study, future work investigating ways to reduce stockpiling could be beneficial in limiting avoidable glucocorticoid exposure in RA populations.

Our study is among the first and largest to evaluate the performance of claims-based algorithms for glucocorticoid use compared to self-report. It is also, to our knowledge, the first to directly assess potential barriers to developing accurate algorithms, such as stockpiling and prescribing not captured in claims. Limitations include that we assessed prednisone use only among patients with RA who were both enrolled in a prospective registry and had a rheumatology visit during the assessment interval. Thus, our results may not be generalizable to patients who do not receive regular rheumatologic care or who receive care at lower-resourced sites. Relatedly, we found a relatively low prevalence of prednisone use in general and high-dose use (≥10 mg/day) in particular. Although this finding is consistent with American College of Rheumatology recommendations to taper glucocorticoids promptly and avoid high-dose use,[3] it also limits our ability to validate claims-based algorithms assessing high-dose use. Additionally, although self-reported current prednisone dose assessed the day before the questionnaire reasonably approximates the mean dose for an active dispensing episode in our sample, these measures do not reflect the fact that prednisone dose can vary from day to day within a single dispensing episode, for example, when a taper is prescribed. We were not able to capture prednisone prescribed by non-VA providers, though prior work suggests veterans with RA are relatively unlikely to use medical care.[19] To maximize algorithm performance, we did not assess prednisone use for RA specifically, such that our results likely capture prescriptions given by nonrheumatologists for other indications (eg, chronic obstructive pulmonary disease flares). Our analysis was limited to veterans and used VA claims; however, our results largely agree with and support prior evidence in non-veteran populations.[5,6,14] We cannot exclude the possibility of survey bias in our questionnaire responses, for example, recall bias (which may be increased among older participants) or demand bias from underreporting use unrelated to RA treatment or reporting a prescribed dose when taking differently. We attempted to mitigate this by limiting our evaluation to prednisone use, which constitutes >95% of oral glucocorticoids prescribed to this population. We also note that prior work evaluating the validity of self-reported medication use in the past year, including evaluations focused on older adults, shows excellent agreement with pharmacy claims.[20,21]

In conclusion, we found that algorithms derived from VA claims can acceptably assess ongoing prednisone use in patients with RA compared to a reference standard of patient self-report. Such algorithms may also be useful for estimating mean daily doses ≤7.5 mg/day of the prednisone equivalent, though we had limited ability to assess their utility for higher doses due to low prevalence. Our algorithm somewhat underestimated prednisone dose, contextualized by our finding of high self-reported glucocorticoid stockpiling and prescribing by non-VA providers. These findings may be used to estimate the performance of claims-based algorithms for glucocorticoid use in research and clinical contexts and to estimate the impact of use that cannot be captured in claims.

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 Wallace 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. Wallace BI, Lin P, Kamdar N, et al. Patterns of glucocorticoid prescribing and provider-level variation in a commercially insured incident rheumatoid arthritis population: a retrospective cohort study. Semin Arthritis Rheum 2020;50(2):228–236.
2. George MD, Baker JF, Wallace B, et al. Variability in glucocorticoid prescribing for rheumatoid arthritis and the influence of provider preference on long-term use of glucocorticoids. Arthritis Care Res (Hoboken) 2021;73(11):1597–1605.
3. Fraenkel L, Bathon JM, England BR, et al. 2021 American College of Rheumatology guideline for the treatment of rheumatoid arthritis. Arthritis Care Res (Hoboken) 2021;73(7):924–939.
4. Palmowski A, Nielsen SM, Boyadzhieva Z, et al. Safety and efficacy associated with long-term low-dose glucocorticoids in rheumatoid

arthritis: a systematic review and meta-analysis. Rheumatology (Oxford) 2023;62(8):2652–2660.

5. Galvao RW, Curtis JR, Harrold LR, et al. Accuracy of administrative claims prescription fill data to estimate glucocorticoid use and dose in patients with rheumatoid arthritis. Pharmacoepidemiol Drug Saf 2023;32(11):1271–1279.
6. Joseph RM, van Staa TP, Lunt M, et al. Exposure measurement error when assessing current glucocorticoid use using UK primary care electronic prescription data. Pharmacoepidemiol Drug Saf 2019; 28(2):179–186.
7. Solomon DH, Stedman M, Licari A, et al. Agreement between patient report and medical record review for medications used for rheumatoid arthritis: the accuracy of self-reported medication information in patient registries. Arthritis Rheum 2007;57(2):234–239.
8. Mikuls TR, Baker JF, Cannon GW, et al. The Veterans Affairs Rheumatoid Arthritis Registry: a unique population in rheumatoid arthritis research. Semin Arthritis Rheum 2025;70S:152580.
9. George MD, Sauer BC, Teng CC, et al. Biologic and glucocorticoid use after methotrexate initiation in patients with rheumatoid arthritis. J Rheumatol 2019;46(4):343–350.
10. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas 1960;20(1):37–46.
11. Ruopp MD, Perkins NJ, Whitcomb BW, et al. Youden Index and optimal cut-point estimated from observations affected by a lower limit of detection. Biom J 2008;50(3):419–430.
12. Cohen-Mekelburg S, Yu X, Costa D, et al. Variation in provider connectedness associates with outcomes of inflammatory bowel diseases in an analysis of data from a national health system. Clin Gastroenterol Hepatol 2021;19(11):2302–2311.e1.
13. Wallace BI, Tsai HJ, Lin P, et al. Prevalence and prescribing patterns of oral corticosteroids in the United States, Taiwan, and Denmark, 2009–2018. Clin Transl Sci 2023;16(12):2565–2576.
14. Solomon DH, Ayanian JZ, Yelin E, et al. Use of disease-modifying medications for rheumatoid arthritis by race and ethnicity in the National Ambulatory Medical Care Survey. Arthritis Care Res (Hoboken) 2012;64(2):184–189.
15. Lamore RF III, Hechenbleikner EM, Ha C, et al. Perioperative glucocorticoid prescribing habits in patients with inflammatory bowel disease: a call for standardization. JAMA Surg 2014;149(5):459–466.
16. Brüne M, Emmel C, Meilands G, et al. Self-reported medication intake vs information from other data sources such as pharmacy records or medical records: identification and description of existing publications, and comparison of agreement results for publications focusing on patients with cancer - a systematic review. Pharmacoepidemiol Drug Saf 2021;30(5):531–560.
17. Grimaldi-Bensouda L, Rossignol M, Aubrun E, et al; the PGRx Study Group. Agreement between patients' self-report and physicians' prescriptions on nonsteroidal anti-inflammatory drugs and other drugs used in musculoskeletal disorders: the international Pharmacoepidemiologic General Research eXtension database. Pharmacoepidemiol Drug Saf 2012;21(7):753–759.
18. Richardson K, Kenny RA, Peklar J, et al. Agreement between patient interview data on prescription medication use and pharmacy records in those aged older than 50 years varied by therapeutic group and reporting of indicated health conditions. J Clin Epidemiol 2013; 66(11):1308–1316.
19. Schwab P, Sayles H, Bergman D, et al. Utilization of care outside the Veterans Affairs health care system by US veterans with rheumatoid arthritis. Arthritis Care Res (Hoboken) 2017;69(6):776–782.
20. West SL, Savitz DA, Koch G, et al. Recall accuracy for prescription medications: self-report compared with database information. Am J Epidemiol 1995;142(10):1103–1112.
21. Hafferty JD, Campbell AI, Navrady LB, et al. Self-reported medication use validated through record linkage to national prescribing data. J Clin Epidemiol 2018;94:132–142.

Arthritis Care & Research
Vol. 77, No. 11, November 2025, pp 1359–1367
DOI 10.1002/acr.25582

American College of Rheumatology
Empowering Rheumatology Professionals

Rheumatoid Arthritis and Risk of Thyroid Cancer: A Nationwide Cohort Study

Mi Hee Cho,[1] Jinhyung Jung,[2] Tae Hyuk Kim,[3] Kyungdo Han,[4] Dagyeong Lee,[5] Keun Hye Jeon,[6] In Young Cho,[7] and Dong Wook Shin[8]

Objective. Despite extensive cancer-related research in patients with rheumatoid arthritis (RA), there remains a paucity of studies on thyroid cancer in patients with RA. We investigated the risk of thyroid cancer in patients with RA using updated definitions to identify patients with RA and incident cases of thyroid cancer with consideration of RA serologic status.

Methods. Using a nationwide database, we identified 40,895 patients with newly diagnosed RA between 2010 and 2017 and matched them by sex and age at a 1:5 ratio to a control population of 204,475 individuals without RA. The association of thyroid cancer and RA with consideration of serostatus was investigated using Cox regression analyses. Stratified analyses by sex and age were conducted using the same Cox modeling.

Results. During 5.5 years of follow-up, compared to the matched control group, the adjusted hazard ratio (aHR) of thyroid cancer in patients with overall RA was 0.98 (95% confidence interval [CI] 0.85–1.13). Compared to the seronegative patients with RA, seropositive patients with RA did not show a significantly different risk of thyroid cancer (aHR 1.03, 95% CI 0.78–1.36). Stratified analyses by age or sex showed no statistical interaction.

Conclusion. Patients with newly diagnosed RA did not show a different risk of thyroid cancer compared to a matched control group. The risk of thyroid cancer incidence was not affected by serologic status of RA, sex, or age.

INTRODUCTION

In patients with rheumatoid arthritis (RA), various comorbidities, including cardiovascular disease, infections, and cancers, increase the disease burden and reduce life expectancy.[1,2] Cancer is one of the leading causes of death in patients with RA,[1,2] and the autoimmune and chronic inflammatory pathogenesis of RA along with common environmental and genetic risk factors between RA and cancers may increase the risk of cancer in patients with RA.[3,4] However, although previous meta-analyses reported an increased risk of overall cancer in patients with RA (pooled standardized incidence ratio [SIR] 1.09, 95% confidence interval [CI] 1.06–1.13), associations between RA and site-specific cancers, especially solid cancers, have varied in individual studies.[3,5]

Previous research regarding the relationship between RA and thyroid diseases has mostly focused on the autoimmune feature because autoimmune thyroid diseases (AITDs) are common worldwide.[6–8] However, malignancies may also develop in the thyroid, accounting for 3.0% of the global cancer incidence.[9] Although there have been extensive research on the associations regarding RA and cancer,[3–5] RA and AITDs,[6–8] or AITDs and thyroid cancer,[10–12] the association between RA and thyroid cancer remains inconclusive. In studies of patients with RA, risk of thyroid cancer varied depending on the study design and the cohort of individual studies (Supplementary Table S1).[12–17] Studies using single-center data in Japan or Korea[14,16] or using Taiwanese National Health insurance data[13,15] showed conflicting results:

Partly supported by the Institution of Quality of Life in Cancer, funded by Samsung Fire & Marine Insurance.

[1]Mi Hee Cho, MD, PhD: Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea; [2]Jinhyung Jung, PhD: Samsung Biomedical Research Institute, Sungkyunkwan University School of Medicine, Suwon, Republic of Korea; [3]Tae Hyuk Kim, MD, PhD: Thyroid Center, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea; [4]Kyungdo Han, PhD: Soongsil University, Seoul, Republic of Korea; [5]Dagyeong Lee, MD: Hallym University Dongtan Sacred Heart Hospital, Hwaseong, Republic of Korea; [6]Keun Hye Jeon, MD: CHA Gumi Medical Center, CHA University School of Medicine, Gumi, Republic of Korea; [7]In Young Cho, MD, MPH: Samsung Medical Center and Samsung Advanced Institute for Health Science & Technology, Sungkyunkwan University, Seoul, Republic of Korea; [8]Dong Wook Shin, MD, DrPH, MBA: Samsung Medical Center and Samsung Advanced Institute for Health Science & Technology, Sungkyunkwan University and Center for Trend Monitoring-Risk Modeling, Institution of Quality of Life in Cancer, Samsung Medical Center, Seoul, Republic of Korea.

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.25582).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25582.

Address correspondence via email to In Young Cho, MD, MPH, at ciyoung0604@gmail.com; or to Dong Wook Shin, MD, DrPH, MBA, at dwshin.md@gmail.com.

Submitted for publication December 11, 2024; accepted in revised form May 1, 2025.

SIGNIFICANCE & INNOVATIONS

- To the best of our knowledge, this study is the first to evaluate risk of thyroid cancer in patients with rheumatoid arthritis (RA) with concomitant consideration of RA serostatus.
- Compared to the non-RA population, patients with newly diagnosed RA did not show a different risk of thyroid cancer.
- The risk of thyroid cancer incidence was not affected by serologic status of RA, sex, or age.

some showed an increased incidence,[13,16] whereas others did not.[14,15] These studies[13–16] included a small number of thyroid cancer cases (~23 cases) and reported SIR which could not be adjusted for important confounding factors, such as smoking and alcohol drinking.[18,19] Two recent studies using the Korean National Health Insurance Database evaluated risk of thyroid cancer compared with a matched control group with adjustment for important confounders, but the results were inconsistent; the adjusted hazard ratios (aHRs) in the overall RA group were 1.01 (95% CI 0.69–1.46) in the study by Choi et al[17] and 1.76 (95% CI 1.07–2.90) in the study by Park et al[12] However, the study by Choi et al did not consider serologic status,[17] and the study by Park et al included only patients with seropositive RA (SPRA).[12] Also, these studies defined thyroid cancer outcome only using diagnostic codes,[12,17] which are potentially subject to overestimation of cancer cases.

Therefore, the objective of this study was to investigate the risk of thyroid cancer in patients with RA using updated definitions to identify patients with RA and incident cases of thyroid cancer with appropriate consideration of major confounding factors. In addition, stratified analyses were conducted according to serologic status, sex, and age to examine the association between RA and thyroid cancer.

METHODS

Data source. This retrospective cohort study was conducted using the National Health Information Database (NHID) provided by the Korean National Health Insurance Service (NHIS) for research purposes. The Korean NHIS provides universal health insurance for 97% of the Korean population and medical aid for the remaining 3% of the population in the lowest income bracket through financial subsidization by the government. In addition, the beneficiaries ≥40 years of age or employees at any age are eligible to participate in free health screening examination programs offered by the NHIS at least once every two years. Hence, the NHID includes all demographic and health care information, such as outpatient clinic visits, hospitalizations, diagnostic codes, and medication prescriptions, as well as national health examination results.[20,21]

Identification of patients with RA. Separate operational definitions were applied to identify patients with newly diagnosed SPRA and seronegative RA (SNRA), as in previous research.[22–25] The Korean NHIS provides the Rare and Intractable Disease (RID) program for patients with certain diseases, including cancers and systemic autoimmune disease. Patients registered in the RID program receive beneficial cost reduction for relevant medical expenses. For patients with RA, only patients with SPRA who meet the following conditions are eligible for the RID program: a positive result for rheumatoid factor (RF) or anti–citrullinated protein antibodies (ACPAs) and a doctor's official certification of RA.

In this study, patients with SPRA were defined using the diagnostic code M05 from the *International Classification of Diseases, Tenth Revision* (ICD-10) and the RID registration code V223 with prescription records of any disease-modifying antirheumatic drugs (DMARDs) (Supplementary Table S2). Patients who visited medical institutions, including hospitalization and outpatient visits, with claims records of ICD-10 code M05 for the first time between 2010 and 2017, followed by registration in the RID program, and who also had prescription records of DMARDs for ≥270 days were defined as patients with newly diagnosed SPRA. Patients with initial claims records of ICD-10 code M06, except M06.1 and M06.4, between 2010 and 2017 and prescription records of DMARDs for ≥270 days were defined as patients with newly diagnosed SNRA.

Study population. We initially selected individuals who had no claims records of either RA-related ICD-10 or RID codes before January 1, 2010, and who were not diagnosed with any other rheumatologic diseases identified using ICD-10 or RID codes (Supplementary Table S3). After applying the aforementioned operational definition for patients with RA, the initial cohort of patients with newly diagnosed RA between 2010 and 2017 and aged ≥40 years numbered 92,336 people. From the initial cohort, we excluded patients who did not participate in the national health examination within the two years before RA diagnosis (n = 37,426), those who had missing data from the national health checkup (n = 1,925), and those who were diagnosed with any cancer (n = 2,902) before the index date. We additionally excluded those with a diagnosis of thyroid cancer within a lag period of one year from the index date (n = 491). The same inclusion and exclusion criteria described here were also applied to the non-RA population. Afterward, each identified patient with RA was matched 1:5 to individuals without RA on age and sex,[26] and the final study population included 40,895 patients with RA and 204,475 individuals without RA (Figure 1).

The study protocol was approved by the Institutional Review Board of Samsung Medical Center (IRB no. SMC 2022-06-141). Due to the anonymity of the NHIS data, the requirement for informed consent from individual participants was waived.

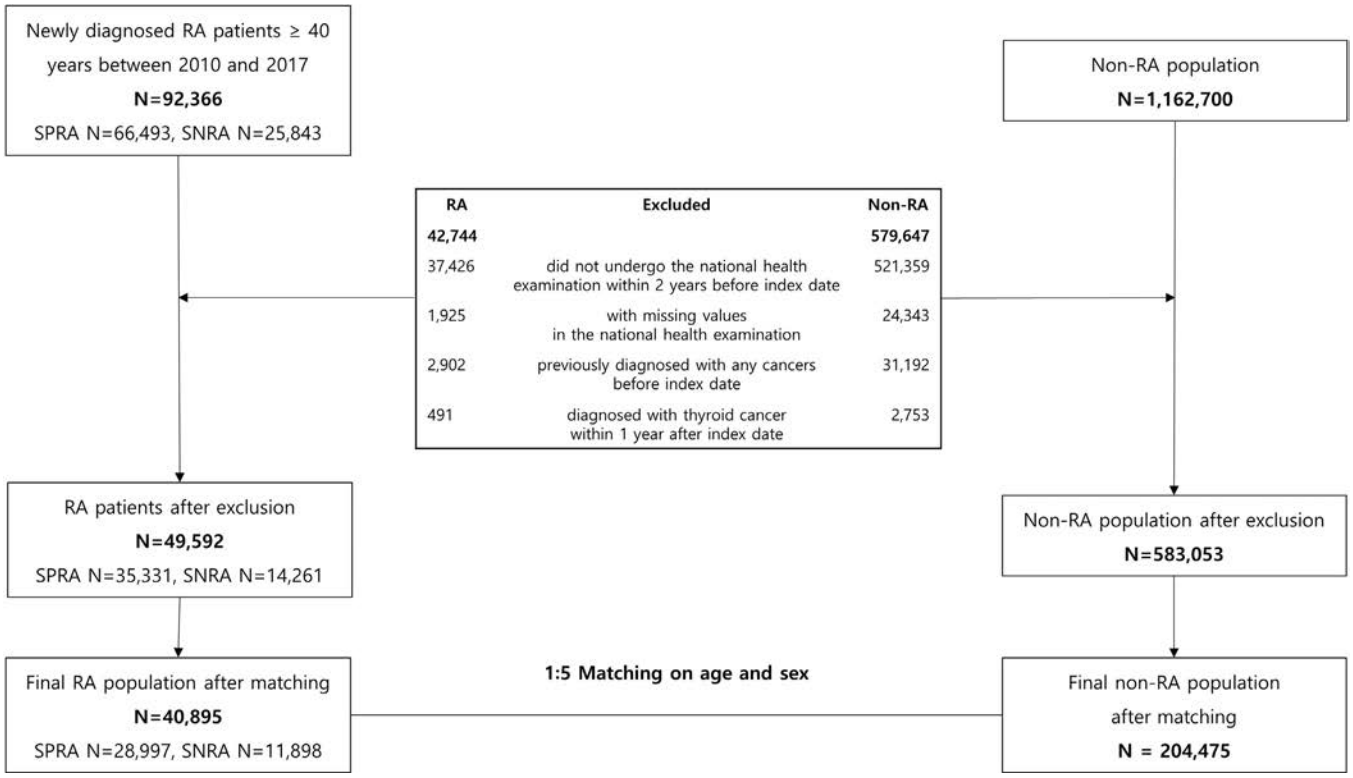


Figure 1. Flow diagram of the study population. Patients with RA were matched by sex and age at a 1:5 ratio with patients without RA. RA, rheumatoid arthritis; SNRA, seronegative rheumatoid arthritis; SPRA, seropositive rheumatoid arthritis.

Outcomes: Thyroid cancer incidence. The end point of this study was newly diagnosed thyroid cancer. Incidence of thyroid cancer was identified with a combination of the ICD-10 code C73 in claims records and registration for the critical illness copayment reduction program. In Korea, patients with cancer are eligible for the critical illness copayment reduction program, which provides a beneficial cost reduction to 5% copayment for cancer-related medical expenses compared to 30% copayment for other general medical conditions; nearly all patients with cancer are registered in this program. Ascertainment of cancer cases using this information has been validated in the Korean health care setting.[27] The date when the RA-related diagnostic code was first registered was defined as the index date. The follow-up started one year after the index date and ended at the date of thyroid cancer diagnosis (incidence), death from any cause, or December 31, 2020, whichever came first.

Covariates. Information from anthropometric examination (height, weight, and blood pressure) and lifestyle (smoking status, alcohol consumption, and exercise) based on self-reported questionnaires was obtained from a national health examination performed within two years before the index date. Individuals who reported alcohol intake >0 g/day were defined as drinkers. Smoking status was categorized into five groups: never-smoker, ex-smoker at <20 pack-years (PY), ex-smoker at ≥20 PY, current smoker at <20 PY, and current smoker at ≥20 PY. Moderate exercise for five or more days or vigorous exercise for three or more days in a week was defined as regular exercise. Body mass index (BMI) was calculated by dividing weight by the square of height (kg/m^2), and individuals with BMI ≥25 were categorized as having obesity.[28,29]

Comorbidities were defined as follows: (1) diabetes, either diabetes-related ICD-10 codes (E11–E14) with prescription records of antidiabetic medications or a fasting glucose level ≥126 mg/dL recorded in a national health examination; (2) hypertension, either hypertension-related ICD-10 codes (I10–I13, I15) with prescription records of antihypertensive medications or high blood pressure (systolic blood pressure ≥140 mm Hg or diastolic blood pressure ≥90 mm Hg) measured in a national health examination; (3) dyslipidemia, dyslipidemia-related ICD-10 code (E78) with prescription records of antidyslipidemia medications or total cholesterol level ≥240 mg/dL recorded in a national health examination; and (4) chronic kidney disease, an estimated glomerular filtration rate <60 mL/min/1.73 m^2, calculated using the Modification of Diet in Renal Disease equation in a national health examination.

Statistical analysis. Baseline characteristics of the study population were presented as numbers with percentages for categorical variables and as means with SDs for continuous variables. Comparisons between groups were performed using

Table 1. Baseline characteristics of the study population*

		RA			Serologic status of RA		
	Total (N = 245,370)	No (n = 204,475)	Yes (n = 40,895)	*P*	Seropositive (n = 28,997)	Seronegative (n = 11,898)	*P*
Sex, male, n (%)	61, 566 (25.1)	51,305 (25.1)	10,261 (25.1)	1.000	7,048 (24.3)	3,213 (27.0)	<0.001
Age, n (%)							
40–64 y	142,914 (58.2)	119,095 (58.2)	23,819 (58.2)	1.000	16,587 (57.2)	7,232 (60.8)	<0.001
≥65 y	102,456 (41.8)	85,380 (41.8)	17,076 (41.8)		12,410 (42.8)	4,666 (39.2)	
Mean age (SD), y	57.8 (9.4)	57.8 (9.4)	57.8 (9.4)	1.000	58.0 (9.4)	57.1 (9.5)	<0.001
Smoking, n (%)				<0.001			<0.001
Never	195,078 (79.5)	163,201 (79.8)	31,877 (78.0)		22,649 (78.1)	9,228 (77.6)	
Ex-smoker, <20 PY	14,361 (5.9)	11,912 (5.8)	2,449 (6.0)		1,551 (5.3)	898 (7.6)	
Ex-smoker, ≥20 PY	9,445 (3.8)	7,547 (3.7)	1,898 (4.6)		1,299 (4.5)	599 (5.0)	
Current, <20 PY	13,089 (5.3)	10,960 (5.4)	2,129 (5.2)		1,506 (5.2)	623 (5.2)	
Current, ≥20 PY	13,397 (5.5)	10,855 (5.3)	2,542 (6.2)		1,992 (6.9)	550 (4.6)	
Alcohol drinking, n (%)	71,871 (29.3)	61,828 (30.2)	10,043 (24.6)	<0.001	6,920 (23.9)	3,123 (26.3)	<0.001
Regular exercise, n (%)	50,611 (20.6)	43,069 (21.1)	7,542 (18.4)	<0.001	5,207 (18.0)	2,335 (19.6)	<0.001
Income, lowest quartile, n (%)	54,714 (22.3)	45,463 (22.2)	9,251 (22.6)	0.086	6,571 (22.7)	2,680 (22.5)	0.765
Obesity, n (%)	81,417 (33.2)	68,978 (33.7)	12,439 (30.4)	<0.001	8,459 (29.2)	3,980 (33.5)	<0.001
BMI, mean (SD)	23.9 (3.2)	23.9 (3.2)	23.6 (3.2)	<0.001	23.5 (3.2)	23.9 (3.3)	<0.001
SBP, mean (SD), mm Hg	123.3 (15.3)	123.4 (15.4)	122.4 (15.2)	<0.001	122.4 (15.2)	122.3 (15.2)	0.636
DBP, mean (SD), mm Hg	76.1 (10.0)	76.2 (10.0)	75.6 (9.9)	<0.001	75.7 (9.9)	75.6 (9.9)	0.335
Comorbidities, n (%)							
Diabetes	30,844 (12.6)	25,734 (12.6)	5,110 (12.5)	0.616	3,523 (12.2)	1,587 (13.3)	0.001
Hypertension	91,217 (37.2)	75,140 (36.8)	16,077 (39.3)	<0.001	10,977 (37.9)	5,100 (42.9)	<0.001
Dyslipidemia	80,618 (32.9)	67,030 (32.8)	13,588 (33.2)	0.080	9,212 (31.8)	4,376 (36.8)	<0.001
Chronic kidney disease	15,699 (6.4)	12,581 (6.2)	3,118 (7.6)	<0.001	2,064 (7.1)	1,054 (8.9)	<0.001
Thyroid cancer event, n (%)	1,507 (0.6)	1,263 (0.6)	244 (0.6)	0.619	175 (0.6)	69 (0.6)	0.779
Follow-up duration, y							
Mean (SD)	5.5 (2.2)	5.5 (2.2)	5.4 (2.2)	<0.001	5.4 (2.2)	5.4 (2.2)	0.066
Median (Q1–Q3)	5.3 (3.6–7.3)	5.4 (3.6–7.4)	5.3 (3.5–7.3)	<0.001	5.2 (3.5–7.3)	5.3 (3.5–7.2)	0.100

* BMI, body mass index; DBP, diastolic blood pressure; PY, pack-year; Q, quartile; RA, rheumatoid arthritis; SBP, systolic blood pressure.

t-tests for continuous variables and chi-square tests for categorical variables. The Kaplan-Meier plot was used to illustrate the cumulative incidence probability of thyroid cancer in patients with RA compared with the non-RA population.

Cox proportional hazards analyses were performed to estimate hazard ratios and 95% CIs of thyroid cancer in patients with RA. The proportional hazard assumption was assessed using Schoenfeld's residuals. Incidence rates of thyroid cancer were calculated as per 1,000 person-years. The analysis models were adjusted for the following variables sequentially. Model 1 was the crude model. Model 2 was adjusted for well-recognized factors associated with thyroid cancer in the general population: age,[30] sex,[30] smoking status,[18,19] alcohol drinking,[18,19] physical activity,[31] income level,[32] BMI,[33] and comorbid diabetes.[34] Model 3 was adjusted for variables in model 2 plus comorbid hypertension, dyslipidemia, and chronic kidney disease.[35–37] Although hypertension and dyslipidemia are not established risk factors for thyroid cancer, previous studies conducted in Korea reported that metabolic syndrome and its components were associated with an increased risk of thyroid cancer.[36,37] Potential interactions between covariates were examined by adding cross-combinations of variables to the final analysis model (model 3). For sensitivity analysis, additional analyses using the stratified Cox model with the same covariates were conducted.

Analyses stratified by sex and age were conducted using the same Cox proportional hazard modeling. In addition, competing risk analyses using the Fine–Gray proportional subdistribution hazards model were performed to calculate subdistribution hazard ratios and 95% CIs for the outcome events with death as a competing risk. Statistical analyses were performed using SAS version 9.4 (SAS Institute Inc), and two-tailed *P* values <0.05 were considered to indicate statistical significance.

RESULTS

Baseline characteristics. In a total of 40,895 patients with newly diagnosed RA, 30,634 (74.9%) were female and 28,997 (70.9%) had SPRA (Table 1). The mean age of the participants was 57.8 (SD 9.4) years at baseline. Compared with the matched non-RA group, patients with RA showed a higher prevalence of smoking but were less likely to be drinkers or have obesity (Table 1). Compared with patients with SNRA, patients with SPRA showed higher current smoking rates but lower alcohol consumption rates (Table 1).

Risk of thyroid cancer in patients with RA. During a mean follow-up of 5.5 (SD 2.2) years after a 1-year lag period, the number of diagnosed thyroid cancer cases was 244 in the RA group and 1,263 in the matched non-RA group (Table 1). Incidence probabilities of thyroid cancer in patients with RA during follow-up are presented in Figure 2. Compared with the matched

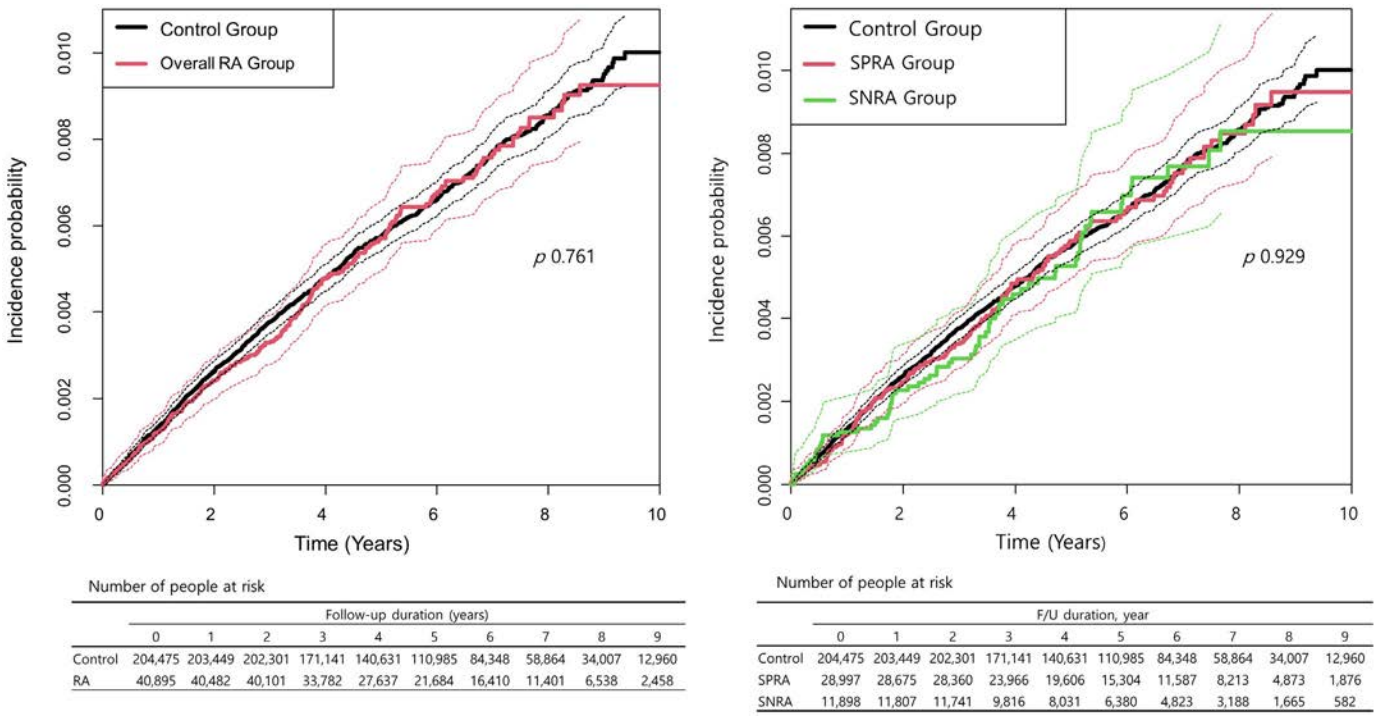


Figure 2. Kaplan-Meier curves showing the incidence probability of thyroid cancer in overall patients with RA (left) and in patients with SPRA and SNRA (right) compared to the control group. F/U, follow-up; RA, rheumatoid arthritis; SNRA, seronegative rheumatoid arthritis; SPRA, seropositive rheumatoid arthritis.

Table 2. HRs and 95% CIs for thyroid cancer incidence in patients with RA*

	N	Event, n	Follow-up duration, person-years	IR per 1,000	Model 1, HR (95% CI)	Model 2, HR (95% CI)	Model 3, HR (95% CI)
Overall							
Control	204,475	1,263	1,122,390	1.13	1 (Reference)	1 (Reference)	1 (Reference)
RA	40,895	244	221,210	1.10	0.98 (0.85–1.12)	0.98 (0.86–1.13)	0.98 (0.85–1.13)
Serologic status of RA							
Control	204,475	1,263	1,122,390	1.13	1 (Reference)	1 (Reference)	1 (Reference)
SNRA	11,898	69	63,980	1.08	0.96 (0.75–1.22)	0.97 (0.76–1.23)	0.96 (0.75–1.22)
SPRA	28,997	175	157,230	1.11	0.99 (0.84–1.16)	0.99 (0.85–1.16)	0.99 (0.84–1.16)
SNRA vs. SPRA							
SNRA	11,898	69	63,980	1.08	1 (Reference)	1 (Reference)	1 (Reference)
SPRA	28,997	175	157,230	1.11	1.04 (0.78–1.37)	1.03 (0.78–1.36)	1.03 (0.78–1.36)

* Model 1: unadjusted; model 2: adjusted for age, sex, smoking, alcohol drinking, regular physical activity, income level, body mass index, and diabetes; model 3: adjusted for variables in model 2, hypertension, dyslipidemia, and chronic kidney disease. CI, confidence interval; HR, hazard ratio; IR, incidence rate; RA, rheumatoid arthritis; SNRA, seronegative rheumatoid arthritis; SPRA, seropositive rheumatoid arthritis.

non-RA group, the aHR of thyroid cancer in patients with overall RA was 0.98 (95% CI 0.85–1.13 in model 3) (Table 2). Results using the stratified Cox model were consistent with those using the main analysis model (Supplementary Table S4). According to serologic status of RA, the aHRs of thyroid cancer were 0.99 (95% CI 0.84–1.16 in model 3) for patients with SPRA and 0.96 (95% CI 0.75–1.22 in model 3) for patients with SNRA compared with the matched non-RA group (Table 2). Compared with the patients with SNRA, the patients with SPRA did not show a significantly different risk of thyroid cancer (aHR 1.03, 95% CI 0.78–1.36 in model 3) (Table 2). Results from competing risk analyses also showed consistent results from main analyses (Supplementary Table S5).

Subgroup analyses. In the subgroup analysis stratified by sex, compared with the matched non-RA group, the aHRs of thyroid cancer were 0.95 (95% CI 0.82–1.10 in model 3) in female patients with RA and 1.30 (95% CI 0.85–2.00 in model 3) in male patients with RA (*P* for interaction = 0.179) (Table 3). Analyses stratified by age group (40–59 years vs ≥60 years) showed no statistical interaction (*P* for interaction = 0.129), and the aHRs were 1.05 (95% CI 0.89–1.23 in model 3) in the younger group and 0.820 (95% CI 0.62–1.08 in model 3) in the older group (Table 3).

DISSCUSSION

In this nationwide cohort study, the risk of thyroid cancer in patients with newly diagnosed RA was not significantly different from that of the matched non-RA group regardless of RA serologic status. Compared with the patients with SNRA, the risk of thyroid cancer was not increased in the patients with SPRA. In the subgroup analyses stratified by sex and age, a significant difference in a risk of thyroid cancer was not observed in patients with RA compared with the counterpart control population.

Table 3. Associations between thyroid cancer and RA according to sex and age*

	N	Event, n	Follow-up duration, (person-years)	IR per 1,000	Model 1, HR (95% CI)	*P* for interaction	Model 2, HR (95% CI)	*P* for interaction	Model 3, HR (95% CI)	*P* for interaction
Sex										
Male										
Control	51,305	103	274,628	0.38	1 (Reference)	0.171	1 (Reference)	0.180	1 (Reference)	0.179
RA	10,261	26	53,158	0.49	1.30 (0.85–2.00)		1.30 (0.85–2.01)		1.30 (0.85–2.00)	
Female										
Control	153,170	1,160	847,762	1.37	1 (Reference)		1 (Reference)		1 (Reference)	
RA	30,634	218	168,053	1.30	0.95 (0.82–1.10)		0.96 (0.83–1.11)		0.95 (0.82–1.10)	
Age										
40–59 y										
Control	119,095	892	664,268	1.34	1 (Reference)	0.140	1 (Reference)	0.128	1 (Reference)	0.129
RA	23,819	185	132,502	1.40	1.04 (0.89–1.22)		1.05 (0.90–1.23)		1.05 (0.89–1.23)	
≥60 y										
Control	85,380	371	458,122	0.81	1 (Reference)		1 (Reference)		1 (Reference)	
RA	17,076	59	88,708	0.67	0.82 (0.62–1.08)		0.82 (0.62–1.08)		0.82 (0.62–1.08)	

* Model 1: unadjusted; model 2: adjusted for age, sex, smoking, alcohol drinking, regular physical activity, income level, body mass index, and diabetes; model 3: adjusted for variables in model 2, hypertension, dyslipidemia, and chronic kidney disease. CI, confidence interval; HR, hazard ratio; IR, incidence rate; RA, rheumatoid arthritis.

To the best of our knowledge, this study is the first to evaluate the risk of thyroid cancer in patients with RA with concomitant consideration of RA serostatus. Other strengths of our study include the following: (1) the nationwide cohort in a region with high incidence of thyroid cancer,[38] which enhances statistical power; (2) the use of a nationwide database that provides reliable prescription records from a compulsory, single national insurance system and various medical information from the national health examination; and (3) the application of operational definitions for identification of patients with newly diagnosed RA using both RA-related disease codes and prescription records to minimize overestimation of RA cases.

Autoimmunity and chronic inflammation are cardinal features of RA.[39] Chronic inflammation has also been accepted as a main pathophysiology of cancer.[3–5] In autoimmune diseases, the proinflammatory environment may support the initiation and growth of precancerous lesions, leading to an increased risk of cancer in patients with autoimmune diseases.[40] The relationship between RA and diseases of the thyroid gland has raised interest because of the autoimmune pathogenesis that underlies both RA and AITD, where the bidirectional association between AITD and RA was reported.[6,8] Although the exact mechanism is not clearly understood, genetic susceptibility and defects in immune tolerance due to autoimmune conditions may play roles in the co-occurrence of AITD and RA.[6,41] Despite extensive cancer-related research in patients with RA, however, there remained a paucity of studies on the association between RA and thyroid cancer.

Based on previous findings, we hypothesized that the risk of thyroid cancer might be increased in patients with RA. In addition, albeit controversial, an association between seropositivity and disease activity has been reported,[42,43] and the presence of RF and/or ACPA have been suggested as poor prognostic factors in patients with RA.[44,45] A previous study reported that the increased risk of lung cancer was significantly associated with seropositivity and also differed by subtypes of SPRA.[46] Thus, we expected that the risk of thyroid cancer might vary by serologic status due to difference in disease severity and autoimmunity or by sex due to higher prevalence of AITD in female patients with RA. Compared with the non-RA group, however, the risk of thyroid cancer was not increased in patients with RA regardless of serologic status or sex in our study.

Through a literature review, inconsistent results regarding association between RA and thyroid cancer have been found, even across studies using the same national database in Korea and in Taiwan (Supplementary Table S1). These inconsistent results may be explained by several reasons: (1) the RA population could be misclassified if defined using diagnostic codes alone or by short periods of DMARD prescription[12,13]; (2) the number of cancer cases could be overestimated when only using diagnostic codes because cancer-related codes can be up-coded as rule-out codes for the cancer screening test, especially during the thyroid cancer "epidemic" in Korea[47,48]; and (3) some studies have been limited by insufficient consideration of key confounding factors or by a small number of study participants.[13–16] Furthermore, none of the previous studies considered serologic status of RA. In our study, such biases could be reduced because of the strict operational definitions for both RA[25] and cancer cases[27] and consideration of various confounding factors using a large nationwide database.

An interesting finding in our study was that the risk of thyroid cancer was not increased in patients with SPRA compared with the non-RA group or even patients with SNRA. Previous studies have reported a significant association between AITD and thyroid cancer,[10–12,49] but a risk of bias and a high level of heterogeneity among previous studies existed.[49] The lack of a significant association between RA, especially SPRA, and thyroid cancer in our study may indirectly suggest that autoimmunity is not a main pathophysiology of thyroid cancer in patients with RA.

There are limitations of our study. First, because we used secondary data from the administrative database provided by the NHIS, we were not able to consider clinical factors, such as RA severity and the stage and pathology of thyroid cancer, which may affect cancer risks in patients with RA. Second, the use of immunosuppressive medication in patients with RA was not considered in this study. Third, the comorbidities of benign thyroid diseases, including thyroid nodules, goiter, hyperthyroidism, and hypothyroidism, were not considered in this study. Although the causal relation between these benign thyroid diseases and thyroid cancer is still unclear, their association has been reported in previous epidemiology studies.[11,47] Lastly, there is a possibility of overdiagnosis of thyroid cancer due to extensive screening with ultrasonography in Korea, which may have an effect on the association between RA and thyroid cancer. However, because the likelihood of overdiagnosis of thyroid cancer increases with greater access to health care[47] and patients with RA generally visit hospitals on a regular basis, leading to greater opportunities for cancer surveillance, the null association between RA and thyroid cancer in our study could be further supported by the higher detection bias in patients with RA.

In conclusion, in this nationwide cohort study, patients with newly diagnosed RA did not show a different risk of thyroid cancer compared to a matched control group. The risk of thyroid cancer incidence was not affected by serologic status of RA, sex, or age. Further studies that consider effects of immunosuppressive RA medication and comorbidity of thyroid diseases in patients with RA are needed to clarify risk of thyroid cancer in patients with RA.

ACKNOWLEDGMENT

This study used data from the Korean NHIS.

AUTHOR CONTRIBUTIONS

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

ROLE OF THE STUDY SPONSOR

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

REFERENCES

1. Widdifield J, Paterson JM, Huang A, et al. Causes of death in rheumatoid arthritis: how do they compare to the general population? Arthritis Care Res (Hoboken) 2018;70(12):1748–1755.
2. Lee YK, Bae SC. Mortality in Korean patients with rheumatoid arthritis. J Rheum Dis 2021;28(3):113–118.
3. Simon TA, Thompson A, Gandhi KK, et al. Incidence of malignancy in adult patients with rheumatoid arthritis: a meta-analysis. Arthritis Res Ther 2015;17(1):212.
4. Turesson C, Matteson EL. Malignancy as a comorbidity in rheumatic diseases. Rheumatology (Oxford) 2013;52(1):5–14.
5. Smitten AL, Simon TA, Hochberg MC, et al. A meta-analysis of the incidence of malignancy in adult patients with rheumatoid arthritis. Arthritis Res Ther 2008;10(2):R45.
6. Conigliaro P, D'Antonio A, Pinto S, et al. Autoimmune thyroid disorders and rheumatoid arthritis: a bidirectional interplay. Autoimmun Rev 2020;19(6):102529.
7. Taylor PN, Albrecht D, Scholz A, et al. Global epidemiology of hyperthyroidism and hypothyroidism. Nat Rev Endocrinol 2018;14(5):301–316.
8. Weetman AP. Diseases associated with thyroid autoimmunity: explanations for the expanding spectrum. Clin Endocrinol (Oxf) 2011;74(4):411–418.
9. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021;71(3):209–249.
10. Silva de Morais N, Stuart J, Guan H, et al. The impact of Hashimoto thyroiditis on thyroid nodule cytology and risk of thyroid cancer. J Endocr Soc 2019;3(4):791–800.
11. Dias Lopes NM, Mendonça Lens HH, Armani A, et al. Thyroid cancer and thyroid autoimmune disease: A review of molecular aspects and clinical outcomes. Pathol Res Pract 2020;216(9):153098.
12. Park SK, Ryoo JH, Kim MH, et al. Association between eight autoimmune diseases and thyroid cancer: a Nationwide Cohort Study. Thyroid 2024;34(2):206–214.
13. Chen YJ, Chang YT, Wang CB, et al. The risk of cancer in patients with rheumatoid arthritis: a nationwide cohort study in Taiwan. Arthritis Rheum 2011;63(2):352–358.
14. Yamada T, Nakajima A, Inoue E, et al. Incidence of malignancy in Japanese patients with rheumatoid arthritis. Rheumatol Int 2011;31(11):1487–1492.
15. Yu KH, Kuo CF, Huang LH, et al. Cancer risk in patients with inflammatory systemic autoimmune rheumatic diseases: a nationwide population-based dynamic cohort study in Taiwan. Medicine (Baltimore) 2016;95(18):e3540.
16. Lee H. The risk of malignancy in Korean patients with rheumatoid arthritis. Yonsei Med J 2019;60(2):223–229.
17. Choi HG, Kang HS, Lim H, et al. Potential cancer risk in patients with rheumatoid arthritis: a longitudinal Korean population-based analysis. J Pers Med 2022;12(6):965.
18. Yeo Y, Han K, Shin DW, et al. Changes in smoking, alcohol consumption, and the risk of thyroid cancer: a population-based Korean cohort study. Cancers (Basel) 2021;13(10):2343.
19. Yeo Y, Shin DW, Han K, et al. Smoking, alcohol consumption, and the risk of thyroid cancer: a population-based Korean cohort study of 10 million people. Thyroid 2022;32(4):440–448.
20. Cheol Seong S, Kim YY, Khang YH, et al. Data resource profile: the National Health Information Database of the National Health Insurance Service in South Korea. Int J Epidemiol 2017;46(3):799–800.
21. Shin DW, Cho J, Park JH, et al. National general health screening program in Korea: history, current status, and future direction. Precision and Future Medicine 2022;6(1):9–31.
22. Jeon KH, Han K, Jung J, et al. Rheumatoid arthritis and risk of depression in South Korea. JAMA Netw Open 2024;7(3):e241139.
23. Chung C, Kim H, Han K, et al. Does rheumatoid arthritis increase the risk of COPD?: a nationwide retrospective cohort study. Chest 2024;165(6):1362–1371.
24. Kang J, Eun Y, Jang W, et al. Rheumatoid arthritis and risk of Parkinson disease in Korea. JAMA Neurol 2023;80(6):634–641.
25. Cho SK, Sung YK, Choi CB, et al. Development of an algorithm for identifying rheumatoid arthritis in the Korean National Health Insurance claims database. Rheumatol Int 2013;33(12):2985–2992.
26. Iwagami M, Shinozaki T. Introduction to matching in case-control and cohort studies. Ann Clin Epidemiol 2022;4(2):33–40.
27. Yang MS, Park M, Back JH, et al. Validation of cancer diagnosis based on the National Health Insurance Service database versus the national cancer registry database in Korea. Cancer Res Treat 2022;54(2):352–361.
28. Tan K; WHO Expert Consultation. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. Lancet 2004;363(9403):157–163.
29. Kim KK, Haam JH, Kim BT, et al; Committee of Clinical Practice Guidelines, Korean Society for the Study of Obesity (KSSO). Evaluation and treatment of obesity and its comorbidities: 2022 update of clinical practice guidelines for obesity by the Korean Society for the Study of Obesity. J Obes Metab Syndr 2023;32(1):1–24.
30. Boucai L, Zafereo M, Cabanillas ME. Thyroid cancer: a review. JAMA 2024;331(5):425–435.
31. Bui AQ, Gunathilake M, Lee J, et al. Relationship between physical activity levels and thyroid cancer risk: a prospective cohort study in Korea. Thyroid 2022;32(11):1402–1410.
32. Lee JH, Youn S, Jung S, et al. A national database analysis for factors associated with thyroid cancer occurrence. Sci Rep 2020;10(1):17791.
33. Shin A, Cho S, Jang D, et al. Body mass index and thyroid cancer risk: a pooled analysis of half a million men and women in the Asia cohort consortium. Thyroid 2022;32(3):306–314.
34. Schmid D, Behrens G, Jochem C, et al. Physical activity, diabetes, and risk of thyroid cancer: a systematic review and meta-analysis. Eur J Epidemiol 2013;28(12):945–958.
35. Wong G, Staplin N, Emberson J, et al. Chronic kidney disease and the risk of cancer: an individual patient data meta-analysis of 32,057 participants from six prospective studies. BMC Cancer 2016;16(1):488.

36. Park JH, Choi M, Kim JH, et al. Metabolic syndrome and the risk of thyroid cancer: a nationwide population-based cohort study. Thyroid 2020;30(10):1496–1504.
37. Park JH, Cho HS, Yoon JH. Thyroid cancer in patients with metabolic syndrome or its components: a nationwide population-based cohort study. Cancers (Basel) 2022;14(17):4106.
38. Shank JB, Are C, Wenos CD. Thyroid cancer: global burden and trends. Indian J Surg Oncol 2022;13(1):40–45.
39. Cross M, Smith E, Hoy D, et al. The global burden of rheumatoid arthritis: estimates from the global burden of disease 2010 study. Ann Rheum Dis 2014;73(7):1316–1322.
40. Elkoshi Z. Cancer and autoimmune diseases: a tale of two immunological opposites? Front Immunol 2022;13:821598.
41. Lazúrová I, Jochmanová I, Benhatchi K, Sotak S. Autoimmune thyroid disease and rheumatoid arthritis: relationship and the role of genetics. Immunol Res 2014;60(2–3):193–200.
42. Aletaha D, Alasti F, Smolen JS. Rheumatoid factor, not antibodies against citrullinated proteins, is associated with baseline disease activity in rheumatoid arthritis clinical trials. Arthritis Res Ther 2015; 17(1):229.
43. Rönnelid J, Wick MC, Lampa J, et al. Longitudinal analysis of citrullinated protein/peptide antibodies (anti-CP) during 5 year follow up in early rheumatoid arthritis: anti-CP status predicts worse disease activity and greater radiological progression. Ann Rheum Dis 2005; 64(12):1744–1749.
44. Smolen JS, Landewé RBM, Bijlsma JWJ, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2019 update. Ann Rheum Dis 2020;79(6):685–699.
45. Alemao E, Bao Y, Weinblatt ME, et al. Association of seropositivity and mortality in rheumatoid arthritis and the impact of treatment with disease-modifying antirheumatic drugs: results from a real-world study. Arthritis Care Res (Hoboken) 2020;72(2):176–183.
46. Chatzidionysiou K, di Giuseppe D, Soderling J, et al. Risk of lung cancer in rheumatoid arthritis and in relation to autoantibody positivity and smoking. RMD Open 2022;8(2):e002465.
47. Kitahara CM, Schneider AB. Epidemiology of thyroid cancer. Cancer Epidemiol Biomarkers Prev 2022;31(7):1284–1297.
48. Ahn HS, Kim HJ, Welch HG. Korea's thyroid-cancer "epidemic"–screening and overdiagnosis. N Engl J Med 2014;371(19): 1765–1767.
49. Abbasgholizadeh P, Naseri A, Nasiri E, et al. Is Hashimoto thyroiditis associated with increasing risk of thyroid malignancies? A systematic review and meta-analysis. Thyroid Res 2021;14(1):26.

Arthritis Care & Research
Vol. 77, No. 11, November 2025, pp 1368–1373
DOI 10.1002/acr.25579

American College of Rheumatology
Empowering Rheumatology Professionals

BRIEF REPORT

Sleep Matters: Exploring the Link Between Sleep Disturbances and Fatigue in Rheumatoid Arthritis

Natalia V. Chalupczak,[1] Burcu Aydemir,[2] Ariel Isaacs,[3] Lutfiyya N. Muhammad,[2] Jing Song,[2] Kathryn J. Reid,[2] Daniela Grimaldi,[2] Mary Carns,[2] Kathleen Dennis-Aren,[2] Dorothy D. Dunlop,[2] Beth I. Wallace,[4] Phyllis C. Zee,[2] and Yvonne C. Lee[2]

Objective. Fatigue is a prevalent and debilitating symptom for patients with rheumatoid arthritis (RA). Although patients and rheumatologists often attribute fatigue to inflammation, other factors such as sleep disturbances are frequently overlooked. This study aims to explore the relationship between subjective (self-reported) and objective (actigraphy based) sleep parameters and self-reported fatigue in patients with RA.

Methods. This cross-sectional analysis included data from 48 adult patients with RA from a single academic rheumatology practice. Sleep data were obtained daily over 14 days with actigraphy (objective) and the Karolinska Sleep Diary (subjective). Fatigue was assessed using the Patient-Reported Outcome Measurement Information System (PROMIS) fatigue computerized adaptive test. Spearman's correlations and linear regression analyses were used to examine associations between sleep parameters and fatigue, adjusting for swollen joint count, pain intensity, and symptoms of depression.

Results. Subjective sleep parameters showed significant correlations with PROMIS fatigue. Longer total sleep time ($\rho = -0.4$, $P < 0.01$), higher sleep efficiency ($\rho = -0.42$, $P < 0.01$), and better sleep quality ($\rho = -0.5$, $P < 0.01$) were associated with lower levels of fatigue. Objective actigraphy-based sleep parameters were not significantly associated with PROMIS fatigue. Separate linear regression models demonstrated that each subjective sleep parameter remained significantly associated with fatigue after adjusting for covariates.

Conclusion. Self-reported poor sleep duration, efficiency, and quality were significantly associated with fatigue in patients with RA, whereas objective actigraphy-based sleep parameters were not, supporting the integration of self-reported assessment of sleep disturbances into RA treatment plans to improve patient outcomes.

INTRODUCTION

Over 50% of patients with rheumatoid arthritis (RA) report fatigue, often on a daily basis.[1] Despite fatigue being one of the most commonly reported and psychologically distressing symptoms experienced by patients with RA, it is rarely directly targeted for treatment.[2] Patients with RA and rheumatologists often attribute fatigue to disease activity and inflammation and rely on standard-of-care therapy with disease-modifying antirheumatic drugs (DMARDs) to manage fatigue. However, several studies have reported that the association between fatigue and systemic markers of inflammation, such as erythrocyte sedimentation rate and C-reactive protein, are modest at best,[3] and although treatment with DMARDs improves fatigue, the magnitude of effect is small.[4]

This focus on RA-related factors has often contributed to the oversight of other important contributors, such as sleep disturbances. Dar et al found that patients with RA predominantly

Supported by an Aspire grant from Pfizer. Dr Aydemir's work was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS), NIH (grants T32-AR-007611 and K12-TR-005104). Dr Muhammad, Ms Song, and Dr Dunlop's work was also supported by the NIAMS, NIH (grant P30-AR-072579). Dr Wallace's work was supported by the VA Office of Research and Development (CX-002430). Dr Lee's work was supported by the NIAMS, NIH (grants K24-AR-080840 and P30-AR-072579).

[1]Natalia V. Chalupczak, BS: Chicago Medical School, Rosalind Franklin University, North Chicago, Illinois; [2]Burcu Aydemir, PhD; Lutfiyya N. Muhammad, PhD, MPH, Jing Song, MS, Kathryn J. Reid, PhD, Daniela Grimaldi, MD, PhD, Mary Carns, MS, Kathleen Dennis-Aren, MPH, Dorothy D. Dunlop, PhD, Phyllis C. Zee, MD, PhD, Yvonne C. Lee, MD, MMSc: Northwestern University Feinberg School of Medicine, Chicago, Illinois; [3]Ariel Isaacs, BA: University of Missouri School of Medicine, Columbia; [4]Beth I. Wallace, MD, MSc: University of Michigan and VA Ann Arbor Healthcare System, Ann Arbor.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25579.

Address correspondence via email to Yvonne C. Lee, MD, MMSc, at yvonne.lee@northwestern.edu.

Submitted for publication October 25, 2024; accepted in revised form May 21, 2025.

SIGNIFICANCE & INNOVATIONS

- Use of daily sleep diaries: Most studies of rheumatoid arthritis (RA) have used retrospective assessments of sleep quality over the past week or month. By employing daily sleep diaries over a 14-day period, this research provides a more nuanced assessment of the relationship between specific self-reported sleep parameters and fatigue.
- Independent impact of sleep on fatigue: This study demonstrates that self-reported measures of poor sleep (short duration, low efficiency, and poor quality) are independently associated with fatigue in patients with RA, even after adjusting for inflammation, pain, and symptoms of depression. This finding challenges the traditional view that fatigue in RA is primarily driven by inflammatory disease activity and pain.
- Targeting sleep in RA management: Our results highlight the importance of routinely assessing and addressing self-reported sleep disturbances as part of comprehensive RA treatment plans, as interventions aimed at improving self-reported sleep could help alleviate fatigue and enhance overall patient well-being.
- Foundation for future research: This study showcases the need for future longitudinal research to establish causal relationships between sleep disturbances and fatigue in RA and develop effective sleep-focused interventions for managing fatigue in patients with RA.

attributed fatigue to joint inflammation and pain, with minimal consideration given to sleep quality.[5] Similarly, da Silva et al reported that patients and rheumatologists often prioritize rheumatic disease management over addressing sleep-related issues.[6]

This study explores the relationship between objective (actigraphy) and self-reported sleep parameters and fatigue in patients with RA. We aim to identify modifiable sleep-related factors that could become specific targets for interventions aimed at mitigating fatigue. We hypothesize that impairments in specific sleep parameters (eg, low sleep duration, poor sleep efficiency, impaired wake after sleep onset [WASO], and poor sleep quality) are significantly associated with higher fatigue, even when controlling for RA-related factors such as inflammation and pain.

METHODS

We conducted a cross-sectional analysis of 48 adult patients with RA enrolled in the Sleep, Pain, and Autonomic Function in RA (SPAN-RA) study.[7] All participants were from a single academic rheumatology practice, ≥18 years old, and met the American College of Rheumatology/EULAR 2010 criteria for RA.[8] Because the original SPAN-RA study was designed to assess heart rate variability, the enrollment criteria excluded patients with diagnoses of cardiac arrhythmias and obstructive sleep apnea (OSA), as well as anyone routinely taking as well as anyone routinely taking beta blockers, central acting pain medications, opioids, and sedatives. Ethics approval was obtained from the Institutional Review Board of Northwestern University. All participants provided written informed consent.

Actigraphy-based sleep measures. Actigraphy data were recorded using the Actiwatch Spectrum (Phillips Respironics) and worn on participants' nondominant wrists for 14 days. Data were sampled at 30-second intervals and processed using the manufacturer's proprietary software (Actiware, version 6.0). Main resting intervals were manually scored. Any 24-hour recording period including ≥4 hours of nonwear was excluded. Sleep parameters included sleep duration, sleep efficiency, and WASO. Sleep duration, measured as the total time spent asleep in the main rest interval, was calculated in minutes. Sleep efficiency measured the percentage of time spent asleep while in bed and was calculated as 100% × total sleep time over time in bed. WASO, defined as the time, in minutes, spent awake after the onset of sleep, was calculated as total time of the intervals scored as wake between sleep onset and sleep offset in the sleep interval. Sleep measures were averaged across all valid days in the recording period.

Self-reported sleep measures. Self-reported measures of sleep were collected using the Karolinska Sleep Diary.[9] Participants were instructed to complete this diary every morning for 14 days. Subjective sleep parameters included duration (subjective sleep length in 10-minute intervals), efficiency (derived from sleep length/time spent in bed), WASO (minutes spent awake after falling asleep), and quality. Sleep quality was derived from a summated index (score range 4–20, higher scores indicating better sleep quality) constructed from four items in the Karolinska Sleep Diary log (sleep quality, restless sleep, difficulties falling asleep, and premature [final] awakening).[9] All measures were averaged over the total number of sleep diary days.

Other measures. All individuals completed the Patient-Reported Outcome Measurement Information System (PROMIS) fatigue and depression computerized adaptive tests. PROMIS T scores are standardized metrics derived from item response theory–based assessments. T scores have a mean of 50 (representing the general population) and an SD of 10. Higher T scores indicate more of the concept being measured. For example, PROMIS fatigue T scores above 55 reflect moderate fatigue, and scores above 60 indicate severe fatigue.[10] Pain intensity was assessed via a pain intensity numerical rating scale ranging from 0 to 10. Swollen joint count was determined by a single trained assessor using a standardized 28-joint count.

Statistical analysis. Descriptive statistics were computed to summarize the demographic and clinical characteristics of the

cohort. Spearman's correlations (rho) were performed to determine the strength of correlations between each wrist actigraphy- and diary-based sleep parameter and the outcome of fatigue. Only sleep parameters that demonstrated significant correlations with fatigue were included in subsequent linear regression analyses.

Separate linear regression models were conducted to assess the associations between different sleep parameters and fatigue, with the PROMIS fatigue T score as the outcome. For the purposes of analysis, sleep duration was expressed in 10-minute increments to provide a more interpretable measure of its association with fatigue. Both unadjusted and adjusted models were constructed, with the latter controlling for potential confounders of depression, swollen joint count, and pain intensity.

All data analyses were performed using R (version 4.2.2). The strength and direction of associations were determined using regression coefficients (beta) with 95% confidence intervals (CIs). An alpha level was set at $P < 0.05$ for all analyses. With a sample size of 48, we have 80% power to detect a medium effect size of Cohen's $f^2 = 0.19$ using a multiple regression model with four independent variables at an α level of 0.05. Additionally, we have adequate power (greater than 80%) to detect correlations of 0.42 or greater using a correlation test at an α level of 0.05.

RESULTS

Baseline participant characteristics. The study cohort comprised 48 patients, all with complete sleep and fatigue data (Table 1). Participants were predominantly female (95.8%), White (52.1%), seropositive (70.8%), and using DMARDs (89.6%). Actigraphy-derived sleep characteristics were reported in a previous publication.[7]

Correlations between sleep parameters and fatigue. *Self-reported sleep measures.* Participants who reported longer total sleep time and higher sleep efficiency had significantly lower fatigue ($\rho = -0.4$, $P < 0.01$, and $\rho = -0.42$, $P < 0.01$, respectively; Figure 1B and D). Similarly, a better sleep quality index score was associated with lower fatigue ($\rho = -0.5$, $P < 0.01$; Figure 1C). Remaining self-reported sleep measures, including WASO, onset latency, number of awakenings, and other self-reported sleep log measures, were not significantly associated with fatigue.

Actigraphy-based sleep measures. Actigraphy-based measures of sleep duration ($\rho = -0.08$, $P = 0.608$), sleep efficiency ($\rho = 0.00$, $P = 0.986$), WASO ($\rho = 0.01$, $P = 0.921$), sleep fragmentation ($\rho = 0.09$, $P = 0.564$), number of awakenings ($\rho = 0.1$, $P = 0.491$), and sleep onset latency ($\rho = 0.01$, $P = 0.921$) were not significantly associated with PROMIS fatigue. Because the actigraphy-based measures were not significantly associated with PROMIS fatigue, they were not included in subsequent analyses.

Table 1. Baseline characteristics of study participants, N = 48*

	Value
Demographics	
Age, y	55.0 (41.5–64.8)
BMI	26.6 (22.4–32.8)
Female, %	95.8
Race and ethnicity, %	
White	52.1
Asian	6.3
Black or African American	12.5
Hispanic or Latino	25.0
Other	4.2
Clinical factors	
Clinical Disease Activity Index	5.0 (2.0–12.0)
Swollen joint count (0–28)	0.0 (0.0–2.0)
Tender joint count (0–28)	1.0 (0.0–4.2)
Patient global (patient global assessment of disease activity) (0–10)	2.0 (1.0–4.0)
Assessor global (assessor global assessment of disease activity) (0–10)	1.0 (0.0–2.0)
Pain intensity NRS (0–10)	3.0 (1.0–5.0)
PROMIS depression	47.2 (39.6–51.8)
PROMIS fatigue	48.6 (42.5–57.1)
Steroid use, %	12.5
RA duration, y	9.8 (4.0–16.2)
Seropositivity, %	70.8
Any DMARD use, %	89.6
Conventional DMARD, %	62.5
Biologics and targeted synthetic DMARD, %	54.2
Sleep diary measures	
Sleep duration, min	427.4 (402.1–454.1)
Sleep efficiency, %	89.9 (86.0–94.0)
Wake after sleep onset, min	11.9 (6.5–30.1)
Sleep quality index	14.1 (12.4–16.3)

* Values are median (Q1–Q3), unless noted otherwise. BMI, body mass index; DMARD, disease-modifying antirheumatic drug; NRS, numeric rating scale; PROMIS, Patient-Reported Outcome Measurement Information System; RA, rheumatoid arthritis.

Linear regression analysis of fatigue on sleep diary measures. Shorter self-reported sleep duration (β = –0.66, 95% CI –1.2 to –0.09), lower sleep efficiency (β = –0.67, 95% CI –1.10 to –0.25), and worse sleep quality (β = –1.50, 95% CI –2.40 to –0.62) were significantly associated with worse fatigue in unadjusted models. The association between WASO and fatigue was not statistically significant (β = 0.06, 95% CI –0.09 to 0.21). After adjustment for symptoms of depression, swollen joint count, and pain intensity, significant associations between self-reported sleep duration (β = –0.50, 95% CI –0.94 to –0.07), sleep efficiency (β = –0.41, 95% CI –0.76 to –0.05), sleep quality (β = –0.81, 95% CI –1.60 to –0.05), and fatigue persisted, although these associations were attenuated (Table 2). The association between WASO and fatigue was not statistically significant after covariate adjustment (β = 0.02, 95% CI –0.10 to 0.13).

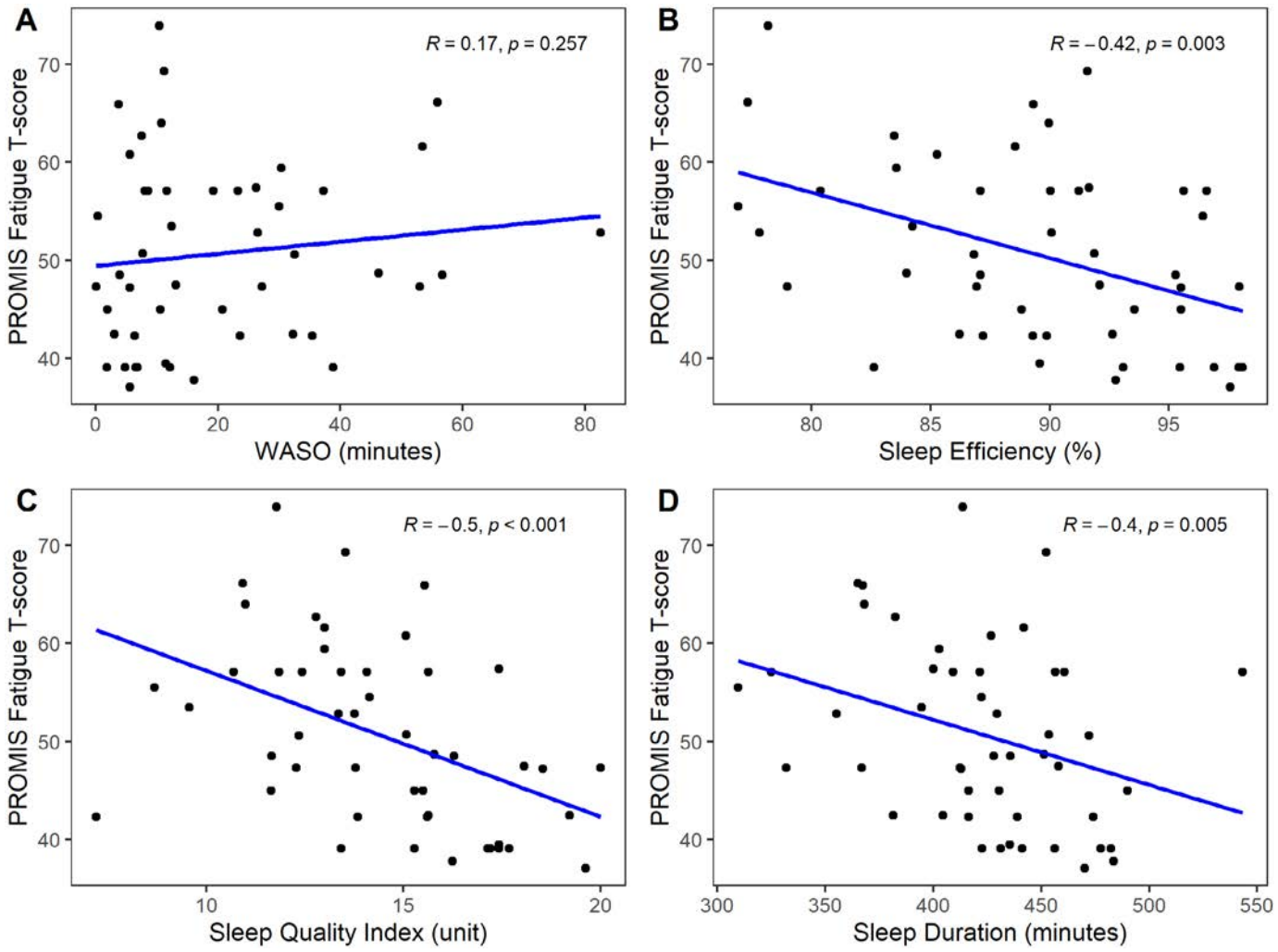


Figure 1. Scatter plots illustrating the relationship between self-reported sleep parameters and PROMIS fatigue. Each plot shows the correlation between PROMIS fatigue and (A) WASO, (B) sleep efficiency, (C) sleep quality index, and (D) sleep duration, based on Spearman's correlation coefficients. PROMIS, Patient-Reported Outcome Measurement Information System; WASO, wake after sleep onset.

DISCUSSION

In this study of patients with established RA, longer self-reported sleep duration, higher self-reported sleep efficiency, and better self-reported sleep quality were associated with lower levels of self-reported fatigue. These relationships persisted after adjusting for joint inflammation, pain, and depressive symptoms. These findings suggest that self-reported poor sleep contributes independently to the experience of fatigue among patients with RA, challenging the commonly held view that fatigue is mainly driven by inflammatory disease activity or pain in this population.[2]

Our study builds on previous studies examining the association between sleep quality and fatigue. One of the largest studies (N = 158) was a cross-sectional analysis by Katz et al, which showed that poor sleep quality (assessed by the Pittsburgh Sleep Quality Index), inactivity, depression, and obesity were all associated with fatigue.[11] Our study extends these results by providing information on specific self-reported sleep parameters (eg, sleep duration and sleep efficiency) through daily sleep diaries. This more nuanced view of self-reported sleep disturbances enabled us to identify sleep duration and sleep efficiency (in addition to the more nebulous concept of sleep quality) as specific targets for interventions to improve fatigue.

Our study is consistent with a mobile health study performed in 254 patients with self-reported RA recruited from the National Rheumatoid Arthritis Society, a patient organization in the United Kingdom.[12] In this study, McBeth et al reported significant associations between both subjective and objective sleep parameters and health-related quality of life. Although the authors did not specifically report associations between sleep parameters and fatigue, the association between sleep and health-related quality of life diminished when fatigue was included in the models, implicating fatigue as a potential mediating factor. Furthermore,

Table 2. Linear regression analysis of PROMIS fatigue on sleep diary measures, N = 48*

Characteristic	Unadjusted			Adjusted		
	β	95% CI	*P* value	β	95% CI	*P* value
Main predictor: sleep duration						
Sleep duration, 10-min intervals	−0.66	−1.20 to −0.09	0.024	−0.50	−0.94 to −0.07	0.025
Swollen joint count	–	–	–	0.72	−0.33 to 1.8	0.2
Pain intensity numeric rating scale	–	–	–	1.0	0.27 to 1.8	0.008
PROMIS depression T score	–	–	–	0.42	0.18 to 0.66	<0.001
Main predictor: sleep efficiency						
Sleep efficiency	−0.67	−1.1 to −0.25	0.002	−0.41	−0.76 to −0.05	0.028
Swollen joint count	–	–	–	0.33	−0.78 to 1.4	0.5
Pain intensity numeric rating scale	–	–	–	0.93	0.18 to 1.7	0.016
PROMIS depression T score	–	–	–	0.43	0.19 to 0.67	<0.001
Main predictor: WASO						
WASO, min	0.06	−0.09 to 0.21	0.4	0.02	−0.10 to 0.13	0.7
Swollen joint count	–	–	–	0.70	−0.43 to 1.8	0.2
Pain intensity numeric rating scale	–	–	–	1.0	0.22 to 1.8	0.013
PROMIS depression T score	–	–	–	0.45	0.20 to 0.71	<0.001
Main predictor: sleep quality						
Sleep quality index	−1.5	−2.4 to −0.62	0.001	−0.81	−1.6 to −0.05	0.038
Swollen joint count	–	–	–	0.72	−0.35 to 1.8	0.2
Pain intensity numeric rating scale	–	–	–	0.72	−0.06 to 1.5	0.071
PROMIS depression T score	–	–	–	0.43	0.19 to 0.67	<0.001

* CI, confidence interval; PROMIS, Patient-Reported Outcome Measurement Information System; WASO, wake after sleep onset.

the importance of sleep duration was highlighted by an interventional in-laboratory study, in which 27 patients with RA underwent partial night sleep deprivation. The morning following partial night sleep deprivation, patients with RA reported significant increases in fatigue and other related symptoms, such as pain and self-reported assessments of disease activity.[13] Together, these findings suggest that interventions aimed at increasing total sleep duration—and possibly enhancing sleep efficiency—may be effective for improving fatigue and other health-related outcomes.

The associations between self-reported sleep duration, sleep efficiency, sleep quality, and the outcome of fatigue remained significant, even after adjustment for other hypothesized predictors of fatigue, including joint inflammation, pain, and symptoms of depression. Pain and symptoms of depression were also significantly associated with fatigue, but joint inflammation, measured by swollen joint count, was not. These findings are consistent with other studies, suggesting that despite the focus on inflammation as a cause for fatigue, the experience of fatigue is likely rooted in multiple causes, which may differ from individual to individual. In a population of patients with long-standing treated RA, inflammation may not be the primary contributor. Routine assessments of self-reported sleep, pain, and depression should be incorporated in a comprehensive approach to identify and target possible sources of fatigue in patients with RA.

Contrary to the observation that self-reported measures of sleep were associated with fatigue, actigraphy-based measures of sleep were not. This discrepancy raises an important question: Is it the actual sleep, as measured objectively, that is most relevant to fatigue, or is it the perception of sleep that matters more? Prior studies suggest that self-reported sleep quality may be a stronger predictor of fatigue than actigraphy-derived sleep measures because it captures an individual's experience of restfulness, sleep satisfaction, and nighttime discomfort, which actigraphy cannot fully assess.[14] Additionally, self-reported sleep disturbances may reflect broader psychosocial factors, including stress, pain perception, and cognitive biases, that influence fatigue perception.[15] It is common for self-reported measures to be more strongly associated with other self-reported measures than with objectively measured data.[16] This may partly be due to bias in the way individuals respond to questionnaires. Another potential reason for the discrepancy is that self-reported and actigraphy-based measures assess different aspects of sleep. Actigraphy primarily captures movement-based parameters, whereas self-reports reflect an individual's perception of sleep quality and disturbances. Both are imperfect assessments, but both can provide valuable information. Our findings highlight the importance of addressing patients' perceptions of sleep disturbances in clinical management rather than relying solely on objective sleep assessments.

Strengths of this study include the comprehensive assessment of sleep characteristics through daily sleep diary assessments, as well as the consideration of potential confounders. Limitations include the cross-sectional design, which restricts the ability to infer causality. Additionally, the relatively healthy study population with low disease activity and mild sleep disturbances may not be representative of all patients with RA, potentially limiting the generalizability of these findings. The original SPAN-RA study enrollment criteria excluded patients with

self-reported OSA. As OSA may not always be clinically diagnosed or captured via actigraphy, its potential impact on sleep quality and fatigue cannot be entirely ruled out. Future research should aim to include a broader range of disease activity levels and sleep disturbances to better understand these relationships. We also plan to evaluate the relationship of physical activity, assessed by actigraphy, and fatigue in this population.

This study highlights the role of self-reported sleep disturbances in the experience of fatigue among patients with RA. Despite the traditional focus on inflammation and pain, our findings suggest that poor sleep is a significant and independent contributor to fatigue. Notably, our sample was biased toward individuals without certain sleep disorders and not taking specific medications. Associations between sleep disturbances and fatigue may be even more pronounced in patients with existing sleep disturbances. Rheumatologists should recognize the importance of assessing and managing sleep disturbances to improve fatigue and overall well-being in patients with RA. Additional longitudinal studies are needed to establish causal relationships and develop effective sleep-focused interventions for this population.

ACKNOWLEDGMENT

Grammarly, an artificial intelligence writing aid, was used to review grammar and syntax.

AUTHOR CONTRIBUTIONS

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

REFERENCES

1. Pope JE. Management of fatigue in rheumatoid arthritis. RMD Open 2020;6(1):e001084.
2. Hewlett S, Cockshott Z, Byron M, et al. Patients' perceptions of fatigue in rheumatoid arthritis: overwhelming, uncontrollable, ignored. Arthritis Rheum 2005;53(5):697–702.
3. Druce KL, Basu N. Predictors of fatigue in rheumatoid arthritis. Rheumatology (Oxford) 2019;58(suppl 5):v29–v34.
4. Wohlfahrt A, Bingham CO III, Marder W, et al. Responsiveness of Patient-Reported Outcomes Measurement Information System Measures in rheumatoid arthritis patients starting or switching a disease-modifying antirheumatic drug. Arthritis Care Res (Hoboken) 2019; 71(4):521–529.
5. Dar WR, Mir IA, Siddiq S, et al. The assessment of fatigue in rheumatoid arthritis patients and its impact on their quality of life. Clin Pract 2022;12(4):591–598.
6. da Silva JAP, Ramiro S, Pedro S, et al. Patients- and physicians- priorities for improvement. The case of rheumatic diseases. Acta Reumatol Port 2010;35(2):192–199.
7. Aydemir B, Muhammad LN, Song J, et al. Characterization of sleep disturbance in established rheumatoid arthritis patients: exploring the relationship with central nervous system pain regulation. BMC Rheumatol 2024;8(1):33.
8. Aletaha D, Neogi T, Silman AJ, et al. Rheumatoid arthritis classification criteria: an american college of rheumatology/european league against rheumatism collaborative initiative. Arthritis Rheum 2010;62: 2569–2581.
9. Akerstedt T, Hume K, Minors D, et al. The subjective meaning of good sleep, an intraindividual approach using the Karolinska Sleep Diary. Percept Mot Skills 1994;79(1 Pt 1):287–296.
10. Bingham CO III, Gutierrez AK, Butanis A, et al. PROMIS fatigue short forms are reliable and valid in adults with rheumatoid arthritis. J Patient Rep Outcomes 2019;3(1):14.
11. Katz P, Margaretten M, Trupin L, et al. Role of sleep disturbance, depression, obesity, and physical inactivity in fatigue in rheumatoid arthritis. Arthritis Care Res (Hoboken) 2016;68(1):81–90.
12. McBeth J, Dixon WG, Moore SM, et al. Sleep disturbance and quality of life in rheumatoid arthritis: prospective mHealth study. J Med Internet Res 2022;24(4):e32825.
13. Irwin MR, Olmstead R, Carrillo C, et al. Sleep loss exacerbates fatigue, depression, and pain in rheumatoid arthritis. Sleep 2012;35(4): 537–543.
14. Russell C, Wearden AJ, Fairclough G, et al. Subjective but not actigraphy-defined sleep predicts next-day fatigue in chronic fatigue syndrome: a prospective daily diary study. Sleep 2016;39(4): 937–944.
15. Suarez EC. Self-reported symptoms of sleep disturbance and inflammation, coagulation, insulin resistance and psychosocial distress: evidence for gender disparity. Brain Behav Immun 2008;22(6):960–968.
16. Moore MN, Wallace BI, Song J, et al. Correlation of Fibromyalgia Survey Questionnaire and quantitative sensory testing among patients with active rheumatoid arthritis. J Rheumatol 2022;49(9): 1052–1057.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 77, No. 11, November 2025, pp 1374–1383
DOI 10.1002/acr.25569

Avoidant or Restrictive Food Intake Disorder Symptoms in Adults With Systemic Sclerosis: A Nationwide Study in Spain

Luis G. Alcala-Gonzalez,[1] Helen Burton-Murray,[2] Micaela Atkins,[2] Alfredo Guillen-del-Castillo,[3] Carolina Malagelada,[4] Michael Hughes,[5] Zsuzsanna H. McMahan,[6] and Carmen P. Simeón-Aznar[3]

Objective. Patients with systemic sclerosis (SSc) may restrict food intake to manage their symptoms (particularly gastrointestinal [GI]). Whether some patients may develop nutritional and/or quality-of-life impairments indicative of an eating disorder, avoidant or restrictive food intake disorder (ARFID), is unknown. We aimed to (1) identify the prevalence and characteristics of ARFID symptoms in patients with SSc and (2) explore the relationship among ARFID symptoms, GI symptom burden, and health-related quality of life.

Methods. In a cross-sectional internet survey nationwide in Spain, patients with SSc completed the Spanish Nine-Item ARFID Screen and assessments of gastrointestinal symptom burden (University of California Los Angeles Scleroderma Clinical Trial Consortium GI tract 2.0 [UCLA SCTC GIT 2.0]) and health-related quality of life (12-item Short Form Survey [SF-12]).

Results. Of 200 patients with SSc, 99 patients (49.5%) screened positive for ARFID. Just over half of those who screened positive for ARFID (n = 53) screened positive based on having a fear of aversive consequences around eating (eg, GI discomfort). A positive ARFID screen was associated with a greater frequency of self-reported enteral nutrition, weight loss, and self-initiated (vs provider-monitored) diet restrictions. ARFID symptoms were moderately associated with worse GI symptom severity by UCLA GIT 2.0 total score ($r = 0.408$, $P < 0.001$) but not for the reflux subscale ($r = 0.058$, $P = 0.420$) and constipation subscale ($r = 0.090$, $P = 0.209$) and with worse health-related quality of life in all domains and both the physical and mental components of the SF-12 (all $P < 0.05$).

Conclusion. ARFID symptoms were relatively common in patients with SSc. Future research is needed to identify when a positive screen for ARFID reflects an adaptive response to disease or pathologic restriction indicative of ARFID warranting behavioral treatment.

INTRODUCTION

Systemic sclerosis (SSc) is a complex multisystem autoimmune disease. Key features include systemic vasculopathy and excessive collagen deposition, resulting in fibrosis of the skin and internal organ dysfunction. Pulmonary complications are the leading cause of death in patients with SSc, but gastrointestinal (GI) involvement is the most prevalent manifestation, affecting more than 90% of patients.[1,2] In patients with SSc, the majority of GI symptoms and complications arise as a result of dysmotility affecting the normal function of GI organs.[2,3] Some patients with SSc may naturally attempt to manage their GI symptoms with dietary changes, but there is a lack of clear guidance for dietary management in SSc despite existing evidence that patients with SSc

Dr Alcala-Gonzalez's work was supported by a grant from the neurogastroenterology and motility workgroup of the Asociación Española de Gastroenterologia (recipient March 2023). Dr Burton-Murray's work was supported by the National Institute of Diabetes and Digestive and Kidney Diseases, NIH (grant K23-DK-131334). Dr McMahan's work was supported by the NIH (grant 1-R01-AR-081382-01A1).

[1]Luis G. Alcala-Gonzalez, MD: Digestive System Research Unit, Digestive Diseases Department, Vall d'Hebron University Hospital, Barcelona, Spain; [2]Helen Burton-Murray, PhD: Micaela Atkins, PhD: Massachusetts General Hospital and Harvard Medical School, Boston; [3]Alfredo Guillen-del-Castillo, MD, PhD, Carmen P. Simeón-Aznar, MD: Systemic Autoimmune Diseases Unit, Internal Medicine Department, Vall d'Hebron University Hospital, Barcelona, Spain; [4]Carolina Malagelada, MD: Digestive System Research Unit, Digestive Diseases Department, Vall d'Hebron University Hospital, Centro de Investigación Biomédica en Red Enfermedades Hepáticas y Digestivas, and Department of Medicine, Universitat Autònoma de Barcelona, Barcelona, Spain; [5]Michael Hughes, BSc (Hons), MSc, MBBs, MRCP: Division of Musculoskeletal and Dermatological Sciences, The University of Manchester and Manchester Academic Health Science Centre, Manchester, UK, and Northern Care Alliance NHS Foundation Trust, Salford Care Organisation, Salford, UK; [6]Zsuzsanna H. McMahan, MD, MHS: University of Texas Health Science Center at Houston.

Drs Alcala-Gonzalez and Burton-Murray are co-first authors and contributed equally to this work.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25569).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25569.

Address correspondence via email to Luis G. Alcala-Gonzalez, MD, at luisgerardo.alcala@vallhebron.cat.

Submitted for publication November 7, 2024; accepted in revised form April 29, 2025.

SIGNIFICANCE & INNOVATIONS

- Patients with systemic sclerosis (SSc) and gastrointestinal (GI) involvement often self-initiate restrictive diets without professional guidance, increasing the risk of malnutrition.
- Almost half of those patients with SSc who participated in our study screened positive for a nonbody image–based eating disorder called avoidant or restrictive food intake disorder (ARFID).
- ARFID symptoms were associated with increased GI symptom burden, particularly upper GI symptoms, worse health-related quality of life, and a higher frequency of reliance on nutritional support (eg, enteral nutrition).
- Our study sheds initial light on ARFID symptoms in the context of SSc, but further research is needed including the use of diagnostic interviews for ARFID and prospective designs.

commonly have clinically significant weight loss and/or malnutrition,[4,5] and the risks for disordered eating and nutritional issues is currently unknown.

Clinical guidelines for the management of SSc and expert consensus recommend dietary modifications as part of general lifestyle measures in the treatment of SSc.[6,7] However, literature about specific dietary recommendations is scarce, and there is currently no empirical dietary treatment for disease management, including GI symptoms.[8,9] Despite a lack of empirical evidence for the use of diet modifications in SSc, a recent international survey found that approximately 90% of patients with SSc seek dietary advice for GI symptom management.[10]

Although dietary management is likely to be helpful and adaptive for many patients, self-initiated, unmonitored (restrictive) dietary changes may potentially be harmful for some. In patients with chronic GI disorders, there is increasing concern about the risk of avoidant or restrictive food intake disorder (ARFID), a non-body image–based eating disorder.[11] ARFID is characterized by overall caloric intake reduction and/or food avoidance that leads to significant medical (eg, weight loss, malnutrition) and/or psychosocial impairment. Food restriction or avoidance in ARFID is motivated by one or more of three prototypical presentations: aversion to the sensory characteristics of food, a lack of interest in eating or low appetite, and fear of aversive consequences.[12]

In patients with SSc, GI discomfort may drive dietary changes, but the risk of developing ARFID is not yet known. Poor appetite is a significant predictor of malnutrition in this population.[13] Notably, early satiety in patients with SSc has been associated with unintentional weight loss, higher nutritional risk, and lower caloric intake, with some research showing that patients restrict eating by reporting consuming only one meal per day.[14] As a result, many patients do not meet their nutritional needs through oral intake alone.[15] However, to our knowledge, no study to date has yet investigated ARFID symptoms in individuals with SSc.

In the current study, we performed an exploratory, cross-sectional, nationwide study of adult patients with SSc. We aimed to (1) identify the prevalence and characteristics of dietary restrictions and ARFID symptoms and (2) evaluate the associations of ARFID symptoms with GI symptoms and health-related quality of life (HrQoL). We hypothesized that greater ARFID severity would be associated with greater GI symptom burden and worse HrQoL.

MATERIALS AND METHODS

Study design. Two researchers (LGAG and CPSA) met with patients (n = 10) from the Spanish association of SSc (AEE) in an open discussion to learn about their concerns about GI symptom management. A main theme of discussion was the need for nutritional and dietary counseling. The research group, after identifying a lack of specific nutritional recommendations in patients with SSc, designed a cross-sectional study using validated questionnaires to explore eating symptoms and nutritional characteristics including ARFID and their relationship with GI symptoms and HrQoL. The project was approved by the AEE scientific committee, and the final survey was developed with input from three patients with SSc to ensure it accurately reflected their symptoms and concerns.

Participants. Patients aged ≥18 years with self-reported diagnosed SSc who belonged to the AEE were invited to participate. Participation was voluntarily, and respondents could discontinue at any point.

Procedures and measures. The AEE distributed the survey to its members using internal emails and group chats. Study data were collected and managed at Vall d'Hebron Research Institute with REDCap, a secure, web-based software platform designed to support data capture for research studies.[16] For the purpose of this study, we used three validated questionnaires to assess ARFID symptoms, GI symptoms, and HrQoL (Spanish Nine-Item ARFID Screen [S-NIAS], University of California Los Angeles Scleroderma Clinical Trial Consortium GI tract 2.0 [UCLA GIT 2.0], and 12-item Short Form Survey [SF-12], respectively), along with a patient survey to capture self-reported demographic and clinical characteristics.

ARFID symptoms. The S-NIAS[17] comprises 11 items on a response scale of 1 to 5 (total score range 1–55), with higher scores indicating greater ARFID severity. Three subscale scores align with the three underlying motivations for avoidant or restrictive eating: (1) the picky eating subscale measures sensory aversion to food (three items), (2) the appetite subscale measures

lack of interest in eating or food (four items), and (3) the fear subscale measures fear of aversive consequences of eating (four items). The S-NIAS has been used to evaluate ARFID symptoms in patients with chronic GI diseases, including patients with gastroparesis, functional dyspepsia,[18,19] achalasia, eosinophilic esophagitis,[20] and inflammatory bowel disease.[21,22] See the Supplementary Methods Material for the complete questionnaire. Internal consistency in our sample was high for all subscales as indicated by Cronbach's α 0.791 for picky eating, 0.815 for low appetite, and 0.825 for fear.

GI symptoms. We used the Spanish version of the UCLA SCTC GIT 2.0. In summary, this is a seven multi-item scale that evaluates reflux, distention or bloating, diarrhea, fecal soiling, constipation, emotional well-being, and social functioning, giving single scores ranging from 0 (best) to 3.00 (worst), except for diarrhea (from 0 to 2.00) and constipation (from 0 to 2.50). The total score is obtained by averaging all subscales except constipation and ranges from 0 (best) to 2.83 (worst). The Spanish version of the UCLA SCTC GIT 2.0 has been used to evaluate GI symptoms in Spanish patients with SSc.[23–25]

HrQoL. The SF-12 (Spanish version[26]) comprises 12 items encompassing the following eight domains of HrQoL: physical function, body pain, general health, vitality, mental health, role limitation due to physical health, role limitation due to mental health, and social functioning. The results of the SF-12 can be grouped into two main component summaries: the physical component summary and the mental component summary. Each domain and both summaries' results were standardized to responses from the Spanish general population, with a mean score fixed at 50 and an SD of 10. The Spanish version of the SF-12 has been validated and standardized for use in the Spanish population.[27]

Demographics and clinical characteristics. Participants self-reported the following: clinical characteristics related to SSc including cutaneous subset, age at diagnosis, use of dietary supplementations or nutritional support, self-reported significant weight loss, history of personal consultation(s) with a dietitian, perceived relationship between GI symptoms and food intake, and self-reported food restrictions due to perceived association to GI symptoms. Demographics reported included reporting Spanish autonomous community of residence; no information about race and ethnicity was obtained. See the Supplementary Methods Material for the complete questionnaire.

Statistical analysis. All statistical analyses were performed using JASP Version 0.18.3 (JASP Team). The normality of data distribution was evaluated by the Shapiro-Wilk test. Data were presented as medians and interquartile ranges (IQRs) for nonnormal continuous variables, means and SDs for normal continuous variables, and numbers and percentages for discrete variables. We operationalized a S-NIAS–positive screen by endorsement of a 4 ("agree") or 5 ("strongly agree") on at least two items within at least one subscale. We compared demographic and questionnaire responses among patients who screened positive for ARFID (S-NIAS) and correlated ARFID symptom scores with GI and HrQoL scores. Comparisons of parametric, normally distributed data were made by Student's *t*-test and by the Mann-Whitney U-test for nonparametric data; for categorical variables, we used the chi-square test. Correlations between the NIAS subscales and total score, UCLA SCTC GIT 2.0 subscales and total score, and SF-12 8 domains and two main component summaries were performed using Pearson's r. Correlations between ARFID symptoms (by S-NIAS) and specific GI symptoms (by UCLA SCTC GIT 2.0) were performed using Spearman's rho. To quantify differences between comparisons, we calculated effect sizes using Cohen's d for Student's *t*-tests and Bserial rank correlation for Mann-Whitney U-tests, categorizing the results as small (0.2–0.49), medium (0.5–0.79), or large (≥0.8). For chi-square tests, effect sizes were calculated using odds ratios (ORs). Given the exploratory nature of this study, no adjustments were made for multiple comparisons.

Patients with SSc who belong to the AEE participated in the selection of topics, approved the content of the survey, and participated in distributing for completion between members. All methods were performed in accordance with relevant guidelines and regulations. The study was approved by the Clinical Research Committee from the AEE. Electronic consent to participate in the survey was obtained, and participation was voluntary. All data were managed in a secure REDCap database, and data sets analyzed were anonymized.

The data that support the findings of this study are available on request from the corresponding author on reasonable request.

RESULTS

Of the 506 patients with SSc who belonged to the AEE, 239 patients started the survey, of whom 200 patients completed the S-NIAS and UCLA GIT 2.0 and were included for analysis (Table 1). Participants were mostly middle aged (mean 54.6 [SD ±10.3] years old) and female (89%). The most frequent SSc subset was diffuse SSc (47%), followed by limited cutaneous SSc (45%), with a median disease duration of 12 (IQR 5–6) years. The geographic distribution of the included participants is shown in Figure 1. Participants from 16 Spanish autonomous communities participated in the survey, although most were from Madrid and Cataluña.

Evaluating weight status, 67 patients (35%) were overweight (body mass index [BMI] ≥25), and 13 patients (8%) were underweight (BMI <18.5). Fifty-one participants (26%) reported unintentional weight loss in the previous six months. Forty-four participants (22%) could not eat normal meals, and 55 participants (28%) reported using some kind of nutritional support. Overall, 140 participants (70%) reported restricting or avoiding

Table 1. Demographic and clinical characteristics, GI symptom burden, and health-related quality of life by positive ARFID screen*

	Total sample (N = 200)	S-NIAS screen negative (n = 101)	S-NIAS screen positive (n = 99)	Effect size[a]	95% CI	*P* value
Clinical characteristics						
Female sex, n (%)	177 (89)	87 (87)	90 (92)	0.1	−0.2 to 0.3	0.269
Age, mean ± SD, y	54.6 ± 10.3	54.2 ± 9.9	54.9 ± 10.7	−0.1	−0.4 to 0.2	0.656
BMI, mean ± SD	23.9 ± 4.5	24.1 ± 4.8	23.7 ± 4.1	0.1	−0.2 to 0.4	0.497
Underweight (BMI < 18.5), n (%)	13 (8)	6 (6)	7 (8)	1.2	0.4 to 3.8	0.716
Disease duration,[b] median (IQR), y	12 (5–16)	12(6–16)	12 (4–15)	−0.1	−0.4 to 0.2	0.482
Age at first symptom,[b] median (IQR), y	38 (30–49)	39 (30–47)	38 (29–50)	0.0	−0.2 to 0.2	0.996
Cutaneous subset, n (%)						
Diffuse cutaneous	94 (47)	46 (46)	48 (49)	1.1	0.7 to 1.8	0.771
Limited cutaneous	89 (45)	48 (48)	41 (42)	0.9	0.5 to 1.5	0.618
Sine scleroderma, n (%)	13 (7)	5 (5)	8 (8)	1.6	0.5 to 6.2	0.395
Does not know, n (%)	3 (2)	2 (2)	1 (1)	0.5	0.0 to 5.8	0.591
Smoking status, n (%)						
Never	77 (39)	44 (44)	33 (33)	0.8	0.5 to 1.3	0.322
History	115 (58)	55 (54)	60 (61)	1.1	0.7 to 1.8	0.648
Active status	8 (4)	2 (2)	6 (6)	3.1	0.6 to 15.5	0.177
Raynaud phenomenon, n (%)	189 (95)	97 (97)	92 (94)	0.5	0.1 to 2.0	0.292
Not able to eat normal meals, n (%)	44 (22)	12(12)	32 (32)	3.3	1.5 to 6.9	<0.001
Use of nutritional support, n (%)	55 (28)	18 (18)	37 (37)	2.1	1.1 to 3.9	0.002
Vitamin and mineral supplements, n (%)	34 (17)	13 (13)	21 (21)	1.8	0.9 to 3.9	0.116
Enteral supplements, n (%)	24 (12)	5 (5)	19 (19)	4.6	1.6 to 12.8	0.002
Parenteral nutrition, n (%)	3 (2)	0 (0)	3 (3)	7.4	0.4 to 144.4	0.078
Restriction of foods, n (%)	140 (70)	63 (63)	77 (78)	2.1	1.1 to 3.8	0.022
Self-initiated diet restriction, n (%)	119 (60)	50 (50)	69 (70)	2.3	1.3 to 4.2	0.004
Dietitian advice, n (%)	35 (18)	17 (18)	18 (18)	1.0	0.5 to 2.0	0.904
Medical advice, n (%)	20 (10)	9 (9)	11 (11)	1.3	0.5 to 3.2	0.604
Recent weight loss, n (%)	51 (26)	12 (12)	39 (39)	4.8	2.3 to 10.0	<0.001
Up to 4 kg, n (%)	31 (16)	9 (9)	22 (22)	2.5	1.1 to 5.7	0.028
5–9 kg, n (%)	13 (7)	2 (2)	11 (11)	5.7	1.2 to 26.2	0.026
More than 10 kg, n (%)	7 (4)	1 (1)	6 (6)	6.1	0.7 to 51.8	0.096
GI symptom burden,[c] mean ± SD						
Reflux subscale	0.9 ± 0.7	0.9 ± 0.7	0.9 ± 0.7	0.0	−0.2 to 0.3	0.730
Bloating subscale	1.6 ± 0.9	1.5 ± 0.9	1.8 ± 0.9	0.4	0.2 to 0.7	0.002
Fecal soilage subscale	0.7 ± 1.0	0.6 ± 0.9	0.9 ± 1.1	0.3	0.1 to 0.6	0.019
Diarrhea subscale	0.6 ± 0.7	0.5 ± 0.6	0.8 ± 0.7	0.4	0.1 to 0.7	0.005
GI social functioning subscale	0.7 ± 0.6	0.5 ± 0.5	0.9 ± 0.7	0.7	0.5 to 1.0	<0.001
GI well-being subscale	1.0 ± 0.8	0.7 ± 0.7	1.3 ± 0.9	0.7	0.4 to 1.0	<0.001
Constipation subscale	0.7 ± 0.7	0.6 ± 0.6	0.8 ± 0.7	0.2	−0.1 to 0.5	0.139
UCLA SCTC GIT 2.0 total score	0.9 ± 0.5	0.8 ± 0.5	1.1 ± 0.6	0.7	0.4 to 0.9	<0.001
Health-related quality of life[d]						
Physical component summary,[b] median (IQR)	35.3 (29.9–40.3)	36.1 (30.0–41.7)	33.6 (29.8–37.1)	−0.3	−0.6 to −0.1	0.014
Mental component summary,[b] median (IQR)	40.1 (32.3–48.9)	42.6 (36.7–51.4)	37.7 (26.7–44.3)	−0.4	−0.7 to −0.1	0.003
Physical function, mean ± SD	36.9 ± 12.8	40.3 ± 13.3	33.3 ± 11.3	−0.6	−0.9 to −0.3	<0.001
Role limitation due to physical health, mean ± SD	34.8 ± 10.5	37.6 ± 10.7	31.8 ± 9.3	−0.6	−0.9 to −0.3	<0.001
Body pain	37.3 ± 12.6	33.4 ± 12.6	41.6 ± 11.3	0.7	0.4 to 1.0	<0.001
General health	30.8 ± 9.0	33.9 ± 9.2	27.4 ± 7.4	−0.8	−1.1 to −0.5	<0.001
Vitality	41.9 ± 9.8	44.9 ± 9.6	38.7 ± 9.0	−0.7	−1.0 to −0.4	<0.001
Social functioning	36.1 ± 10.9	38.7 ± 10.7	33.2 ± 10.4	−0.5	−0.8 to −0.2	<0.001
Role limitation due to emotional health, mean ± SD	34.7 ± 12.1	36.6 ± 12.0	32.7 ± 12.0	−0.3	−0.6 to 0.0	0.031
Mental health	42.0 ± 10.4	43.9 ± 9.8	40.0 ± 10.7	−0.4	−0.7 to −0.1	0.012

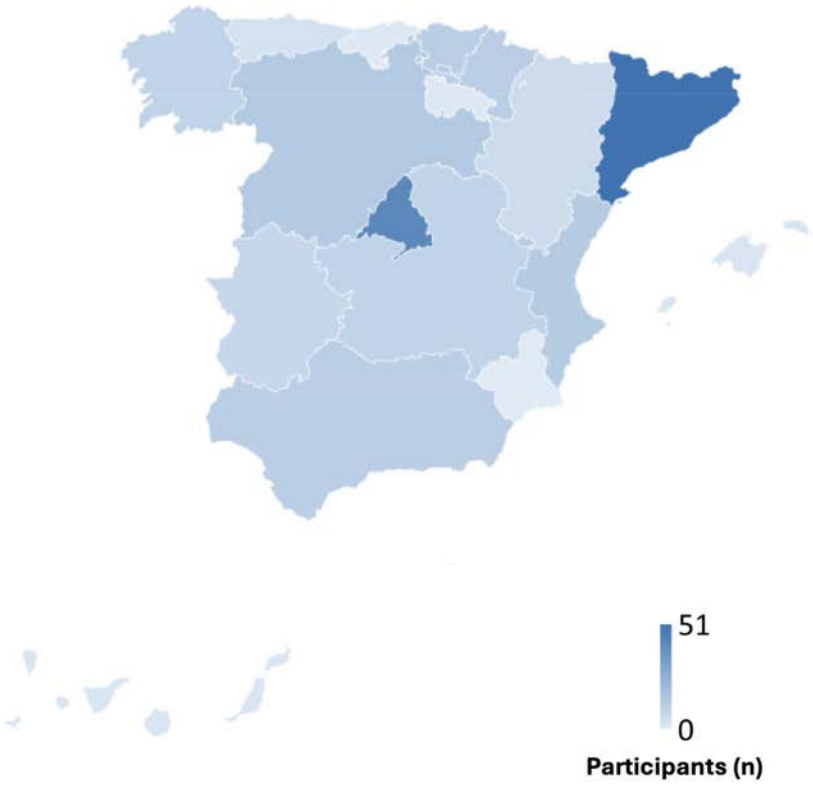
* A positive S-NIAS screen status was determined if participants rated at least two items with a 4 ("Agree") or 5 ("Strongly Agree") within any one of the subscales (fear, picky eating, or appetite). ARFID, avoidant or restrictive food intake disorder; BMI, body mass index; CI, confidence interval; GI, gastrointestinal; IQR, interquartile range; S-NIAS, Spanish Nine-Item Avoidant or Restrictive Food Intake Disorder Screen; UCLA SCTC GIT 2.0, University of California Los Angeles Scleroderma Clinical Trial Consortium GI tract 2.0.

[a] Cohen's d used for *t*-tests, Biserial rank correlation for Mann-Whitney U-tests, and odds ratios for chi-square tests.

[b] Mann-Whitney U-test used instead of *t*-test as the data were not normally distributed.

[c] UCLA GIT 2.0 subscales range from 0 (best) to 3.00 (worst) except for diarrhea, which ranges from 0 (best) to 2.00 (worst), and constipation, which ranges from 0 (best) to 2.50 (worst), UCLA GIT 2.0 total score is obtained from averaging all subscales except constipation and ranges from 0 (best) to 2.83 (worst).

[d] 12-item Short Form Survey scores for the eight domains and for the physical and mental component summaries are standardized. A score below 50 reflects worse health-related quality of life compared to the average of the general Spanish population. Only 181 patients completed the 12-item Short Form Survey, 94 NIAS-positive patients and 87 NIAS-negative patients.



Spanish autonomous communities	Proportion
Andalucía	6%
Aragón	3%
Asturias	2%
Canarias	1%
Cantabria	1%
Castilla-La Mancha	5%
Castilla y León	7%
Cataluña	26%
Extremadura	4%
Galicia	5%
Islas Baleares	1%
La Rioja	1%
Comunidad de Madrid	21%
Región de Murcia	0%
Navarra	6%
País Vasco	6%
Comunidad Valenciana	7%

Figure 1. Geographic distribution of residence across Spain of patients (N = 200) with systemic sclerosis. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25569/abstract.

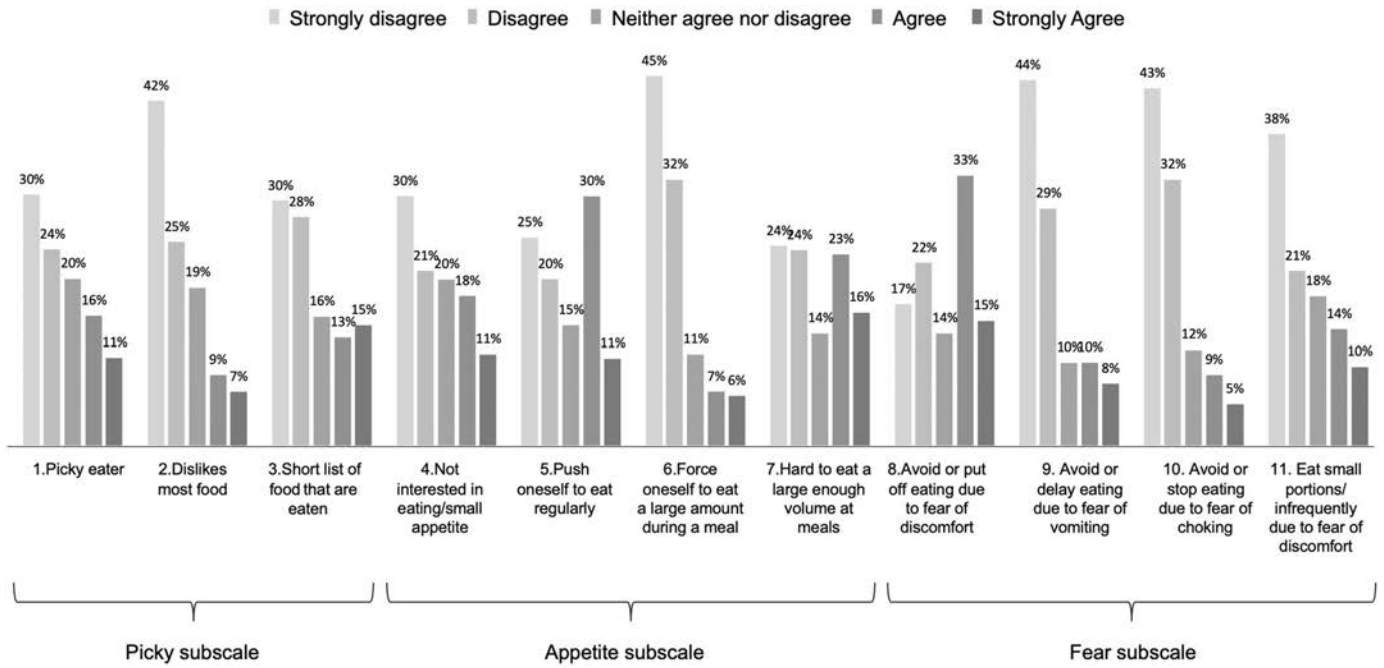


Figure 2. Score distribution of avoidant or restrictive food intake disorder (ARFID) symptom survey items (N = 200). The 11 items of the Spanish-Nine-Item ARFID Screen are displayed in panel A. Each bar represents the percentage of patients who chose a response for the indicated items across the three subscales. We operationalized a positive screen on a subscale as selecting "agree" or "strongly agree" for at least two items within the subscale.

certain food types, and most participants (60%) reported self-initiating diet restrictions, with only 18% of participants doing so after counseling with a dietitian.

Among the cohort, 99 participants (50%) screened positive for one or more S-NIAS subscales. Figure 2 illustrates the distribution of individual scores for ARFID symptoms as assessed by the S-NIAS, and Figure 3 depicts the proportion of participants with positive screen results. Notably, positive responses on the S-NIAS (defined as "agree" or "totally agree" responses) were most common for the item assessing food avoidance due to fear of discomfort (48%), whereas the appetite subscale had the highest frequency of positive screens, with 72 participants (36%) screening positive. Compared to those who screened negative, participants who screened positive on the S-NIAS were more likely to report restriction of foods (63% vs 78%, OR 2.1, 95% confidence interval [CI] 1.1–3.8), recent weight loss (12% vs 39%, OR 2.1, 95% CI 1.3–3.4), and require some type of nutritional support (18% vs 37%, OR 2.1, 95% CI 1.3–3.4) and were more likely to report not being able to eat normal meals (12% vs 32%, OR 3.3, 95% CI 1.5–6.9) (Table 1).

We evaluated ARFID symptoms in relation to GI symptom burden and HrQoL by (1) comparing those with and without a positive ARFID screen and (2) exploring the association of ARFID severity to GI symptom burden (by UCLA SCTC GIT 2.0) and HrQoL severity (SF-12). Compared to those who screened negative, participants who screened positive for ARFID had greater GI symptom burden with large effects, except for the reflux and constipation subscales (Table 1); similar relationships were found when correlating ARFID scores with GI symptom scores (Figure 4). Those with ARFID also had worse HrQoL in all domains and in both the physical and mental component summaries (by SF-12) (Table 1), with similar relations found when correlating ARFID scores with HrQoL scores (Figure 4).

DISCUSSION

To our knowledge, this study is the first to assess the prevalence of ARFID symptoms in patients with SSc. Our findings indicate that ARFID symptoms, as measured by a self-reported survey (the S-NIAS), are notably prevalent, with approximately half of our cohort having elevated scores suggestive of possible ARFID. A positive ARFID screen was associated with greater self-reported weight loss, food restriction, and challenges in maintaining normal meal patterns. Additionally, ARFID symptoms were associated with greater GI symptom burden and worse HrQoL. This study highlights a potentially overlooked issue in SSc that may significantly impact HrQoL and represents a promising target for improving holistic care and management.

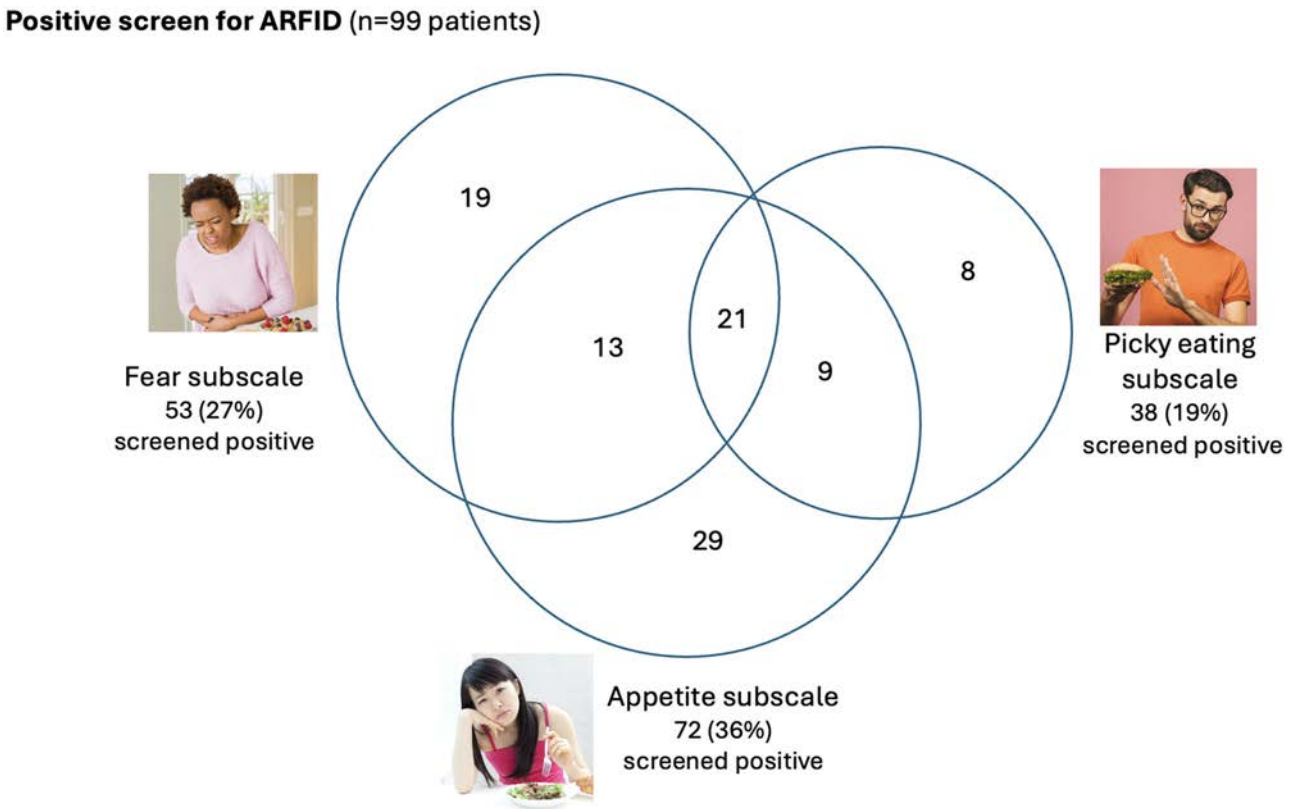


Figure 3. Diagram of positive screening rates by Spanish Nine-Item Avoidant or Restrictive Food Intake Disorder (ARFID) Screen subscale among the 99 patients (50%) who screened positive for ARFID. Each circle corresponds to a subscale. The numbers within the circles show the frequency of those who only screened positive on one subscale or multiple. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25569/abstract.

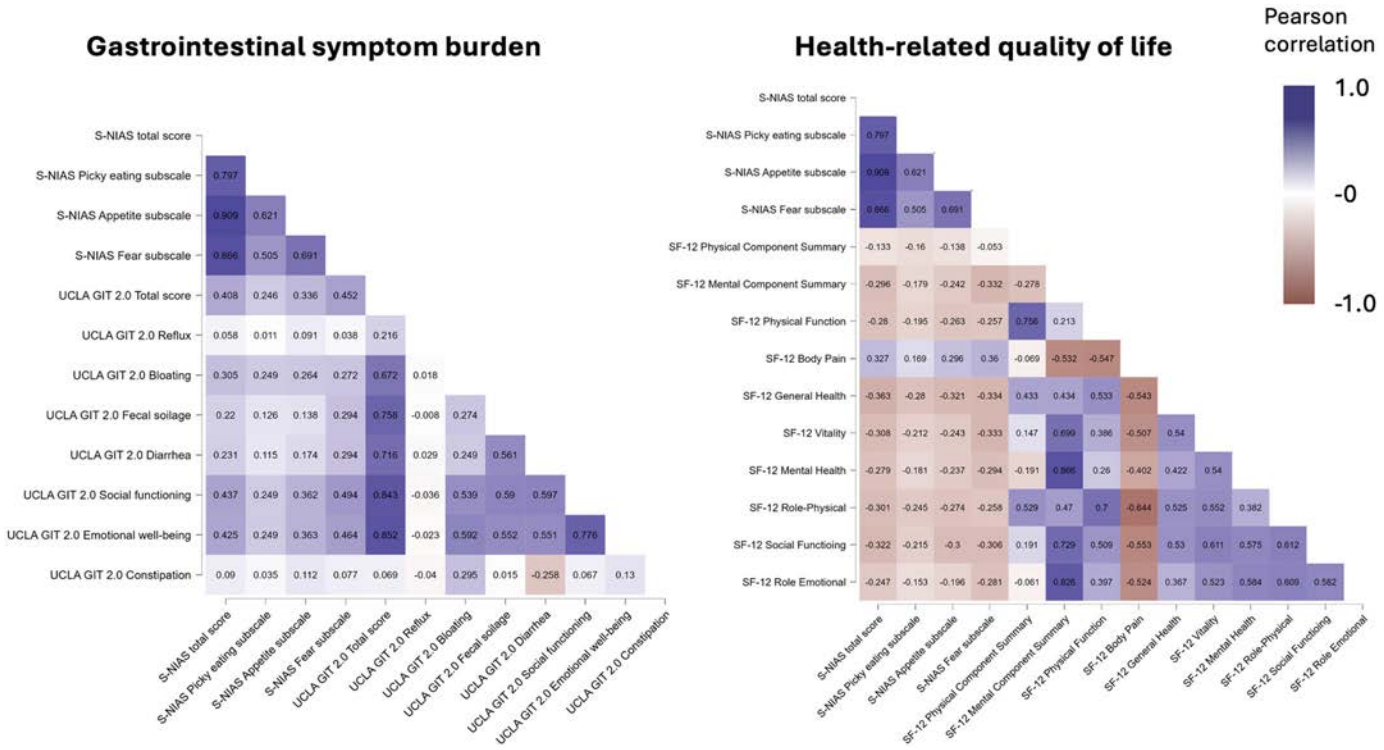


Figure 4. ARFID subscales correlations with gastrointestinal symptom burden and health-related quality of life (N = 200). (A) Pearson correlations between ARFID severity by S-NIAS total and subscale scores and gastrointestinal symptom severity subscale and total scores. (B) Pearson correlations between ARFID severity scales and health-related quality of life (SF-12 physical and mental summary scores as well as subscales). ARFID, avoidant or restrictive food intake disorder; SF-12, 12-item Short Form Survey; S-NIAS, Spanish Nine-Item avoidant or restrictive food intake disorder Screen; UCLA SCTC GIT 2.0, University of California Los Angeles Scleroderma Clinical Trial Consortium GI tract 2.0.

Overall, health studies that rely on patient self-participation show lower response rates for online surveys compared to in-person or postal recruitment, with an average response rate[28] of 46%. Specifically, published online surveys targeting patients with SSc report response rates[29,30] of approximately 30%. In our study, we achieved a 37% response rate, corresponding to 200 patients with SSc, and approximately half of the included cohort screened positive for ARFID on one or more subscales. The three primary motivations underlying restrictive eating in ARFID (sensory aversion, lack of interest or low appetite, and fear of aversive consequences) are intended to be captured by the three NIAS subscales.[31,32] Our findings suggest that a subset of patients with SSc endorse symptoms most associated with lack of interest or low appetite presentations and fear of aversive consequences—notably endorsing avoiding or delaying eating due to fear of discomfort (48%), having to push themselves to eat (41%), difficulty eating enough at meals (39%), a lack of interest in eating or small appetite (29%), and having only a short list of foods that they will eat (28%). Notably, the appetite subscale was the most frequently positive screen for ARFID in our sample, differing from patients with organic GI disorders who most commonly screen positive on the NIAS fear subscale.[20] The most likely explanation for this, fear of "discomfort," is used in all the fear subscale items in the NIAS (English version) but only in two of the four S-NIAS fear items, with the other two items being specific to fear of a discrete event (choking or vomiting). Another reason could be due to our strict definition of a positive screen, requiring the endorsement of at least two symptoms from each subscale, rather than using a total score.

The association between ARFID and fear of discomfort may be the result of but could also perpetuate greater GI symptom severity in SSc. We found that those with a positive ARFID screen had worse GI symptom severity, which appeared to be driven by worse GI-related social functioning and emotional well-being. Previous research has demonstrated that in patients with SSc, poor appetite, a higher burden of GI symptoms such as early satiety and nausea, and abnormal bowel habits are associated with an increased risk of malnutrition.[4,13] Although our study was cross-sectional, it is possible that some patients may restrict their diet in response to heightened GI discomfort resulting from GI motor dysfunction in the setting of SSc (eg, delayed gastric emptying and colonic dysmotility). Dietary restriction may initially be adaptive and temporarily alleviate distress around GI symptoms, for example, increasing a sense of control or reducing anxiety around symptoms. However, in the longer term, continued food avoidance may perpetuate habit processes that could actually contribute to heightened GI pain (eg, poor gastric accommodation), as has been proposed for functional or motility GI disorders.[33]

Although seeking diet and nutritional advice is common in patients with SSc,[10] only a small proportion of our sample (18%) reported seeking dietitian guidance. Although dietary management strategies may be adaptive and helpful for many patients with SSc, there may be a subset who are at an elevated risk for dietary management to become maladaptive. In fact, although rates of seeking dietitian advice did not differ between groups, we found that a positive ARFID screen was associated with a higher frequency of food restriction and a tendency to do so without seeking professional help. Although we did not capture data specific to the types of dietary approaches patients were using, whole diet exclusions (eg, gluten-free, avoidance of foods high in fermentable oligo-, di-, mono-saccharides, and polyols (FODMAPS)) may particularly elevate the risk for ARFID without a dietitian's guidance.[33] Gastroenterology patients who reported previously trialing exclusion diets (and mostly without dietitian guidance) are reportedly three times more likely to present with ARFID symptoms than those who have not.[34] Although guidelines and expert consensus consider dietary and lifestyle as part of the management of GI symptoms in SSc,[6,7,35] there is a lack of evidence for the use of empirical diet therapies and their benefit in patients with SSc.[8,9] However, research is growing, with one diet trial for 18 patients with advanced SSc providing dietary strategies that varied across patients, personalized to address weight loss (ie, increased caloric intake) and applicable GI symptoms.[36] Future research is needed on the dietary management of patients with SSc, but we believe health care providers should tailor their guidance to each patient's needs, ideally with the guidance of a dietitian experienced in the treatment of patients with SSc.

If dietary restrictions significantly impair nutritional status and/or reduce quality of life, the possibility of ARFID becomes a relevant concern. In our study, we found that a positive ARFID screen was associated with a higher frequency of recent weight loss and a higher frequency of nutritional support use, including enteral nutrition, suggesting that elevated S-NIAS scores aligned with worse nutritional status. In addition to ARFID being associated with worse emotional well-being and social functioning around GI symptoms, we determined that those with a positive ARFID screen had overall worse HrQoL, specifically with greater ARFID severity moderately associated with worse social and emotional functioning due to patient's physical health. Although longitudinal research is needed to understand the temporality, worse HrQoL (including GI-specific) could be a result of or be exacerbated by nutrition difficulties, including psychologic distress around eating.

Given that up to 20% of patients with SSc are at risk of malnutrition,[13] it is crucial to ensure that these patients maintain adequate oral intake and avoid unnecessary dietary restrictions. If a patient reports that they are using or have been recommended a whole exclusion diet, we suggest that clinicians ensure that patients are doing so under the guidance of a dietitian.[33] Although a formal diagnostic evaluation (eg, with a mental health professional) is needed to diagnose ARFID, we recommend clinicians consider screening for ARFID symptoms through clinical questioning. The NIAS may serve as a valuable tool for clinicians to identify patients experiencing eating-related distress, prompting further evaluation by a mental health professional, dietitian, or other eating disorder specialist.

Given the high prevalence of ARFID symptoms in patients with SSc identified in this study, a key next step is to assess the predictive value of the NIAS for clinical outcomes, including malnutrition risk and quality of life impairments. A prospective, in-clinic evaluation of ARFID screening using the NIAS in patients with SSc, followed by consultation with an experienced behavioral health clinician (eg, a psychologist) for those who screen positive, would help distinguish maladaptive eating behaviors from adaptive dietary modifications driven by SSc-related GI discomfort. Furthermore, to enhance the clinical utility of the NIAS in SSc, prospective cohort studies should investigate whether ARFID symptoms fluctuate in response to disease activity, GI symptom severity, or dietary modifications and whether improvement in ARFID symptoms correlates with enhancements in other HrQoL outcomes.

A major strength of our study is the inclusion of a robust sample of Spanish patients with SSc through a national patient society. However, our study has some limitations. First, we relied on self-reported questionnaire data to capture clinical features and SSc-specific characteristics, possibly resulting in some misclassification. Second, we were unable to evaluate specific disease features, particularly immunologic markers, the effects of concomitant medication, or the presence of overlapping GI diseases. Third, although we used an online survey tool to reach patients across the country, this approach potentially excluded individuals without internet access and/or those unable to respond via computer or mobile device. In addition, voluntary surveys are subject to participation bias, and despite achieving a good response rate, more than half of the society's members did not access the survey. Last, although the S-NIAS was developed and initially validated in Spanish,[17] there are not yet validated clinical screening cutoffs. To address this, we implemented a strict criterion for a positive screen, requiring participants to endorse at least two items from each subscale. We recognize that the S-NIAS may overestimate the prevalence of ARFID in patients with SSc, as some ARFID symptoms might reflect normal responses to GI involvement rather than a distinct eating disorder. Consequently, more research is needed to accurately determine the prevalence of ARFID in patients with SSc. Furthermore, future research could also incorporate examination of patterns energy expenditure and for differences in ARFID in patients with specific organ-based involvements (eg, significant cardiac or lung disease).[37]

In summary, our findings show that half of the patients with SSc who participated in our study endorsed high ARFID symptoms, which were associated with worse GI symptom burden,

worse HrQoL, and a greater frequency of some markers of nutritional inadequacy. Although some patients with SSc may adaptively restrict their diet to manage uncomfortable symptoms (eg, GI discomfort) related to their disease, a subset may experience significant impairments in HrQoL and/or nutritional consequences (eg, weight loss, reliance on supplemental nutrition) indicative of ARFID. Our findings emphasize the need for more research, particularly longitudinally, on diet and eating behaviors in patients with SSc, hopefully to guide future treatment recommendations and disease management strategies.

ACKNOWLEDGMENT

We would like to acknowledge Beatriz García from the AEE for her special contribution in the development of this work.

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 Alcala-Gonzalez 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. Bandini G, Alunno A, Ruaro B, et al. Significant gastrointestinal unmet needs in patients with systemic sclerosis: insights from a large international patient survey. Rheumatology Oxford 2024;63(3):e92–e93.
2. McMahan ZH, Kulkarni S, Chen J, et al. Systemic sclerosis gastrointestinal dysmotility: risk factors, pathophysiology, diagnosis and management. Nat Rev Rheumatol 2023;19(3):166–181.
3. Kaniecki T, Abdi T, McMahan ZH. A practical approach to the evaluation and management of gastrointestinal symptoms in patients with systemic sclerosis. Best Pract Res Clin Rheumatol 2021;35(3):101666.
4. Volkmann ER, McMahan ZH, Smith V, et al; SENSCIS Trial Investigators. Risk of malnutrition in patients with systemic sclerosis-associated interstitial lung disease treated with nintedanib in the randomized, placebo-controlled SENSCIS Trial. Arthritis Care Res (Hoboken) 2023;75(12):2501–2507.
5. Hughes M, Heal C, Siegert E, et al; EUSTAR Collaborators. Significant weight loss in systemic sclerosis: a study from the EULAR Scleroderma Trials and Research (EUSTAR) database. Ann Rheum Dis 2020;79(8):1123–1125.
6. Gwinnutt JM, Wieczorek M, Balanescu A, et al. 2021 EULAR recommendations regarding lifestyle behaviours and work participation to prevent progression of rheumatic and musculoskeletal diseases. Ann Rheum Dis 2023;82(1):48–56.
7. Denton CP, De Lorenzis E, Roblin E, et al. The 2024 British Society for Rheumatology guideline for management of systemic sclerosis. Rheumatol (Oxford) 2024;63(11):2956–2975.
8. Parodis I, Tsoi A, Gomez A, et al. Lifestyle interventions in the management of systemic sclerosis: a systematic review of the literature. Rheumatol Adv Pract 2024;8(2):rkae037.
9. Smith E, Pauling JD. The efficacy of dietary intervention on gastrointestinal involvement in systemic sclerosis: a systematic literature review. Semin Arthritis Rheum 2019;49(1):112–118.
10. Østbø N, Yakes Jimenez E, Carrier ME, et al; Scleroderma Patient-centred Intervention Network (SPIN) Diet and Nutrition Education Patient Advisory Team, on behalf of the SPIN Investigators. Use and perceptions of nutrition information resources in systemic sclerosis: a Scleroderma Patient-centred Intervention Network (SPIN) cohort study. Clin Exp Rheumatol 2024;42(8):1606–1614.
11. Staller K, Abber SR, Burton Murray H. The intersection between eating disorders and gastrointestinal disorders: a narrative review and practical guide. Lancet Gastroenterol Hepatol 2023;8(6):565–578.
12. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision (DSM-5-TR). Psychiatry Online. Accessed October 20, 2024. https://psychiatryonline.org/doi/book/10.1176/appi.books.9780890425787
13. Baron M, Hudson M, Steele R; Canadian Scleroderma Research Group. Malnutrition is common in systemic sclerosis: results from the Canadian scleroderma research group database. J Rheumatol 2009;36(12):2737–2743.
14. Burlui AM, Cardoneanu A, Macovei LA, et al. Diet in scleroderma: is there a need for intervention? Diagnostics (Basel) 2021;11(11):2118.
15. Krause L, Becker MO, Brueckner CS, et al. Nutritional status as marker for disease activity and severity predicting mortality in patients with systemic sclerosis. Ann Rheum Dis 2010;69(11):1951–1957.
16. Harris PA, Taylor R, Thielke R, et al. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42(2):377–381.
17. Medina-Tepal KA, Vazquez-Arevalo R, Trujillo-ChiVacuán EM, et al. Cross-cultural adaptation and validation of the Nine Item ARFID Screen (NIAS) in Mexican youths. Int J Eat Disord 2023;56(4):721–726.
18. Burton Murray H, Jehangir A, Silvernale CJ, et al. Avoidant/restrictive food intake disorder symptoms are frequent in patients presenting for symptoms of gastroparesis. Neurogastroenterol Motil 2020;32(12):e13931.
19. Hollis E, Murray HB, Parkman HP. Relationships among symptoms of gastroparesis to those of avoidant/restrictive food intake disorder in patients with gastroparesis. Neurogastroenterol Motil 2024;36(2):e14725.
20. Fink M, Simons M, Tomasino K, et al. When is patient behavior indicative of avoidant restrictive food intake disorder (ARFID) vs reasonable response to digestive disease? Clin Gastroenterol Hepatol 2022;20(6):1241–1250.
21. Robelin K, Senada P, Ghoz H, et al. Prevalence and clinician recognition of avoidant/restrictive food intake disorder in patients with inflammatory bowel disease. Gastroenterol Hepatol (N Y) 2021;17(11):510–514.
22. Burton-Murray H, Kiser K, Gurung J, et al. Avoidant/restrictive food intake disorder symptoms are not as frequent as other eating disorder symptoms when ulcerative colitis is in remission. J Crohns Colitis 2024;18(9):1510–1513.
23. Alcala-Gonzalez LG, Guillen-Del-Castillo A, Simeón-Aznar CP; Scleroderma GI Research Working Group; Annex. Scleroderma GI Research Working Group. Patient perspectives and treatment gaps in managing gastrointestinal symptoms in systemic sclerosis. Med Clin (Barc) 2025;164(4):196–198.
24. Cano-García L, Redondo-Rodríguez R, Mena-Vázquez N, et al. Severity and impact of digestive impairment perceived by patients with systemic sclerosis: a cross-sectional study. BMJ Open 2024;14(4):e083419.
25. Alcala-Gonzalez LG, Guillen-Del-Castillo A, Aguilar A, et al. Impact of gastrointestinal symptoms and psychological distress on quality of life

in systemic sclerosis: a cross-sectional study. BMJ Open 2024; 14(11):e089725.

26. Vilagut G, Valderas JM, Ferrer M, et al. [Interpretation of SF-36 and SF-12 questionnaires in Spain: physical and mental components]. Med Clin (Barc) 2008;130(19):726–735. In Spanish.
27. Schmidt S, Vilagut G, Garin O, et al. [Reference guidelines for the 12-Item Short-Form Health Survey version 2 based on the Catalan general population]. Med Clin (Barc) 2012;139(14):613–625. In Spanish.
28. Meyer VM, Benjamens S, Moumni ME, et al. Global overview of response rates in patient and health care professional surveys in surgery: a systematic review. Ann Surg 2022;275(1):e75–e81.
29. Toci AL, Hyer JM, Silver RM, et al. Systemic sclerosis and perceptions of quality in primary care. Am J Med Sci 2016;351(5):447–451.
30. Spierings J, van den Ende CHM, Schriemer RM, et al; ARCH Study Group. How do patients with systemic sclerosis experience currently provided healthcare and how should we measure its quality? Rheumatology (Oxford) 2020;59(6):1226–1232.
31. Zickgraf HF, Ellis JM. Initial validation of the Nine Item Avoidant/Restrictive Food Intake disorder screen (NIAS): a measure of three restrictive eating patterns. Appetite 2018;123:32–42.
32. Burton Murray H, Dreier MJ, Zickgraf HF, et al. Validation of the Nine Item ARFID screen (NIAS) subscales for distinguishing ARFID presentations and screening for ARFID. Int J Eat Disord 2021;54(10):1782–1792.
33. Murray HB, Doerfler B, Harer KN, et al. Psychological considerations in the dietary management of patients with DGBI. Am J Gastroenterol 2022;117(6):985–994.
34. Atkins M, Zar-Kessler C, Madva EN, et al. History of trying exclusion diets and association with avoidant/restrictive food intake disorder in neurogastroenterology patients: a retrospective chart review. Neurogastroenterol Motil 2023;35(3):e14513.
35. Pope JE, Denton CP, Johnson SR, et al. State-of-the-art evidence in the treatment of systemic sclerosis. Nat Rev Rheumatol 2023;19(4): 212–226.
36. Doerfler B, Allen TS, Southwood C, et al. Medical Nutrition Therapy for Patients with Advanced Systemic Sclerosis (MNT PASS): a pilot intervention study. JPEN J Parenter Enteral Nutr 2017;41(4):678–684.
37. Hughes M, Harrison E, Herrick AL, et al. The need to accurately measure energy intake and expenditure in patients with systemic sclerosis. J Scleroderma Relat Disord 2022;7(3):217–223.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 77, No. 11, November 2025, pp 1384–1392
DOI 10.1002/acr.25567

Validation of Lung Ultrasound Interpretation Criteria for Interstitial Lung Disease in Systemic Sclerosis and Inflammatory Myopathy

Robert M. Fairchild,[1] Diane A. Mar,[2] Mariani D. Deluna,[1] Melody Chung,[3] Srijana Davuluri,[1] Yumeko Kawano,[4] Henry Guo,[1] Matthew C. Baker,[1] David Fiorentino,[1] and Lorinda Chung[2]

Objective. Interstitial lung disease (ILD) has a high prevalence in patients with systemic sclerosis (SSc) and inflammatory myopathy (IM), and early identification reduces associated morbidity and mortality. We previously developed lung ultrasound (LUS) interpretation criteria for ILD detection in 2020 (LUS-ILD-20) showing excellent sensitivity and specificity in patients with SSc-ILD; herein, we sought to validate a revised LUS-ILD-24 in a large cohort with SSc and IM.

Methods. Patients meeting criteria for SSc and IM, with planned computed tomography (CT) chest imaging underwent LUS imaging interpreted with LUS-ILD-24 by three blinded readers. The sensitivity and specificity for LUS-ILD detection as noted on CT was analyzed for subgroups with SSc, IM, and possible incident ILD. Inter- and intrarater agreements were calculated. Correlations between LUS-ILD-24 severity, CT imaging severity, and pulmonary function tests were assessed.

Results. Ninety-five patients were included in the analyses. Sensitivity and specificity for ILD detection ranged from 92.4% to 95.5% and 82.8% to 86.2% across readers with similar accuracy in all subgroups. Inter- and intrareader reliability showed near perfect agreement ($\kappa = 0.92$ and $\kappa = 0.90$ to 1, respectively). LUS severity correlated with CT imaging severity and inversely correlated with diffusion capacity for carbon monoxide and forced vital capacity.

Conclusion. We validated our revised LUS-ILD-24 in cohorts with SSc and IM and found excellent sensitivity, specificity, and reliability for detection of ILD identified on CT imaging. LUS severity correlated with CT and pulmonary function test markers of ILD severity. Validation of the revised LUS-ILD-24 supports the implementation of LUS imaging in screening algorithms for ILD in patients with SSc and IM.

INTRODUCTION

Interstitial lung disease (ILD) is a leading cause of morbidity and mortality in rheumatic disease with a prevalence of up to 50% in patients with systemic sclerosis (SSc) and inflammatory myopathy (IM).[1,2] Early detection allows for timely intervention and improved outcomes.[3,4] Guidelines recommend screening for ILD in high risk systemic autoimmune rheumatic diseases with a combination of pulmonary function tests (PFTs) and high-resolution computed tomography (CT) imaging of the chest.[5,6] However, PFTs may not be sensitive in the early stages of ILD,[7] whereas CT scans have considerable associated financial and environmental costs, expose patients to significant ionizing radiation, and may not be readily accessible or available in certain areas.[8–10] As a result, only 51% of general rheumatologists and 66% of SSc experts routinely order chest CT in all newly diagnosed patients with SSc.[5,8,11]

Ultrasound is a radiation-free, cost-effective, environmentally-friendly, and readily available imaging modality that is well tolerated.[9,12,13] There is a large body of literature suggesting lung ultrasound (LUS) has use in patients with connective tissue disease (CTD)-associated ILD, with high accuracy for ILD

Drs Fairchild and Chung's work was supported by Boehringer Ingelheim. Dr Chung's work was supported by the Scleroderma Research Foundation.

[1]Robert M. Fairchild, MD, PhD, RhMSUS, Mariani D. Deluna, BS, Srijana Davuluri, MD, Henry Guo, MD, PhD, Matthew C. Baker, MD, MS, David Fiorentino, MD, PhD: Stanford University School of Medicine, Stanford, California; [2]Diane A. Mar, MD, RhMSUS, Lorinda Chung, MD, MS: Stanford University School of Medicine, Stanford, California, and VA Palo Alto Health Care System, Palo Alto, California; [3]Melody Chung, MD: Kaiser Permanente, Hyattsville, Maryland; [4]Yumeko Kawano, MD: Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25567).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25567.

Address correspondence via email to Robert M. Fairchild, MD, PhD, RhMSUS, at rfairchi@stanford.edu.

Submitted for publication November 20, 2024; accepted in revised form April 23, 2025.

SIGNIFICANCE & INNOVATIONS

- This study validates our revised lung ultrasound (LUS) for interstitial lung disease (ILD) detection in 2024 (LUS-ILD-24) interpretation criteria for LUS of systemic sclerosis and inflammatory myopathy-associated ILD with excellent accuracy.
- The revised LUS-ILD-24 criteria performs well in assessing ILD in baseline connective tissue diagnosis, follow-up screening, and evaluation for new suspected ILD.
- LUS-ILD-24–derived severity scoring correlates with both pulmonary function testing and chest computed tomography severity suggesting usage in disease monitoring.

detection.[14–17] The ability of LUS to detect ILD is a result of the peripheral and basilar pattern of disease in nonspecific interstitial pneumonia (NSIP) and usual interstitial pneumonia (UIP), the most frequently encountered ILD subtypes associated with CTD.[18–20] LUS findings of pleural irregularity, pleural line thickening, and the presence of B-lines are known to be associated with ILD (Supplementary Figure 1), and studies on these LUS imaging findings show high diagnostic accuracy compared to CT for ILD detection.[14,21–23] Even with this mounting body of evidence, no validated set of interpretation criteria are available, and few ultrasonographers are familiar with LUS technique or interpretation in ILD. To address this, we previously developed a set of novel LUS interpretation criteria for ILD in patients with SSc (LUS-ILD-20) based on LUS experience at our institution with the goal of providing a simple set of interpretation criteria to detect SSc-ILD in the clinical setting.[14] In a pilot study with 20 patients with SSc, we demonstrated a sensitivity of 100% and specificity of 82% for CT-detected ILD with perfect agreement between two readers. Encouraged by these results, we sought to validate a set of revised LUS interpretation criteria (LUS-ILD-24) for clinical CTD-ILD detection in a large cohort of patients with SSc and IM and to determine whether these criteria could also be used to determine ILD severity.

PATIENTS AND METHODS

LUS interpretation criteria. Based on our clinical and research experience at our academic institution, including LUS evaluation of several hundred patients over seven years in SSc-ILD, IM-ILD, SARS-CoV-2–associated pulmonary disease, and pediatric systemic juvenile idiopathic arthritis related lung disease (SJIA-LD), our continued clinical use of LUS-ILD-20 for CTD-ILD detection,[14,16,24] as well as ongoing research and standardization of definitions in LUS by other researchers,[25,26] we made minor modifications to the LUS-ILD-20 criteria to improve clarity and provide more consistent terminology (see Supplementary Methods). We also added exceptions for pleural abnormalities due to lung fissures and diaphragmatic proximity based on our experience (Supplementary Figure 1). These changes resulted in the revised LUS-ILD-24 (Figure 1). The primary framework and function of the criteria remained unchanged.

Validation cohort. The study was approved by Stanford University's institutional review board, and informed consent was obtained from all patients. For this prospective observational imaging study, we approached all patients seen at our institution aged 18 years and over with diagnoses of SSc, dermatomyositis, or polymyositis with planned chest CT for (1) evaluation of known ILD, (2) evaluation for suspected ILD, or (3) initial baseline ILD screening in newly diagnosed CTD. Patients with a known history of lung cancer, lung surgery, lung infection, heart failure, aspiration, or other known non-ILD pulmonary parenchymal disease were excluded. Patients unable to undergo LUS examination for nondisease-related reasons (eg anxiety, gel allergy, etc) were also excluded.

CT and PFT collection. Volumetric CT data were collected and read by dedicated chest CT radiologists. Volumetric CT data were also analyzed using Computer-Aided Lung Informatics for Pathology Evaluation and Rating (CALIPER) lung textural analysis.[27] CALIPER textural analysis is a validated machine learning algorithm that provides CT analysis for findings of hyperlucency, reticulation, honeycombing, and ground glass opacities, each quantified by lung region. The CALIPER overall percentage of lung volume with ILD (c%[ILD]) combines quantified ground glass, reticulation, and honeycombing. The CALIPER overall percentage of lung volume of fibrosis (c%[fibrosis]), combines quantified reticulation and honeycombing. PFTs obtained at the time of CT acquisition for ILD investigation were analyzed using percent predicted diffusing capacity for carbon monoxide (%DL_{CO}) and percent predicted forced vital capacity (%FVC), when available.

LUS examination and interpretation. LUS examinations were performed within 10 days of CT acquisition. We followed our previously reported LUS acquisition protocol without modification.[14] LUS was obtained by RMF, who has seven years of LUS experience, or by MDD under the direct supervision of RMF. Three LUS readers (RMF, DAM, and MDD) independently read all LUS examinations using the revised LUS-ILD-24 criteria. LUS studies were randomized, and readers were blinded to CT and to each other. Randomized LUS studies were interpreted consecutively and were completed after three reading sessions. At the time of interpretation, RMF and DAM had seven and four years of experience respectively in acquiring and interpreting LUS in patients with CTD-ILD, LUS in patients with SARS-CoV-2, and LUS in patients with pediatric SJIA-LD.[14,16,24] MDD had limited prior LUS interpretation experience. Before LUS interpretation, RMF trained DAM and MDD on the revised LUS-ILD-24

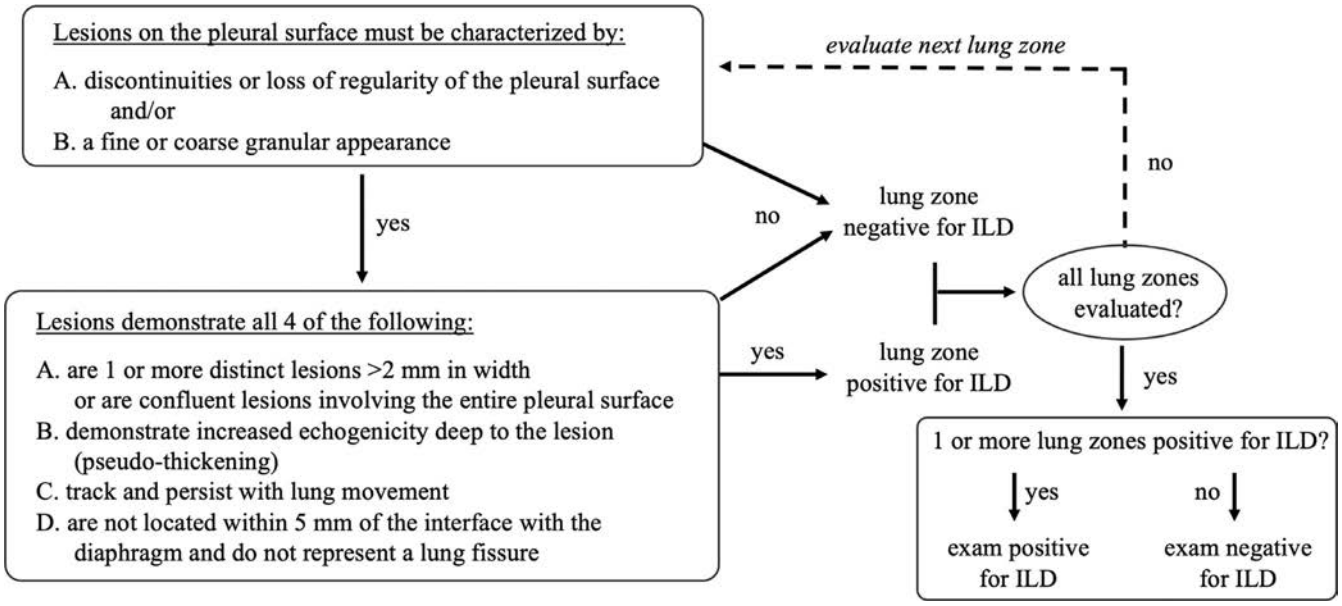


Figure 1. Revised 2024 LUS interpretation criteria for ILD screening in connective tissue disease (revised LUS-ILD-24). We used previously described LUS acquisition protocol, including[14] LUS windows. Connective tissue disease includes systemic sclerosis and inflammatory myopathy. ILD; interstitial lung disease; LUS; lung ultrasound.

criteria with several representative LUS cases not included in the current validation set. To assess intrareader agreement, repeat interpretation was performed on 20 randomly chosen studies six months after initial reading was complete.

Data analysis. For each participant and reader, LUS-ILD-24 criteria interpretation results were analyzed for the presence of ILD (LUS-ILD+ versus LUS-ILD–), as well as the number and location of positive lung zones (LZs) by interpretation criteria (LZ-ILD+ versus LZ-ILD–). LUS severity for each participant was calculated by summing the number of LZ-ILD+-interpreted regions. CT scans were noted as CT+ if ILD was noted by radiologist interpretation. The Mann-Whitney test was used for comparisons for both PFT and CALIPER indices between patients identified as CT+ and patients identified as CT– as well as patients with CT +-prevalent ILD versus patients with incident ILD. Sensitivity and specificity for ILD detection on LUS were calculated using CT+ as the gold standard for the presence of ILD. The performance of the LUS-ILD-24 criteria were analyzed separately for the interpretation of experienced readers RMF and DAM, and consensus interpretation among the three readers (two or more readers identifying LUS-ILD+). Sensitivity and specificity analysis was performed on four patient subsets consisting of (1) all participants, (2) patients with SSc, (3) patients with IM, and (4) patients with possible ILD. The group with possible ILD consisted of patients without known ILD who had either a diagnosis of CTD and were obtaining CT because of an increased clinical concern for ILD or who had new CTD diagnoses undergoing CT for baseline ILD screening. Confidence intervals (CIs) for sensitivity and specificity were calculated using the Wilson/Brown method. Interreader agreement between the three readers was calculated for both LZ-ILD+ interpretation and LUS-ILD+ interpretation using Randolph's free marginal multirater kappa,[28] and intrareader agreement was assessed with Cohen's kappa. Correlations between LUS severity and %DLCO, %FVC, c%(ILD), and c%(fibrosis) were assessed by Spearman correlation.

RESULTS

Between 2020 and 2023, we identified 168 patients with prior or new diagnoses of SSc and IM with planned CT for baseline ILD screening, suspected ILD, or ILD disease monitoring and management. Of these, 100 patients were enrolled and completed their CT and LUS imaging with median months from CTD diagnosis to LUS imaging of 1 month (interquartile range [IQR] 1–3) for the base ILD screening group, 48 months (IQR 24–176) for suspected ILD, and 55 months (IQR 36–146.3) for the ILD disease monitoring group. (Figure 2). The mean age at the time of enrollment was 55 ± 14 years, and the majority (77%) were women (Table 1). Patients had diagnoses of limited SSc (41%), diffuse SSc (30%), dermatomyositis (26%), and polymyositis (3%). Among patients with SSc, the most common associated autoantibodies were Scl-70 (33%), anticentromere (28%), and RNA-polymerase-3 (15%). The most common IM-associated antibodies were MDA5 (14%), TIF1-gamma (7%), NXP-2 (11%), Jo-1 (25%), Ku (11%), Ro 60 (4%), PL7 (14%), Mi-2 (4%). Organizing pneumonia (OP) was seen in seven patients with IM. After CT imaging, five patients were excluded from the primary

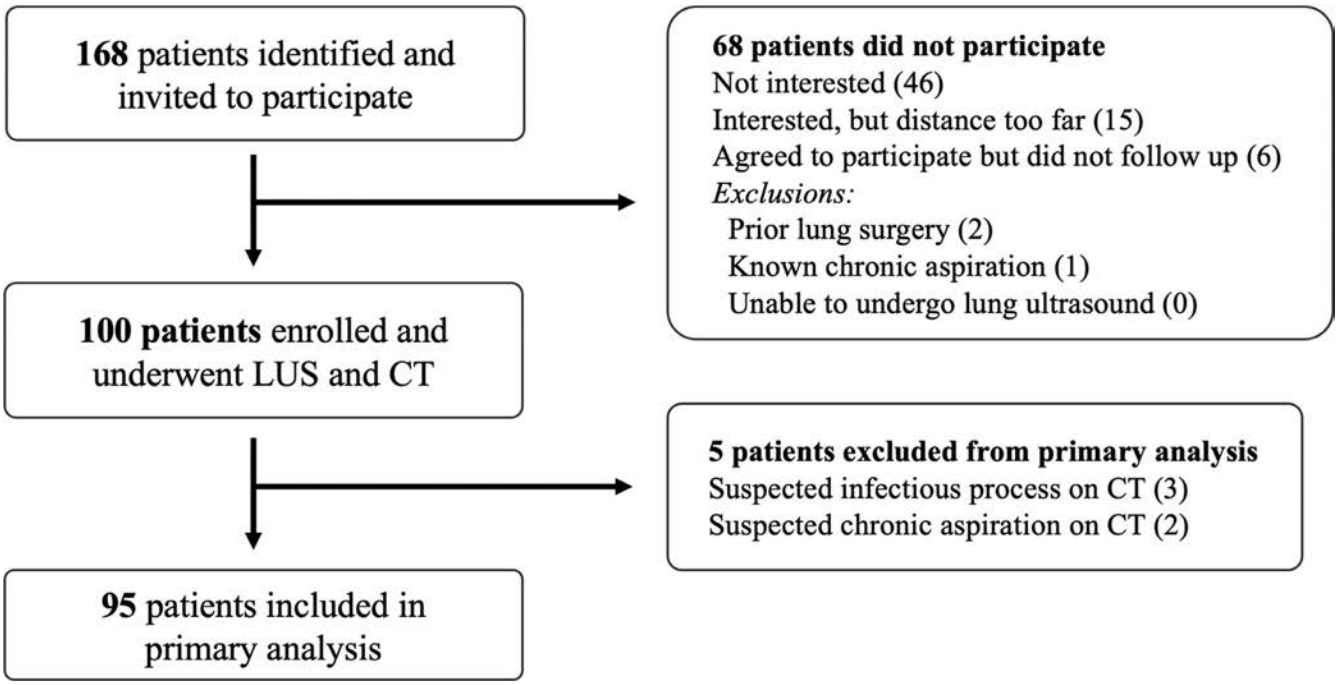


Figure 2. Flow diagram for patient identification, enrollment, and analysis. CT, computed tomography imaging; ILD, interstitial lung disease; LUS, lung ultrasound.

Table 1. Baseline characteristics of enrolled patients*

Characteristics	n (%)
Total enrolled patients	100
Age, mean ± SD	55 ± 14
Sex, female	77 (77)
CTD diagnosis	
Limited SSc	41 (41)
Diffuse SSc	30 (30)
DM	26 (26)
PM	3 (3)
Reason for examination	
No Prior ILD	46 (46)
Baseline evaluation at new CTD diagnosis	19 (19)
Interval evaluation for suspected ILD	27 (27)
Prevalent ILD follow-up	54 (54)
After excluding infection or aspiration on CT	95 (95)[b]
Incident ILD on CT	16 (16.8)
Prevalent ILD on CT[c]	50 (52.6)
Negative for ILD on CT	29 (30.5)
SSc	67 (71)
NSIP	42 (91)
UIP	4 (9)
DM and PM	28 (29)
NSIP	11 (55)
UIP	1 (5)
OP	8 (40)

* CT, computed tomography imaging; CTD, connective tissue disease; DM, dermatomyositis; ILD, interstitial lung disease; NSIP, nonspecific interstitial pneumonia; OP, organizing pneumonia; PM, polymyositis; SSc, systemic sclerosis; UIP, usual interstitial pneumonia.
[a] The five patients excluded from the primary analysis were all within this interval evaluation subgroup.
[b] Cohort selected for analysis.
[c] Four patients with prior diagnoses of ILD had resolution of ILD on CT imaging.

analyses because of new CT findings associated with recent SARS-CoV-2 infection (n = 1), other infectious process (n = 2), and chronic aspiration (n = 2) (Figure 2).

Of the 95 patients remaining in this primary analysis cohort, 66 (69.5%) patients were found to be CT+ for ILD and 29 (30.5%) patients were found to be CT– for ILD. Patients who were identified as CT+ had imaging findings consistent with NSIP (n = 53, 80%), OP (n = 8, 12%), and UIP (n = 5, 8%); OP was only observed in the group with IM. CT was obtained for either baseline ILD screening in patients with new CTD diagnoses or for patients with suspicion of ILD with chronic CTD (n = 41, 43.2%). Of these, 16 (39.0%) patients were identified as CT+ for incident ILD and 25 patients (61.0%) were identified as CT– for ILD. All 16 patients with incident ILD had NSIP on CT imaging. Within the primary analysis cohort, there were 54 (56.8%) patients with known CTD-ILD on prior CT imaging. Within this group, CT imaging obtained at the time of enrollment revealed that 50 patients were identified as CT+ for ILD, and four patients were identified as CT– for ILD, likely representing resolution of ILD findings on CT in patients who were identified as CT–.

CT examinations were also assessed using CALIPER in 69 participants. CALIPER analysis could not be obtained in the remaining 26 patients because of technical issues with imaging input compatibility or CALIPER analysis failure. These technical issues occurred randomly across the cohort. Among participants who were identified as CT+, CALIPER analysis showed the median c%(ILD) of lung volume of 5.7% (IQR 2.5–12.4), whereas

participants who were identified as CT– resulted in a median of 0.3% (IQR 0.2–1.0), showing a statistically significant difference ($P < 0.0001$). Similarly, a statistically significant difference was seen between participants identified as CT+ versus participants identified as CT– c%(fibrosis) of lung volume, ($P < 0.0001$) with a median of 1.4% (IQR 0.6–2.5) versus a median of 0.2% (IQR 0.1–0.6), respectively.

PFTs were available for 79 participants. The median (IQR) for %FVC between patients identified as CT+ versus patients identified as CT– was 85% (72–96.25) versus 90% (81–104), which was not significantly different ($P = 0.08$). Conversely, a statistically significant difference was seen between patients identified as CT+ versus patients identified as CT– %DLCO ($P = 0.0001$) with a median (IQR) of 72% (57.25–89) versus 97% (86–105), respectively.

We also compared CALIPER and PFT indices in patients identified as CT+ among patients previously diagnosed with ILD versus new CT+ diagnoses within the cohort with possible ILD. The %DLCO of the previously diagnosed patients were significantly lower than new ILD diagnosis patients with median (IQR) of 71.0% (54.75–84.25) versus 87.5% (70.0–101.5; $P = 0.022$). Similarly, %FVC was significantly lower in previously diagnosed patients 84.0% (70.0–92.0) versus 93.5% (77.5–102.0; $P = 0.037$). CALIPER indices of c%(ILD) and c%(fibrosis) did not differ significantly between these groups.

LUS detection of CT-ILD. Across the entire primary analysis cohort, the sensitivity of LUS detection of ILD (CT+) as read by RMF and DAM, and consensus was 95.5% (95% CI 87.5–98.9), 93.9% (95% CI 85.4–97.6), and 92.4% (95% CI 83.5–96.7) respectively. Specificities were 82.8% (95% CI 65.5–92.4), 86.2% (95% CI 69.4–94.5), and 82.8% (95% CI 65.5–92.4), respectively. Similar sensitivities and specificities were observed across the subsets of patients with SSc, patients with IM, and patients with possible ILD (Table 2). Consensus reading demonstrated lower sensitivities without any improvement in specificities as compared to the individual experienced readers RMF and DAM. All five patients excluded from the primary analysis cohort due to non-ILD parenchymal pathology on CT were read as LUS-ILD+ across all readers. Inclusion of these patients in the analysis cohort resulted in unchanged sensitivities and reduced specificities (see Supplementary Table 1).

LZ-ILD+ interpretation was most frequently observed in the posterior axillary, subscapular, and paravertebral positions. More generally, LZ-ILD+ interpretation was encountered 1.7 times more often in these posterior positions compared to anterior positions. LZ-ILD+ identification was relatively evenly distributed between the right and left lungs with some increased findings in the right posterior zones compared with left (see Supplementary Table 2). We performed a post hoc analysis eliminating the anterior lung fields from LUS interpretation to determine whether an abbreviated posterior only six lung position examination could be sufficient for screening. Across the entire primary analysis cohort, interpretation by RMF and DAM and consensus excluding the anterior lung fields resulted in a sensitivity/specificity for ILD detection of 90.9%/86.2%, 87.8%/86.2%, and 87.9%/86.2% respectively (see Supplementary Table 3). This mild reduction in sensitivity and increase in specificity was also seen when analyzing the cohorts with SSc and possible ILD. In the cohort with IM, posterior only interpretation resulted in mildly reduced sensitivity and no improvement in specificity.

Table 2. Sensitivity and specificity for LUS detection of ILD on CT using the Revised LUS-ILD-24 criteria in patients with SSc and IM*

Readers and cohort[a]	TP	TN	FP	FN	Sensitivity, %	Specificity, %	Sensitivity, 95% CI, %	Specificity, 95% CI, %
Reader 1 (RMF)								
All	63	24	5	3	95.5	82.8	(87.5–98.8)	(65.5–92.4)
SSc only	41	15	4	2	95.7	81.0	(85.5–99.2)	(60.0–92.3)
IM only	19	7	1	1	95.0	87.5	(76.4–99.7)	(52.9–99.4)
Possible ILD	16	21	4	0	100.0	84.0	(80.6–100.0)	(65.4–93.6)
Reader 2 (DAM)								
All	62	25	4	4	93.9	86.2	(85.4–97.6)	(69.4–94.5)
SSc only	40	16	3	3	93.4	85.7	(82.5–97.8)	(65.4–95.0)
IM only	19	7	1	1	95.0	87.5	(76.4–99.7)	(52.9–99.4)
Possible ILD	15	22	3	1	93.8	88.0	(71.6–99.7)	(70.0–95.8)
Consensus[b]								
All	61	24	5	5	92.4	82.8	(83.5–96.7)	(65.5–92.4)
SSc only	39	15	4	4	91.3	81.0	(79.7–96.6)	(60.0–92.3)
IM only	19	7	1	1	95.0	87.5	(76.4–99.7)	(52.9–99.4)
Possible ILD	15	21	4	1	93.8	84.0	(71.7–99.7)	(65.4–93.6)

* CI, confidence interval; CT, computed tomography; FN, false negative; FP, false positive; ILD, interstitial lung disease; IM, inflammatory myopathy; LUS, lung ultrasound; possible ILD, patients without prior ILD diagnosis undergoing baseline or follow-up assessment; SSc, systemic sclerosis; TN, true negative; TP, true positive.
[a] Cohorts are subdivided as the following: All (n = 95), SSc only (n = 67), IM only (n = 28), and possible ILD (n = 41).
[b] Consensus was defined by identification by two to three out of three readers.

We also assessed the effect of combining LUS and PFTs for ILD screening in 79 patients who had PFTs available. ILD was detected when any of the following criteria were met: if LUS-ILD + was identified using the revised LUS-ILD-24 criteria, %DLCO < 80%, or %FVC < 80%. The addition of PFTs reduced the number of false negatives and increased the false positives to a lesser extent (see Supplementary Table 4). This resulted in sensitivities of 95% or greater for ILD detection across all patients for expert readers and consensus reading but with reduced specificities.

Reader agreement. Interrater reliability for LUS-ILD+ interpretation across the 95 studies was excellent (κ = 0.92; 95% CI 0.85–0.98), with 95.8% overall agreement in interpretation among the three readers (Table 3). Similar high agreement (κ = 0.79; 95% CI 0.76–0.81) was seen among the three readers for the 1,428 interpreted LZs (LZ-ILD+ versus LZ-ILD–) with 89.3% overall percent agreement across the three readers. Using reinterpretation of 20 studies six months after initial reading, intrareader reliability for individual images (LZ+ versus LZ–) for RMF, DAM, and MDD were κ = 0.89 (95% CI 0.78–1.00), κ = 0.84 (95% CI 0.78–1.00), κ = 0.80 (95% CI 0.78–1.00), and for overall studies (LUS+ versus LUS–) intrareader agreement was κ = 0.90 (95% CI 0.71–1.00), κ = 0.90 (95% CI 0.71–1.00), and κ = 1, respectively.

LUS severity. Summing the number of LZ-ILD+ per examination provided a severity score for each patient (0–14). After obtaining an average severity scores across all three readers for each patient, the median severity score across the entire primary analysis cohort was 4.0 (IQR 0.3–7.0) and the median severity score for only those patients who were identified as LUS-ILD+ by consensus reading was 5.7 (IQR 3.3–8.4). Qualitative review of the scoring revealed similar scoring between RMF and DAM and lower average severity scores for the interpretation from MDD.

Using average severity scores for each patient, LUS severity showed strong correlation to CALIPER measures of c%(ILD), r = 0.70 (95% CI 0.54–0.80), and moderate correlations with c% (fibrosis), r = 0.58, (95% CI 0.39–0.72). There was moderate negative correlation of severity scoring to %DLCO, r = –0.50 (95% CI –0.65 to –0.30), and a weak negative correlation between LUS severity scoring and %FVC, r = –0.24 (95% CI –0.44 to –0.01).

DISCUSSION

Our initial LUS-ILD-20 interpretation criteria were formulated with the goal of creating a simple and standardized method for interpreting LUS imaging for the presence or absence of ILD in patients with SSc. An initial pilot using the LUS-ILD-20 showed excellent sensitivity and specificity for ILD detection in patents with SSc.[14] In a more recent study performed by another group, similarly high accuracy was observed in 29 patients with SSc using LUS-ILD-20.[29] With continued experience using these criteria, we made minor modifications to clarify and simplify the LUS-ILD-20 criteria. Among these, the addition of criterion 2d to the interpretation criteria provides exemptions for lung fissures and physiologic irregularities at the very inferior margins of the lungs. These findings are commonly observed in most individuals and can result in false positive interpretation using the LUS-ILD-20 criteria. In this current validation study, the revised LUS-ILD-24 criteria performed well in detecting CT-identified ILD with excellent sensitivity and specificity in a validation cohort of patients with SSc and IM. Importantly, LUS-ILD-24 criteria performed equally as well on a subset of patients with possible ILD referred for CT chest for baseline ILD evaluation in patients with

Table 3. Interreader and intrareader reliability and LUS-ILD-24*

	Individual lung zones, κ (95% CI)	For studies, κ (95% CI)	LUS-ILD-24 severity, *ρ* (95% CI)[a]
Reader reliability			
Interreader reliability[b]	0.79 (0.76–0.81)	0.92 (0.85–0.98)	–
Intrareader reliability[c]			
Reader 1 (RMF)	0.89 (0.83–0.95)	0.90 (0.71–1)	–
Reader 2 (DAM)	0.84 (0.77–0.91)	0.90 (0.71–1)	–
Reader 3 (MDD)	0.80 (0.71–0.89)	1	–
Measures of severity			
PFTs			
%DLCO	–	–	−0.50 (−0.65 to −0.30)
%FVC	–	–	−0.24 (−0.44 to −0.01)
CALIPER			
total % fibrosis	–	–	0.58 (0.39–0.72)
total % ILD	–	–	0.70 (0.54–0.80)

* CALIPER, Computer-Aided Lung Informatics for Pathology Evaluation and Ratings (artificial intelligence quantification of computed tomography ILD findings); CI, confidence interval; DLCO, diffusion capacity of the lungs for carbon monoxide; FVC, forced vital capacity; ILD, interstitial lung disease; LUS, lung ultrasound; PFTs, pulmonary function tests.
[a] Spearman correlations. LUS-ILD-24 severity is defined by revised 2024 LUS interpretation criteria severity assessed by summing positive lung zones for each patient using consensus reading.
[b] Intrareader reliability for three readers assessed using Randolph's free marginal multirater kappa.
[c] Intrareader reliability calculated using Cohen's kappa.

newly diagnosed CTD or suspected ILD in patients with known CTD.

We analyzed sensitivity and specificity in two scenarios: experienced readers (RMF and DAM) and consensus interpretation across three readers. We found excellent agreement across all readers for LZ interpretation as well as overall examination interpretation. Sensitivity and specificity remained high across all patient cohorts for experienced readers. Mildly reduced sensitivity without improved specificity was observed with consensus interpretation. These findings suggest that individual experienced readers may provide the best performance in screening for ILD using the LUS-ILD-24 criteria. The cohort with possible ILD deserves special attention because this group represents patients for whom LUS evaluation may be most useful. Both experienced readers (RMF and DAM) and consensus reading using the LUS-ILD-24 were highly sensitive for ILD detection in this cohort, with only one false negative as interpreted by DAM and consensus readings. To assess if this group had less severe disease which might be more difficult to detect on LUS, we compared PFT and CALIPER indices in patients identified as CT+ between the cohort with possible ILD versus those with prior ILD diagnoses. The group with possible ILD had less prominent reductions in %DLco and %FVC compared with the patients identified as CT+ with prior ILD diagnoses. CT imaging severity by CALIPER quantification was not significantly different between these groups. These findings suggest similar severity in CT imaging between these groups but more significant functional impact in patients with chronic ILD; this may explain why LUS-ILD-24 performs well across all subsets.

Previously, LUS-ILD-20 criteria were evaluated in patients with SSc only. In this study, the LUS-ILD-24 criteria performed similarly well in patients with IM, even in patients found to have OP, an ILD type less explored with LUS. Of the 20 patients with IM found to be CT+, eight had OP, of which seven were LUS+ and one LUS– across all interpretation groups. Further work to understand the accuracy of LUS in patients with IM-related OP is needed.

We also assessed a whether a limited posterior only lung examination could be sufficient for ILD detection using the LUS-ILD-24. Posterior only examinations can be obtained rapidly in clinic with patients only needing to lift the back of their shirts, rather than gowning often required for anterior LUS examination. Interestingly, the sensitivity for ILD detection after excluding anterior views was only reduced by roughly 5% across cohorts and readers, with some improvement in specificities (see Supplementary Table 3). This approach may find use as a rapid screen for rheumatology ultrasonographers in their clinics.

PFTs are currently an integral component of screening for ILD. PFTs have a high false negative rate for ILD detection on their own but are helpful in combination with other modalities.[7] When we combined LUS and PFTs we found increased sensitivity and somewhat reduced specificity for ILD detection (see Supplementary Table 4). Because PFT assessment is standard practice, combining PFTs and LUS may provide additional accuracy. However, this study was not powered to specifically ascertain maximized PFT cutoffs for combined LUS and PFTs detection of ILD and more work is needed.

We also evaluated the effectiveness of an ILD severity scoring system based on the revised LUS-ILD-24, which demonstrated strong correlations with c%(ILD), moderate positive and negative correlations with c%(fibrosis) and %DLCO, respectively, and a weak negative correlation with %FVC. These results indicate that, beyond detecting ILD, the revised LUS-ILD-24 criteria may be a valuable tool for longitudinal monitoring of ILD. Further longitudinal studies incorporating LUS, CT imaging, and PFTs are necessary to better understand the utility of the LUS-ILD-24 in disease monitoring. Because LUS can be performed more frequently than CT, early follow-up LUS imaging after therapy is initiated for ILD has the potential to provide early objective evidence of therapeutic response, or lack thereof, and may allow for earlier modifications to therapy when needed.

Recently, parallel work by the Outcome Measures in Rheumatology (OMERACT) ultrasound working group on ILD attempted to obtain consensus-based definitions of LUS findings associated with ILD.[25,26] Consensus definitions for B-lines and pleural irregularity were achieved. However, only a moderate interreader agreement was seen in a Delphi study review of 80 LUS video clips[25] using these definitions. In contrast to our study, the frequency ranges in this OMERACT study were focused on lower resolution and increased penetration (between 4–10 MHz). In our cohort of nearly 1,400 LUS video clips, most LUS clips were obtained at 12 MHz with occasional use of lower frequencies of 10 or 8 MHz due to body habitus. Lower frequency tends to increase the conspicuousness of B-line artifacts but reduces spatial resolution and sensitivity for identifying pleural irregularity at the pleural surface.[30,31] Across our current and prior LUS-ILD cohorts, we found that B-line artifacts were variably present in ILD-affected lung, whereas directly visualizing nonartifactual pleural irregularity or granularity was more consistently present (Supplemental Figure 1). B-line artifact presence and intensity varies with machine settings such as frequency and can be dramatically reduced with common modern imaging processing algorithms designed to eliminate ultrasound artifacts such as compound spatial imaging.[32] These factors led us to exclude B-lines specifically from our interpretation criteria in the prior LUS-ILD-20 and current LUS-ILD-24.

Improving ILD detection continues to be an important issue for pulmonologists, rheumatologists, and radiologists alike. Low dose CT imaging techniques are one tool that has been proposed.[33,34] Although this reduces the risks associated with ionizing radiation, this technique remains costly, with a high environmental impact, limited availability, and with a reduced sensitivity for ILD detection compared to standard radiation dose protocols.[34]

The LUS-ILD-24 criteria may serve as a springboard for incorporation of LUS into a next generation ILD detection algorithm. The results of this study suggest LUS-ILD-24 would perform well as a highly sensitive initial screen for patients in the possible ILD group with only patients with LUS+ referred on for CT imaging for diagnostic confirmation. We can evaluate this hypothetical scenario post hoc on the possible ILD group, including patients excluded from the primary analysis, 46 patients in total. In this scenario using Reader 1 (RMF) results, 21 CT scans could have been avoided due to LUS– results. Of the patients identified as LUS+, CT imaging would reveal 16 new ILD diagnoses, five patients with infection or aspiration, and four patients without any lung pathology. No CTD-ILD would have been missed. Although false positive LUS would be seen in the five patients with infection and aspiration, identifying these etiologies on the triggered CT imaging would likely also be of clinical use. Reader 2 (DAM) and consensus readers would have missed one patient with incident mild atypically apical ILD. In the analyzed cohort with LUS and PFT-possible ILD (Supplementary Table 4), combining LUS with PFT would lead to no ILD being missed. Overall, this suggests incorporation of LUS with or without PFTs as a screening test can greatly reduce CT imaging burden in this patient population and is highly sensitive for ILD.

To incorporate LUS into clinical practice, efforts to disseminate LUS technique and build a critical mass of experience with this technique will be essential. Rheumatologists are the primary point of contact for CTD patients at risk for ILD and the increasing ability of rheumatology ultrasonographers to perform a variety of diagnostic procedures makes ILD assessment by LUS feasible. In many areas of the world, ultrasound may be the only imaging tool available to health care providers,[10,35] and the LUS-ILD-24 provides a simple set of criteria to detect ILD by LUS that can be performed by providers across any discipline. Furthermore, the development of deep-learning algorithms to assist in human LUS interpretation may help increase implementation of this technique.

The strengths of this study are the large sample size, a short CT to LUS interval, parallel PFT and lung texture analysis quantification using CALIPER, and inclusion of patients with IM in the validation cohort. These factors provide significant power to assess sensitivity, specificity, and correlations between LUS severity with PFTs and CALIPER indices and provide support for the construct and criterion validity of this technique. Additionally, assessing interpretation with three readers and several reading paradigms reinforces the robustness of our results. Limitations to this study include the fact that both the pilot and current validation studies were performed at a single center, which may reduce generalizability. A recent external study using our LUS-ILD-20 in patients with SSc-ILD found similar results with sensitivity of 91% and specificity of 86% for ILD detection,[29] but further external validation using LUS-ILD-24 in a large cohort would be useful. We did not collect patient-reported outcomes for further correlation with severity, nor did we obtain longitudinal imaging to assess sensitivity to change over time, although these studies are currently ongoing. CALIPER and PFT results were not available in all patients.

In conclusion we found excellent sensitivity and specificity for our revised LUS-ILD-24 interpretation criteria in this validation study in patients with SSc and IM. LUS-ILD-24 performed similarly well in cohorts with SSc and IM, as well as incident ILD detection. Our findings provide support for including LUS-ILD-24 in the assessment of ILD in patients with SSc and IM, with potential for use at baseline CTD diagnosis, follow-up screening, and evaluation for new suspected ILD. Correlations between LUS-ILD-24 severity and both PFT indices and CT imaging severity opens the door for future studies to determine if LUS monitoring may have a role in assessing treatment response.

ACKNOWLEDGMENTS

A large language model was used solely for checking spelling and to help format references for this article.

AUTHOR CONTRIBUTIONS

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

ROLE OF THE STUDY SPONSOR

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

REFERENCES

1. Gutsche M, Rosen GD, Swigris JJ. Connective tissue disease-associated interstitial lung disease: a review. Curr Respir Care Rep 2012;1(4):224–232.
2. Joy GM, Arbiv OA, Wong CK, et al. Prevalence, imaging patterns and risk factors of interstitial lung disease in connective tissue disease: a systematic review and meta-analysis. Eur Respir Rev 2023;32(167): 220210.
3. Molberg Ø, Hoffmann-Vold AM. Interstitial lung disease in systemic sclerosis: progress in screening and early diagnosis. Curr Opin Rheumatol 2016;28(6):613–618.
4. Steen VD, Medsger TA. Changes in causes of death in systemic sclerosis, 1972–2002. Ann Rheum Dis 2007;66(7):940–944.
5. Johnson SR, Bernstein EJ, Bolster MB, et al. 2023 American College of Rheumatology (ACR)/American College of Chest Physicians (CHEST) guideline for the screening and monitoring of interstitial lung

disease in people with systemic autoimmune rheumatic diseases. Arthritis Rheumatol 2024;76(8):1201–1213.
6. Khanna SA, Nance JW, Suliman SA. Detection and monitoring of interstitial lung disease in patients with systemic sclerosis. Curr Rheumatol Rep 2022;24(5):166–173.
7. Suliman YA, Dobrota R, Huscher D, et al. Brief report: pulmonary function tests: high rate of false-negative results in the early detection and screening of scleroderma-related interstitial lung disease. Arthritis Rheumatol 2015;67(12):3256–3261.
8. Hoffmann-Vold AM, Maher TM, Philpot EE, et al. The identification and management of interstitial lung disease in systemic sclerosis: evidence-based European consensus statements. Lancet Rheumatol 2020;2(2):e71–e83.
9. McAlister S, McGain F, Petersen M, et al. The carbon footprint of hospital diagnostic imaging in Australia. Lancet Reg Health West Pac 2022;24:100459.
10. Hricak H, Ward ZJ, Atun R, et al. Increasing access to imaging for addressing the global cancer epidemic. Radiology 2021;301(3):543–546.
11. Bernstein EJ, Khanna D, Lederer DJ. Screening high-resolution computed tomography of the chest to detect interstitial lung disease in systemic sclerosis: a global survey of rheumatologists. Arthritis Rheumatol 2018;70(6):971–972.
12. Bierig SM, Jones A. Accuracy and cost comparison of ultrasound versus alternative imaging modalities, including CT, MR, PET, and angiography. J Diagn Med Sonogr 2009;25(3):138–144.
13. Picano E, Mangia C, D'Andrea A. Climate change, carbon dioxide emissions, and medical imaging contribution. J Clin Med 2022;12(1):215.
14. Fairchild R, Chung M, Yang D, et al. Development and assessment of novel lung ultrasound interpretation criteria for the detection of interstitial lung disease in systemic sclerosis. Arthritis Care Res (Hoboken) 2021;73(9):1338–1342.
15. Ferro F, Delle Sedie A. The use of ultrasound for assessing interstitial lung involvement in connective tissue diseases. Clin Exp Rheumatol 2018;36 Suppl 114(5):165–170.
16. Vega-Fernandez P, Ting TV, Mar DA, et al. Lung ultrasound in children with systemic juvenile idiopathic arthritis-associated interstitial lung disease. Arthritis Care Res (Hoboken) 2023;75(5):983–988.
17. Watanabe S, Yomono K, Yamamoto S, et al. Lung ultrasound in the assessment of interstitial lung disease in patients with connective tissue disease: performance in comparison with high-resolution computed tomography. Mod Rheumatol 2024;35(1):79–87.
18. Ha YJ, Lee YJ, Kang EH. Lung involvements in rheumatic diseases: update on the epidemiology, pathogenesis, clinical features, and treatment. BioMed Res Int 2018;2018:6930297.
19. Silver KC, Silver RM. Management of systemic-sclerosis-associated interstitial lung disease. Rheum Dis Clin North Am 2015;41(3):439–457.
20. Jacob TM, Matin TN, Jacob J. Imaging approach to interstitial lung disease. In: Cottin V, Richeldi L, Brown K, et al, eds. Orphan Lung Diseases: A Clinical Guide to Rare Lung Disease. Springer International Publishing; 2023:505–523. https://doi.org/10.1007/978-3-031-12950-6_29
21. Huang Y, Liu T, Huang S, et al. Screening value of lung ultrasound in connective tissue disease related interstitial lung disease. Heart Lung 2023;57:110–116.
22. Song G, Bae SC, Lee YH. Diagnostic accuracy of lung ultrasound for interstitial lung disease in patients with connective tissue diseases: a meta-analysis. Clin Exp Rheumatol 2016;34(1):11–16.
23. Yan JH, Pan L, Gao YB, et al. Utility of lung ultrasound to identify interstitial lung disease: an observational study based on the STROBE guidelines. Medicine (Baltimore) 2021;100(12):e25217.
24. Fairchild RM, Horomanski A, Mar DA, et al. Prevalence and significance of pulmonary disease on lung ultrasonography in outpatients with SARS-CoV-2 infection. BMJ Open Respir Res 2021;8(1):e000947.
25. Delle Sedie A, Terslev L, Bruyn GAW, et al; OMERACT US Task Force on Lung US in Rheumatology. Standardization of interstitial lung disease assessment by ultrasound: results from a Delphi process and web-reliability exercise by the OMERACT ultrasound working group. Semin Arthritis Rheum 2024;65:152406.
26. Gutierrez M, Soto-Fajardo C, Pineda C, et al. Ultrasound in the assessment of interstitial lung disease in systemic sclerosis: a systematic literature review by the OMERACT ultrasound group. J Rheumatol 2020;47(7):991–1000.
27. Zavaletta VA, Bartholmai BJ, Robb RA. High resolution multidetector CT-aided tissue analysis and quantification of lung fibrosis. Acad Radiol. 2007;14(7):772–787.
28. Randolph JJ. Free-marginal multirater kappa (multirater κfree): an alternative to Fleiss Fixed-marginal multirater kappa. Paper presented at: University of Joensuu Learning and Instruction Symposium; October 14–15, 2005; Joensuu, Finland.
29. Mohammad Reza Beigi D, Pellegrino G, Loconte M, et al. Lung ultrasound compared to computed tomography detection and automated quantification of systemic sclerosis-associated interstitial lung disease: preliminary study. Rheumatology (Oxford) 2024;63(5):1240–1245.
30. Dietrich CF, Mathis G, Blaivas M, et al. Lung B-line artefacts and their use. J Thorac Dis 2016;8(6):1356–1365.
31. Soldati G, Demi M, Inchingolo R, et al. On the physical basis of pulmonary sonographic interstitial syndrome. J Ultrasound Med 2016;35(10):2075–2086.
32. Lin DC, Nazarian LN, O'Kane PL, et al. Advantages of real-time spatial compound sonography of the musculoskeletal system versus conventional sonography. AJR Am J Roentgenol 2002;179(6):1629–1631.
33. Ley S, Fidler L, Schenk H, et al. Low dose computed tomography of the lung for detection and grading of interstitial lung disease: a systematic simulation study. Pulmonology 2021;27(1):14–25.
34. Martini K, Jungblut L, Sartoretti T, et al. Impact of radiation dose on the detection of interstitial lung changes and image quality in low-dose chest CT – assessment in multiple dose levels from a single patient scan. Eur J Radiol 2023;166:110981.
35. Lanza GM. Ultrasound imaging: something old or something new? Invest Radiol 2020;55(9):573–577.

Arthritis Care & Research
Vol. 77, No. 11, November 2025, pp 1393–1394

American College of Rheumatology
Empowering Rheumatology Professionals

LETTER

DOI 10.1002/acr.25538

Withholding acetaminophen for older adults with osteoarthritis? Not so fast: comment on the article by Kaur et al

To the Editor:

Kaur et al report that acetaminophen use is associated with a significantly higher risk of gastrointestinal, cardiovascular, and renal adverse effects in older adults with osteoarthritis.[1] The reported associations are likely confounded by the higher prevalence in the acetaminophen group of the use of opioids (70% vs 30%), aspirin (40% vs 20%), and nonsteroidal anti-inflammatory drugs (NSAIDs; 31% vs 12%), drugs known to be associated with these adverse effects.[2,3] Although use of the inverse probability of treatment weighting (IPTW) approach allows the creation of a balanced cohort for the estimation of unbiased average treatment effects,[4] the above imbalances in the use of opioids, aspirin, and NSAIDs were observed in the IPTW cohort.[1]

Because most practicing clinicians are unaware that IPTW cohorts are synthetic or artificial[4] and many might be uncomfortable with studies based on a population of pseudopatients, greater transparency is needed in the reporting of IPTW studies. In addition to clearly mentioning it in the Abstract and Methods sections, the use of subheadings such as "Findings from a synthetic population" may also more visibly increase transparency. This would be akin to subheadings such as "Findings from a randomized trial" or "Findings from a cross-sectional study."

The 2019 American College of Rheumatology/Arthritis Foundation osteoarthritis guideline recommends that for patients with osteoarthritis with intolerance of or contraindications to NSAIDs, acetaminophen may be appropriate for short-term and episodic use.[5] This is especially true for older adults, most of whom are unable to tolerate NSAIDs.[6] Clinicians should not be advised to withhold acetaminophen from older adults with osteoarthritis based on the findings of this study.

Drs Ahmed and Faselis's work was supported in part by the US Department of Veterans Affairs (grant 1I01HX003552).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25538.

Ali Ahmed, MD, MPH
ali.ahmed@va.gov
Washington VA Medical Center
George Washington University
and Georgetown University
Washington, DC

Gail Kerr, MD, FRCP(Edin), MACR
Washington VA Medical Center
and Howard University
Washington, DC
Hans J. Moore, MD
Washington VA Medical Center
George Washington University
Georgetown University
and Uniformed Services University
Washington, DC
Janani Rangaswami, MD
Washington VA Medical Center
and George Washington University
Washington, DC
Charles Faselis, MD
Washington VA Medical Center
George Washington University
and Uniformed Services University
Washington, DC
Richard M. Allman, MD
Wake Forest University
Winston-Salem, NC
University of Alabama at Birmingham
and George Washington University
Washington, DC

1. Kaur J, Nakafero G, Abhishek A, et al. Incidence of side effects associated with acetaminophen in people aged 65 years or more: a prospective cohort study using data from the Clinical Practice Research Datalink. Arthritis Care Res (Hoboken) 2025;77(5):666–675. https://doi.org/10.1002/acr.25471
2. Solomon DH, Rassen JA, Glynn RJ, et al. The comparative safety of analgesics in older adults with arthritis. Arch Intern Med 2010;170(22):1968–1976.
3. Mahady SE, Margolis KL, Chan A, et al. Major GI bleeding in older persons using aspirin: incidence and risk factors in the ASPREE randomised controlled trial. Gut 2021;70(4):717–724.
4. Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. Stat Med 2015;34(28):3661–3679.
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 Care Res (Hoboken) 2020;72(2):149–162.
6. O'Neil CK, Hanlon JT, Marcum ZA. Adverse effects of analgesics commonly used by older adults with osteoarthritis: focus on non-opioid and opioid analgesics. Am J Geriatr Pharmacother 2012;10(6):331–342.

DOI 10.1002/acr.25535

Reply

To the Editor:

We thank Ahmed et al for their interest in this study. We conducted a retrospective cohort study and reported that acetaminophen prescription was associated with a significantly higher risk of gastrointestinal, cardiovascular, and renal adverse effects in adults ≥65 years of age. We acknowledged the concerns raised by Ahmed et al regarding potential confounders, such as the use of opioids, aspirin, and nonsteroidal anti-inflammatory drugs, which are known to be associated with these adverse effects. To mitigate the effect of confounders, including those mentioned by Ahmed et al, we used different analytical approaches: (1) age, sex, and practice-matched Cox regression; (2) multivariable Cox regression; (3) propensity score (PS) matching with Cox regression; and (4) inverse probability of treatment weighting (IPTW) using PS with Cox regression.

As shown in the article, the magnitude of risk estimates reduced when additional covariates and adjustment techniques that can account for confounding were used, although we agree residual confounding could still be present, as acknowledged in the article.[1] Because we undertook a range of analyses, including IPTW (which includes the entire eligible cohort), we disagree that this study should be described as having "Findings from a synthetic population," and we have not come across this description previously.

We agree with Ahmed et al that "Clinicians should not be advised to withhold acetaminophen from older adults with osteoarthritis based on the findings of this study," and we have not suggested this as part of the recommendations of our study. Our study offers evidence on the risks of adverse events associated with oral acetaminophen use in adults aged ≥65 years, including those with osteoarthritis, and is in line with previous research.[2–6] We suggest that the findings of this study be considered by guideline writing groups when recommending first-line analgesia in people with osteoarthritis.

Jaspreet Kaur, BAMS, MPH, PG Diploma, PhD
jaspreet.kaur1@nottingham.ac.ukdrjaspreet@live.co.uk
Abhishek Abhishek, MBBS, MD (Medicine), FRCP (UK), PhD
University of Nottingham
Nottingham, UK
Georgina Nakafero, BScN, MPH, PhD
University of Nottingham
and National Institute for Health and Care Research Biological Research Centre
Nottingham, UK
Christian Mallen, BMedSci, BMBS, MMedSci, MPhil, PhD, FRCGP, FFCI, FFPH
Keele University
Keele, UK
Michael Doherty, MA, MD, FRCP
Weiya Zhang, BMed, MMed, PhD
University of Nottingham
and Pain Centre Versus Arthritis
Nottingham, UK

1. Kaur J, Nakafero G, Abhishek A, et al. Incidence of side effects associated with acetaminophen in people aged 65 years or more: a prospective cohort study using data from the Clinical Practice Research Datalink. Arthritis Care Res (Hoboken) 2025;77(5):666–675. https://doi.org/10.1002/acr.25471
2. Imhof M, Ohmann C, Hartwig A, et al; DUSUK Study Group. Which peptic ulcers bleed? Results of a case-control study. Scand J Gastroenterol 1997;32(2):131–138.
3. Rahme E, Barkun A, Nedjar H, et al. Hospitalizations for upper and lower GI events associated with traditional NSAIDs and acetaminophen among the elderly in Quebec, Canada. Am J Gastroenterol 2008; 103(4):872–882.
4. de Vries F, Setakis E, van Staa TP. Concomitant use of ibuprofen and paracetamol and the risk of major clinical safety outcomes. Br J Clin Pharmacol 2010;70(3):429–438.
5. Spence JD, Grosser T, FitzGerald GA. Acetaminophen, nonsteroidal anti-inflammatory drugs, and hypertension. Hypertension 2022;79(9): 1922–1926.
6. Doherty M, Hawkey C, Goulder M, et al. A randomised controlled trial of ibuprofen, paracetamol or a combination tablet of ibuprofen/paracetamol in community-derived people with knee pain. Ann Rheum Dis 2011;70(9):1534–1541.