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 75 No. 10 October 2023

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 75 • OCTOBER 2023 • NO. 10

Cover image: The image on the cover (from Donohue et al, page 2036), shows mitral valve pathology: hydroethidine stain of mitral valve (10x). High-power view of surface fibrin with mixed inflammatory cells.

Arthritis Care & Research
Vol. 75, No. 10, October 2023, pp 2033–2035
DOI 10.1002/acr.25126

American College of Rheumatology
Empowering Rheumatology Professionals

EDITORIAL

UnDRESSing Systemic Juvenile Idiopathic Arthritis Lung Disease

Edward M. Behrens

One of the current pressing issues facing pediatric rheumatologists is obtaining a better understanding of the cause and treatment of systemic juvenile idiopathic arthritis lung disease (JIA-LD). Systemic JIA-LD can be a progressive, fatal complication with features of endogenous lipoid pneumonia, pulmonary alveolar proteinosis, and interstitial lung disease. Because our perception that this complication of systemic JIA was previously unheard of prior to its first reports in 2013 (1), a very natural hypothesis that some new environmental exposure must be generating a new complication of an old disease has been forwarded. An obvious candidate, the use of interleukin 1 (IL-1) blocking and other biologic agents for the treatment of systemic JIA, which began in the mid-2000s (2), has been proposed.

Some of the initial literature focusing on systemic JIA-LD has suggested that the complication is associated with eosinophilia, reports of drug intolerance symptoms, and HLA–DRB1*15:01 carriage (3–5). Taken together, these findings have been suggested to represent a drug reaction with eosinophilia and systemic symptoms (DRESS) syndrome that leads to lung inflammation and progression of systemic JIA-LD. Yet, many features of systemic JIA-LD are not reminiscent of DRESS syndrome, and many of the features used by the Registry of Severe Cutaneous Adverse Reaction (RegiSCAR) criteria for DRESS syndrome overlap with the exact features present in active systemic JIA. This raises the specter that DRESS syndrome may not be the best explanation for systemic JIA-LD, and thus alternatively the cytokine plasticity hypothesis has been offered. This hypothesis posits that cytokine blockade in systemic JIA redirects cytokine mediated CD4 T cell fate choices in a manner that leads to systemic JIA-LD (6). Importantly, direct evidence for either DRESS syndrome or cytokine plasticity as a causal factor is lacking. The implications are not trivial, as it may change choice of therapy, or even starting a therapy based on HLA risk stratification. If the hypothesis that a drug-linked causal factor is not correct, an inappropriate choice might be made, denying children an otherwise effective therapy.

Importantly, some implications of these hypotheses might be testable through observational data. In the case of the DRESS syndrome hypothesis, one would predict that eosinophilia should temporally follow drug exposure. Likewise, if DRB1*15:01 is the risk allele for such a reaction, carriage rates in the systemic JIA-LD drug-exposed populations should be higher than in the non-systemic JIA-LD population. Articles by Lerman et al (7) and Womba et al (8), included in this issue of *Arthritis Care & Research*, test these hypotheses using retrospective data and conclude that neither of these predicted observations are true. Lerman and colleagues show that eosinophilia precedes drug therapy in 43% of patients with systemic JIA (7). This suggests eosinophilia is a feature of the disease rather than drug reaction. The authors point out that many of the RegiSCAR criteria for DRESS syndrome overlap with features that are part of systemic JIA itself. It appears eosinophilia is yet an another one of these features. In agreement with this observation, Wobma et al find that eosinophilia is common and equally prevalent in systemic and nonsystemic JIA patients using IL-1 or IL-6 inhibitors and that eosinophilia may be a marker of more severe disease (8). These data are in keeping with the idea that systemic inflammation may be responsible for, or at least associated with, eosinophilia rather than acting as a marker of drug reaction. Since patients with more severe systemic disease are more likely to be treated with biologic therapy, these factors will be highly confounded with each other.

The DRB1*15:01 major histocompatibility complex (MHC) allele has also been implicated as a potential causal mechanism for systemic JIA-LD. This possibility was also explored in the aforementioned studies. Wobma et al find that the MHC haplotype does not statistically associate with eosinophilia (8). Furthermore, Lerman et al show that DRB1*15:01 is enriched in systemic JIA more broadly, perhaps as a risk factor for severe disease, but not

Edward M. Behrens, MD: Children's Hospital of Philadelphia and University of Pennsylvania, Philadelphia.

Author disclosures are available online at https://onlinelibrary.wiley.com/doi/10.1002/acr.25126.

Address correspondence via email to Edward M. Behrens, MD, at behrens@email.chop.edu.

Submitted for publication March 27, 2023; accepted in revised form April 4, 2023.

necessarily associated with systemic JIA-LD or drug reaction per se (7). Thus, as with eosinophilia, the fact that DRB1*15:01 was originally associated with systemic JIA-LD may simply be due to its association with severe systemic JIA more broadly rather than a lung disease–specific factor.

It should be noted that falsifying the DRESS syndrome hypothesis does not imply the cytokine plasticity hypothesis. For cytokine plasticity to be a useful construct, it needs to make similar testable predictions that can or cannot be falsified. For instance, we might expect to see different proportions of CD4 T cell populations prior to and after treatment with cytokine blockade, and that these population differences are systematically different in those with systemic JIA-LD. Such an analysis would require a difficult prospective study to collect biospecimens. Moreover, it is difficult to conceive of retrospective data that might allow us to test cytokine plasticity.

However, considering other explanations beyond these 2 competing hypotheses may prove fruitful. A driving motivator for both hypotheses is the idea that lung disease is a new phenomenon. Initially, it seems implausible that we could have missed such a severe complication in our patients, so this must be an emergent complication. Yet, if the data from Womba et al and Lerman et al are to be believed, it seems we have missed eosinophilia as a common systemic JIA complication for many years. Careful search reveals reports of restrictive pulmonary disease in systemic JIA as far back as 1997 (9). A 2001 report discusses a systemic JIA patient with fatal lung disease whose biopsy showed intraalveolar cholesterol granulomas, a pathology reminiscent of the endogenous lipoid pneumonia of systemic JIA-LD (10). This was well before IL-1 or IL-6 blockade was used to treat patients. A complete search of the early literature becomes increasingly difficult because the nomenclature was not precisely established, but the earliest report the author of this editorial could locate is a 1976 review of 100 juvenile rheumatoid arthritis patients followed for 15 years that noted that within the "Acute Febrile Onset/Still's" subset, "chronic pulmonary fibrosis is rarely observed" (11). Thus, the rare, but present complication of lung disease in systemic JIA was noted almost 30 years prior to the use of IL-1 blockade in the disease. Therefore, it seems possible that some form of systemic JIA-LD was present prior to biologic therapies and perhaps may have been a missed complication before it began receiving increased attention.

In a disease without any specific pathognomonic identifying features, one possibility is that the entity we are calling systemic JIA has actually been in flux. Systemic JIA may represent a heterogenous collection of fever syndromes including "canonical Still's disease," but also some syndromes that may occur in a spatiotemporal window, with some phenotypes arising and then fading away, only to be replaced with new phenotypes similar enough to meet criteria but representing different pathology. Viral infections may play a part in triggering a spatiotemporally restricted fever syndrome, as we have recently seen with Multisystem Inflammatory Syndrome in Children (MIS-C) and SARS-CoV-2. We know that systemic JIA-LD is associated with IL-18 elevations, yet not all patients meeting systemic JIA criteria have elevated IL-18 (12), clearly showing heterogeneity within the broad umbrella of systemic JIA. Indeed, even in considering elements of the International League of Associations for Rheumatology (ILAR) criteria, serositis is only found in 10% of patients (13), which suggests significant heterogeneity. Thus, what we are calling systemic JIA-LD may indeed be its own entity, perhaps revealed by an emergent environmental trigger, that we are lumping into the broad umbrella of systemic JIA by virtue of its ability to overlap with the ILAR criteria.

To this end, it may be a mistake to focus on drugs as an inciting trigger, as there are a host of emergent and changing environmental triggers that may have occurred in the last 15 years, leading to systemic JIA-LD syndrome. A premature fixation on cytokine blockers as the culprit may have the undesirable effect of reducing the availability of these efficacious drugs to this population (14,15). There is at least 1 case report of cytokine blockade improving systemic JIA-LD (16), which further complicates the picture. Once we consider a broader source of emergent triggers, the list of possible pathogenic models expands well beyond DRESS syndrome and cytokine plasticity. We need more robust immunologic characterization of the systemic JIA-LD syndrome to make headway in this regard. We are already beginning to see these types of studies, such as the report that serum levels of the adhesion molecule ICAM-5 distinguishes systemic JIA-LD from both macrophage activation syndrome and systemic JIA without lung disease (3). Better characterization of the humoral and cellular biomarkers associated with systemic JIA-LD in both blood and lung are needed to make these advances. Fortunately, we are in an era of increasing availability of high-dimensional, multimodal techniques that should allow us to make this type of progress. Coordination with national, multi-institution research collaboratives will be essential to maintain the logistics required for such a rare disease. Such work may need to be incorporated into empiric clinical trials of therapy, as the need to find clinical solutions is great and may not be able to wait for a full pathogenic model to be complete. We are fortunate to have as large a cohort of well-trained, rigorous, and passionate researchers steeped in quantitative methods in pediatric rheumatology as we have had in any era. Now is the time for us to step up to the task of solving these problems for our patients.

AUTHOR CONTRIBUTIONS

Dr. Behrens drafted the article, revised it critically for important intellectual content, and approved the final version to be published.

REFERENCES

1. Kimura Y, Weiss JE, Haroldson KL, et al. Pulmonary hypertension and other potentially fatal pulmonary complications in systemic juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2013;65:745–52.

2. Allantaz F, Chaussabel D, Stichweh D, et al. Blood leukocyte microarrays to diagnose systemic onset juvenile idiopathic arthritis and follow the response to IL-1 blockade. J Exp Med 2007;204:2131–44.
3. Chen G, Deutsch GH, Schulert GS, et al. Identification of distinct inflammatory programs and biomarkers in systemic juvenile idiopathic arthritis and related lung disease by serum proteome analysis. Arthritis Rheumatol 2022;74:1271–83.
4. Saper VE, Chen G, Deutsch GH, et al. Emergent high fatality lung disease in systemic juvenile arthritis. Ann Rheum Dis 2019;78:1722–31.
5. Saper VE, Ombrello MJ, Tremoulet AH, et al. Severe delayed hypersensitivity reactions to IL-1 and IL-6 inhibitors link to common HLA-DRB1*15 alleles. Ann Rheum Dis 2022;81:406–15.
6. Binstadt BA, Nigrovic PA. The conundrum of lung disease and drug hypersensitivity-like reactions in systemic juvenile idiopathic arthritis [review]. Arthritis Rheumatol 2022;74:1122–31.
7. Lerman A, Mahmud S, Alfath Z, et al. HLA DRB1*15 and eosinophilia are common among patients with systemic juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2023;75:2082–7.
8. Wobma H, Arvila S, Taylor M, et al. Incidence and risk factors for eosinophilia and lung disease in biologic-exposed children with systemic juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2023;75:2063–72.
9. Van der Net J, Kuis W, Prakken AB, et al. Correlates of disablement in systemic onset juvenile chronic arthritis. A cross sectional study. Scand J Rheumatol 1997;26:188–96.
10. Schultz R, Mattila J, Gappa M, et al. Development of progressive pulmonary interstitial and intra-alveolar cholesterol granulomas (PICG) associated with therapy-resistant chronic systemic juvenile arthritis (CJA). Pediatr Pulmonol 2001;32:397–402.
11. Calabro JJ, Holgerson WB, Sonpal GM, et al. Juvenile rheumatoid arthritis: a general review and report of 100 patients observed for 15 years. Semin Arthritis Rheum 1976;5:257–98.
12. Shimizu M, Nakagishi Y, Yachie A. Distinct subsets of patients with systemic juvenile idiopathic arthritis based on their cytokine profiles. Cytokine 2013;61:345–8.
13. Behrens EM, Beukelman T, Gallo L, et al. Evaluation of the presentation of systemic onset juvenile rheumatoid arthritis: data from the Pennsylvania Systemic Onset Juvenile Arthritis Registry (PASOJAR). J Rheumatol 2008;35:343–8.
14. De Benedetti F, Brunner HI, Ruperto N, et al. Randomized trial of tocilizumab in systemic juvenile idiopathic arthritis. N Engl J Med 2012; 367:2385–95.
15. Ruperto N, Brunner HI, Quartier P, et al. Two randomized trials of canakinumab in systemic juvenile idiopathic arthritis. N Engl J Med 2012;367:2396–406.
16. Rood JE, Rezk A, Pogoriler J, et al. Improvement of refractory systemic juvenile idiopathic arthritis-associated lung disease with single-agent blockade of IL-1beta and IL-18. J Clin Immunol 2023;43: 101–8.

Arthritis Care & Research
Vol. 75, No. 10, October 2023, pp 2036–2043
DOI 10.1002/acr.25138

American College of Rheumatology
Empowering Rheumatology Professionals

CLINICOPATHOLOGIC CONFERENCE

Triphasic: Preeclampsia, Systemic Lupus Erythematosus, and Severe Neutropenia With Use of Granulocyte Colony Stimulating Factor in the Partum and Postpartum Period

Sarah Donohue, Shelby Gomez, Tripti Singh, and Shivani Garg

CASE PRESENTATION

History of present illness

A 28-year-old female who self identifies as White, non-Hispanic ethnicity with a diagnosis of childhood-onset systemic lupus erythematosus (SLE) presents with acute onset hypertension, lower extremity edema, shortness of breath, and headache during her second trimester of pregnancy.

Past medical history

The patient was diagnosed at age 8 years in the setting of thrombocytopenia, hemolytic anemia, hypocomplementemia, elevated double-stranded DNA (dsDNA) antibody, anti-Ro/SSA positivity, and triple antiphospholipid antibody (aPL) panel positivity. At age 16 years, she developed hypertension, proteinuria, and hematuria and underwent renal biopsy that demonstrated class II lupus nephritis (LN). Additionally, in adolescence she was found to have mild mitral valve regurgitation stable on annual echocardiogram and previously followed by a pediatric cardiologist. Treatment for her SLE included monthly pulse-dose glucocorticoids, mycophenolate mofetil 1,500 mg twice a day and hydroxychloroquine (HCQ). Overall, SLE was clinically stable, but she continued to have serologic activity as noted by elevated dsDNA, hypocomplementemia, and elevated inflammatory markers. Prior to pregnancy, she was maintained on mycophenolate mofetil 1,500 mg twice a day, HCQ, lisinopril, and nifedipine until preconception counseling (Table 1).

Prepartum and partum stage

At age 28 years she underwent preconception counseling, and her mycophenolate mofetil was changed to azathioprine while HCQ was maintained and low-dose aspirin was started. Additionally, lisinopril was discontinued and the antihypertensive regimen was converted to monotherapy with nifedipine. Given the anti-Ro/SSA positivity, she was followed by Maternal-Fetal Medicine and underwent serial fetal cardiac ultrasounds from weeks 16 to 24 without evidence of a heart block. The first trimester of pregnancy was uncomplicated, with stable blood pressure and SLE clinical and serologic markers. During the second trimester, while creatinine and the glomerular filtration rate (GFR) remained stable, urinalysis revealed mild hematuria with a spot urine:protein:creatinine (UPC) ratio of 0.3 (Table 1). At 25 weeks gestation, she developed persistent hypertension, despite antihypertensive therapy with nifedipine, which was up-titrated with subsequent addition of labetalol. At 26 weeks gestation, given worsening proteinuria (24-hour urine protein of 1.3 grams) and precipitous thrombocytopenia (platelets 23,000/μl) (Table 1), her local health care team, including a rheumatologist, nephrologist, and Maternal-Fetal Medicine, were concerned for SLE flare versus preeclampsia and initiated a course of high-dose intravenous (IV) solumedrol 500 mg for 3 days, followed by oral prednisone 60 mg daily for 10 days, and subcutaneous enoxaparin 40 mg daily for deep vein thrombosis prophylaxis, given positive aPL and proteinuria in the second trimester of pregnancy. A course of IV immunoglobulin (IVIG) 1 gram/kg was administered for 2 days as treatment of thrombocytopenia.

First hospital admission

At 27 weeks gestation she presented to her outpatient obstetrician with acute onset of severe lower-extremity edema, shortness of breath, and headache requiring hospital admission.

Sarah Donohue, MD, Shelby Gomez, PharmD, Tripti Singh, MD, Shivani Garg, MD, MS: University of Wisconsin, Madison.
Author disclosures and graphical abstracts are available online at https://onlinelibrary.wiley.com/doi/10.1002/acr.25138.

Address correspondence via email to Sarah Donohue, MD, at sdonohue@uwhealth.org.
Submitted for publication October 13, 2022; accepted in revised form April 18, 2023.

Physical examination. Vital signs showed blood pressure 160/110 and oxygenation saturation of 88% on room air. She was noted to have a cushingoid appearance without malar rash or nasal or oral ulcers. Diminished breath sounds were noted in bilateral lung bases. Lower extremities demonstrated 3+ pitting edema without overt synovitis, and there was no evidence of livedo reticularis on dermatologic examination.

Laboratory evaluation from first admission. On admission she was noted to have acute kidney injury, with peak creatinine 4.5 mg/dl (baseline creatinine 0.74 mg/dl), 24-hour urine protein collection of 1.3 grams, hematuria, acute microangiopathic hemolytic anemia with hemoglobin 7.7 grams/dl, and worsening thrombocytopenia. Uric acid level was elevated up to 10.6 mg/dl (reference range 2.3–6.3 mg/dl) (Table 1) without evidence of a hepatocellular transaminitis, as aspartate aminotransferase (AST) was 60 U/liter and alanine aminotransferase (ALT) 52 U/liter.

Table 1. Summary of prepartum and partum stage, including first hospitalization*

Prepregnancy medications	Mycophenolate mofetil 1,500 mg BID; HCQ 200 mg BID; lisinopril 10 mg daily; nifedipine| 90 mg BID
Prepregnancy laboratory results	Stable cell counts; creatinine 0.74 mg/dl, GFR 92 ml/minute/1.73 m^2, UPC 0.3, UA with 0–2 RBCs and 3–5 WBCs; C3 ↓74 mg/dl, C4 ↓10.6 mg/dl, dsDNA ↑473 IU/ml, positive anti-Ro/SSA antibodies; anticardiolipin IgG ↑65 CU, anticardiolipin IgM <1.0 CU; β_2 glycoprotein IgG ↑64.6 U/ml, β_2 glycoprotein IgM <9.4 U/ml; lupus anticoagulant present
Pregnancy medications	Azathioprine 75 mg qAM, 50 mg qPM; HCQ 200 mg BID; nifedipine 30 mg daily; aspirin 1 mg daily
Obstetrics history	First trimester–uneventful, laboratory results stable, BP 120–130/80s mm Hg; second trimester–BP elevated; platelets ↓57,000/μl; UA revealed proteinuria with UPC 0.3 and hematuria
First admission laboratory results	Platelets ↓23,000/μl, Hgb ↓7.7 grams/dl, creatinine ↑4.5 mg/dl; UA with 2+ proteinuria, >100 RBCs/HPF, 24-hour urinary protein ↑1.3 grams; AST 60 U/liter, ALT 52 U/liter; serum uric acid ↑10.6 mg/dl (range 2.3–6.3 mg/dl); C3 ↓53 mg/dl, C4 ↓6 mg/dl, dsDNA ↑46 IU
First admission treatment	IV solumedrol 1 gram, IV magnesium, IV furosemide; emergency C-section, preterm delivery at 27 weeks gestation; required intubation and mechanical ventilation

* ALT = alanine aminotransferase; AST = aspartate aminotransferase; BID = twice a day; BP = blood pressure; CU = cardiolipin unit; dsDNA = double-stranded DNA; GFR = glomerular filtration rate; HCQ = hydroxychloroquine; Hgb = hemoglobin; HPF = high-powered field; IU = international units; IV = intravenous; qAM = every morning; qPM = every evening; RBC = red blood cell; UA = urinalysis; UPC = urine:protein:creatinine ratio; WBC = white blood cell.

Differential diagnosis: first hospital admission. The primary clinical concern during the first admission was for: 1) severe preeclampsia given the degree of hypertension, proteinuria ≥300 mg/day, a doubling of the creatinine, severe thrombocytopenia, and new onset severe headache, versus 2) SLE/LN flare, as mean 24-hour proteinuria exceeded 2.9 grams, with hypocomplementemia, elevated dsDNA, thrombocytopenia, and microangiopathic hemolytic anemia, versus 3) obstetric aPL antibody syndrome (obstetric-APS) in the setting of aPL positivity, preterm birth <34 weeks with features of severe preeclampsia, and concern for thrombotic microangiopathy.

First hospital course. On first admission, she was treated with pulse-dose IV solumedrol 1 gram, IV magnesium and IV diuretic. Given concern for severe preeclampsia with systolic blood pressure ≥160 mm Hg and diastolic blood pressure ≥110 mm Hg on 2 separate occasions at least 4 hours apart while at rest, severe thrombocytopenia, proteinuria, and cerebral symptoms of severe headache, she underwent emergent cesarean section (1). Following the cesarean section, her symptoms progressed and she developed acute hypoxic respiratory failure, with plain films demonstrating diffuse, bilateral pulmonary infiltrates requiring intubation and mechanical ventilation. No computed tomography (CT) was obtained outside prior to transfer. Following the cesarean section, laboratory results were notable for progression of thrombocytopenia, with nadir to 11,000/μl from 57,000/μl prior to admission. Given the complexity of her care, she was transferred to our facility for further management.

Second hospital admission (interfacility transfer)

She arrived at our facility on postoperative day 1 from emergent cesarean section and remained intubated in the setting of diffuse bilateral airspace disease consistent with acute repository distress syndrome.

Physical examination. At our facility, initial vital statistics were notable for blood pressure at 99/69, heart rate 49, respiratory rate 14, and sating 94% on mechanical ventilation. Physical examination was notable for anasarca, livedo reticularis, and a new holosystolic apical murmur in a mechanically ventilated patient. Other pertinent positive findings included elevated jugular venous pressure and 3+ peripheral edema of the bilateral upper and lower extremities. Examination did not reveal hepatosplenomegaly.

Laboratory evaluation and imaging from second admission. Laboratory results were notable for progressive hemolytic anemia, thrombocytopenia, leucopenia, proteinuria, hypocomplementemia, and elevated dsDNA (Table 2). Imaging included CT of the chest that demonstrated diffuse centrilobular ground-glass, bilateral pleural effusions and body wall anasarca. A

Table 2. Summary of second hospital admission (interfacility transfer)*

Second admission physical examination	Grade III/VI holosystolic murmur at apex, radiates throughout precordium; anasarca with abdominal wall edema; livedo reticularis–bilateral upper extremities
Second admission laboratory results	Anemia, thrombocytopenia, leucopenia; haptoglobin ↓10 mg/dl, LDH ↑519 U/liter, retic count ↑5.5%; GFR ↓35 ml/minute/1.73 m^2, UA with proteinuria and hematuria, UPC ↑1.23 grams; C3 ↓52 mg/dl, C4 ↓5 mg/dl, dsDNA ↑95 IU
Second admission imaging	TTE: new severe mitral stenosis and mitral regurgitation; CT chest: diffuse centrilobular ground-glass, bilateral pleural effusions, body wall anasarca
Second admission treatment	Pulse-dose steroids, IV diuretic and ultimately required hemodialysis; plasma exchange; rituximab; taken to the operative room with cardiothoracic surgery for mitral valve replacement

* CT = computed tomography; dsDNA = double-stranded DNA; GFR = glomerular filtration rate; IV = intravenous; LDH = lactate dehydrogenase; retic = reticulocyte; TTE = transthoracic echocardiogram; UA = urinalysis; UPC = urine:protein:creatinine ratio.

transthoracic echocardiogram (TTE) revealed severe mitral valve stenosis and regurgitation (Table 2). When compared to her previous echocardiograms from 2013, 2014, and 2015, a significant worsening in mitral valve regurgitation was noted and a new, severe mitral valve stenosis was identified, with a rheumatic-appearing mitral valve. No valvular vegetations were identified on TTE. TEE was deferred at this time as the patient was extubated.

Differential diagnosis: second hospital admission. Given her new mitral valve findings, progression of hemolytic anemia, thrombocytopenia, proteinuria, obstetric-APS, and hypocomplementemia, we formulated a differential diagnosis of SLE/LN flare versus thrombotic microangiopathy with coexisting Libman-Sacks endocarditis, considering the degree of severe mitral valve regurgitation. Other differential diagnoses included culture-negative infective endocarditis and lupus valvulitis.

Second hospital course. When she arrived at our facility, she was continued on treatment with pulse-dose steroids with IV solumedrol 1 gram daily until extubated and then transitioned to 40 mg every 8 hours for 10 days, followed by prolonged oral regimen taper and diuretics, and initiated on plasma exchange daily for 3 days. Despite therapy, she continued to have progressive thrombocytopenia and hemolytic anemia requiring an additional course of IVIG for 5 days and ultimately rituximab dosed at 375 mg/m^2 weekly for 4 weeks, with gradual improvement of cell counts. Her prepartum mycophenolate mofetil was resumed (Table 2). Additionally, she was treated with a 7-day course of antibiotic therapy as sputum cultures returned positive for methicillin-sensitive *S. aureus*. Blood cultures were negative and formal bronchoscopy with bronchoalveolar lavage was not pursued at this time. Given ongoing anasarca and acute repository distress syndrome, she ultimately required continuous hemodialysis that was transitioned to intermittent as she stabilized. Due to the severity of her cardiopulmonary disease, cardiothoracic surgery was consulted in the setting of persistent hypoxemia and recurrent flash pulmonary edema, prompting removal to the operating room for mitral valve replacement surgery. Following mechanical mitral valve replacement, she was initiated on lifelong anticoagulation therapy with Warfarin.

Surgical findings and histopathology. Final mitral valve pathology was consistent with Libman-Sacks endocarditis and healed valvulitis (Figure 1). Final mitral valve tissue and fungal cultures were negative. Repeat blood cultures and serum testing for *Bartonella, Coxiella, Mycoplasma hominis, Brucella*, histoplasmosis, and blastomycosis were obtained and were negative. Final 16S rRNA gene sequencing for bacterial identification of the mitral valve was negative. During admission, a renal biopsy was not pursued given the degree of thrombocytopenia, anticoagulation, and ongoing hemodialysis, indicating high bleeding risk. Following acute prolonged hospitalization, her renal function improved, and she no longer required hemodialysis at the time of discharge.

Third hospital admission

Two weeks following discharge, she was readmitted to our facility with acute onset fever and found to have new severe leucopenia (Table 3).

Physical examination. Vital signs were notable for fever to 103.2° F and tachycardia. A sternotomy incision was clear, dry, and intact without drainage. Mechanical heart sounds were appreciated. The abdomen was notable for suprapubic erythema at the Pfannensteil surgical site, with purulent drainage and surrounding induration with tenderness to palpation (Table 3).

Laboratory evaluation and imaging from third hospital admission. Laboratory results demonstrated leukopenia (white blood cells ↓1,100/μl) with neutropenia (absolute neutrophil count [ANC] ↓600/μl), acute kidney injury (creatinine ↑2.08 mg/dl, GFR ↓32 ml/minute/1.73 m^2) with proteinuria (UPC ↑2.33 grams) in addition to elevated ferritin (↑1,250 ng/ml), fibrinogen (↑511 mg/dl), triglycerides (↑295 mg/dl), and a pattern of hepatocellular transaminitis (AST ↑156 U/liter, ALT ↑472 U/liter) (Table 3). CT of the abdomen and pelvis noted inflammatory changes of the skin and subcutaneous fat consistent surrounding the Pfannensteil surgical site (Table 3).

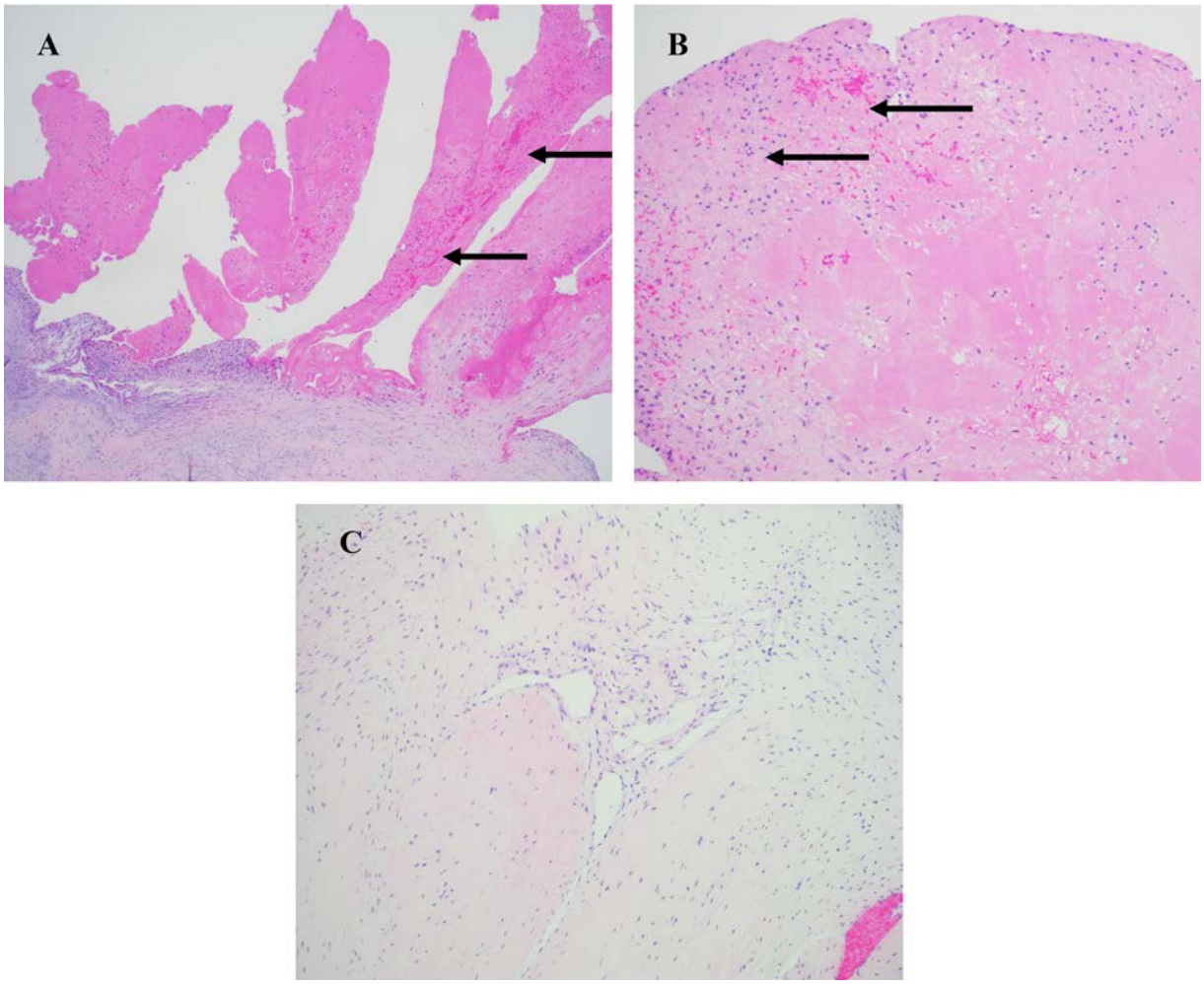

Figure 1. Mitral valve pathology. **A**, Hydroethidine (H&E) stain of mitral valve (10×). High-power view of surface fibrin with mixed inflammatory cells. Arrows indicate fibromyxoid valvular degeneration with healed valvulitis. **B**, H&E stain of mitral valve (100×). Low power showing surface deposition of fibrin. Arrows indicate fibromyxoid valvular degeneration with healed valvulitis. **C**, H&E stain of mitral valve. Areas of valve showing neovascularization, consistent with a healed valvulitis.

Table 3. Summary of third hospital admission*

Third admission physical examination	Febrile to 103.2° F; suprapubic Pfannensteil site with purulent drainage
Third admission laboratory results	WBC ↓1,100/μl, ANC ↓600/μl; creatinine ↑2.08 mg/dl, GFR ↓32 ml/minute/1.73 m^2, UPC ↑2.33 grams; ferritin ↑1,250 ng/ml, fibrinogen ↑511 mg/dl, AST ↑156 U/liter, ALT ↑472 U/liter, TG ↑295 mg/dl
Third admission imaging	CT abdomen/pelvis: inflammatory changes of the anterior abdominal wall concerning for cellulitis
Third admission treatment	G-CSF 300 mcg daily × 2 doses; broad-spectrum IV antibiotics transitioned to oral antibiotics

* ALT = alanine aminotransferase; ANC = absolute neutrophil count; AST = aspartate aminotransferase; CT = computed tomography; G-CSF = granulocyte colony stimulating factor; GFR = glomerular filtration rate; IV = intravenous; TG = triglyceride; UPC = urine:protein:creatinine ratio; WBC = white blood cell.

Differential diagnosis: third hospital admission. The primary clinical concern at the time of the third admission was for severe neutropenic fever in the setting of a surgical site infection in an immunosuppressed patient. However, given laboratory findings of neutropenia, hypertriglyceridemia, and hyperferritinemia a concern for early macrophage activation syndrome/hemophagocytic lymphohistiocytosis was raised. A soluble interleukin-2 was not obtained, given a prolonged turn-around time of approximately 1 week at our institution. Further bone marrow biopsy was deferred by the hematology team given higher concern for neutropenic fever in a setting of infection and immunosuppression.

Hospital course. The patient was treated with IV antibiotics and transitioned to oral antibiotics at the time of discharge (Table 3). Given a suspicion for neutropenic fever, hematology was consulted and recommended initiation of granulocyte colony

stimulating factor (G-CSF) for management of severe leucopenia. She received 2 doses of G-CSF with recovery of the cell count.

Postdischarge follow-up. At 1.5 weeks following discharge and the last dose of G-CSF, the patient presented to our specialty Lupus Nephritis clinic with severe proteinuria as demonstrated by progression from 2.33 grams to 9.97 grams, consistent with SLE/LN flare. Given concern for flare, mycophenolate mofetil dosing was judiciously increased, and oral glucocorticoids were continued with close laboratory and clinical monitoring.

CASE SUMMARY

To summarize, our patient presented with 3 different clinical presentations and 3 different complications of SLE and related diseases. Each time, the treating team faced unique challenges in obtaining the correct diagnosis to guide further management and treatment decision-making, as discussed below:

- **Dilemma one.** Pregnancy complicated by preeclampsia in the setting of aPLs versus an SLE flare, a common dilemma in pregnant patients who are aPL positive, and with SLE presenting with multifactorial features.
- **Dilemma two.** Progressive renal failure with several possible etiologies, including LN flare versus thrombotic microangiopathic renal disease in the setting of newly diagnosed obstetric-APS and concern for Libman-Sacks endocarditis, given the progression of severe mitral valve regurgitation (2).
- **Dilemma three.** Severe neutropenic fever in the setting of surgical site infection in an immunosuppressed patient with SLE. Obstetric-APS promoted the use of G-CSF with the knowledge that G-CSF could induce SLE/LN flare.

DISCUSSION

Pregnancy in patients with existing SLE, LN, and/or APS is prone to complications, including disease flare, preeclampsia, and preterm birth, to name a select few that were demonstrated in our patient's case. Furthermore, active LN during the partum period poses the greatest risk to maternal-fetal complications and requires a multidisciplinary team approach to guide disease management and treatment (3,4). Below, we propose our main take-away points from each dilemma complicating this case.

Dilemma one: prepartum management of patients with SLE and aPL positivity

Pregnancy in the setting of SLE is prone to a 20-fold increased risk in maternal mortality with an odds ratio of 2.4 for preterm birth and 3.0 for preeclampsia (5), and with approximately 25% of patients experiencing a lupus flare during pregnancy (3). As 15% to 30% of SLE patients with a flare during pregnancy had severe disease manifestations commonly involving the kidney and other internal organs, SLE flare during pregnancy associated with poor maternal and fetal outcomes again reiterates the importance of well-controlled SLE/LN prior to conception (6). In a single-center retrospective study of 72 patients, the study authors identified a prevalence of preterm delivery, before 37 weeks of gestation, to be 38% in SLE, as seen in our patient's case. Further, adverse outcomes are associated with active LN and hypertension during the first trimester, and such patients have a higher risk of developing preeclampsia subsequently during the later trimesters (7).

Approximately 30–40% of patients with SLE are positive for aPL antibodies, with approximately 40% of patients with SLE with aPL positivity developing clinical features of APS, and with venous thromboembolism as the most common manifestation (8,9). Pregnancy morbidity in APS includes miscarriage, late intrauterine fetal demise, and severe preeclampsia. Thus all patients with SLE should be screened for aPL positivity during prepartum management, including β_2 glycoprotein IgG and IgM, anticardiolipin antibody IgG and IgM, and lupus anticoagulant (8–11). Key attention should be given to lupus anticoagulant, as it carries the strongest association with thrombotic disease and adverse pregnancy outcomes after 12 weeks of pregnancy as seen in the Predictors of Pregnancy Outcome: Biomarkers in Antiphospholipid Antibody Syndrome and Systemic Lupus Erythematosus study of pregnancy morbidity due to aPL antibodies (12–14).

With regard to treatment of aPL-positive patients without a history of thrombosis during pregnancy, low-dose aspirin should be initiated to reduce the risk of preeclampsia (12). If the patient has a history of obstetric complications, including late fetal demise (>10 weeks), multiple early losses, intrauterine growth restriction, and a history of severe preeclampsia, the treatment of choice includes prophylactic dosing of low molecular weight heparin (dosed twice daily) in addition to low-dose aspirin, which should be initiated as soon as pregnancy is confirmed and switched to unfractioned heparin 4 weeks prior to delivery, with continuation of heparin 6 weeks postpartum (15,16). Finally, if there is a history of a known postthrombotic event, the recommendation is to initiate full-dose heparin and low-dose aspirin (16).

Dilemma two: LN in pregnancy

Patients with well-controlled LN should be transitioned to azathioprine therapy during preconception counseling, and angiotensin converting enzyme inhibitor/angiotensin receptor blocker should be discontinued (11). Our expert panel established that this transition should occur at least 3 months prior to pregnancy to ensure that LN remains controlled (as monitored by urinalysis, urine protein:creatinine ratio, creatinine, and eGFR) following alteration in pharmacotherapy.

As previously described by Smyth et al, active LN poses the greatest risk to the outcome of pregnancy in women with lupus (3). As 15% to 30% of patients with SLE who have a flare during pregnancy had severe disease manifestations, commonly involving the kidney and other internal organs, SLE flare during pregnancy associated with poor maternal and fetal outcomes again reiterates the importance of well-controlled SLE/LN prior to conception (6).

Moreover, among women with SLE, active LN, and serum creatinine of >2.8 mg/dl at the time of conception, up to 85% develop maternal complications, including preeclampsia, 60% develop intrauterine growth restriction, and 70% experience preterm labor (17). Other factors that determine the risk of preeclampsia in patients with SLE include hypertension and thrombocytopenia (18).

Differentiation between preeclampsia and active SLE/LN continues to be a diagnostic quandary, as both may present with progressive proteinuria, declining renal function, hypertension, lower-extremity edema, and thrombocytopenia (19). Preeclampsia is defined as systolic blood pressure ≥140 mm Hg and/or diastolic blood pressure ≥90 mm Hg on at least 2 occasions at least 4 hours apart after 20 weeks of gestation in a previously normotensive patient AND the new onset of 1 or more of the following: proteinuria ≥0.3 grams or dipstick ≥2+ if a quantitative measurement is unavailable, platelet count <100,000/μl, serum creatinine >1.1 mg/dl or doubling of the creatinine concentration (absent other renal disease), liver function tests >2 times the upper limit of normal, pulmonary edema, new-onset and persistent headache not due to alternative diagnoses and not responding to usual analgesic doses, or visual symptoms (1), whereas severe preeclampsia is defined as systolic blood pressure ≥160 mm Hg and/or diastolic blood pressure ≥110 mm Hg on 2 separate occasions at least 4 hours apart while at rest, central nervous system dysfunction with new-onset cerebral or visual disturbances, impaired liver function >2 times the upper limit of normal or severe persistent right upper quadrant or epigastric pain, severe thrombocytopenia <100,00/μl, renal insufficiency with serum creatinine >1.1 mg/dl or a doubling of the serum creatinine, and pulmonary edema, of which 4 of the 6 features were present in our patient (1).

As we learned from our expert panelists, a serum uric acid level >5.5 mg/dl, urinary calcium level <195 mg/dl, and elevated liver enzymes can be helpful in differentiating features of preeclampsia versus SLE/LN and obstetric-APS disease activity and may help guide treatment decision-making. In a retrospective study of 400 patients, 331 women (82.7%) had preeclampsia with severe features, and a serum uric acid level >7 mg/dl was associated with renal involvement and preeclampsia with severe features (20). Further, in a study of 143 obstetric patients, 24-hour urine collection was performed, and the 28 patients who developed preeclampsia had a urine calcium level of <12 mg/dl (21). Nevertheless, differentiating between these entities remains challenging and an unmet need remains for improved noninvasive testing.

There has been great debate regarding the safety of renal biopsy during the partum period. Practice guidelines including the 2012 American College of Rheumatology guidelines for the management of LN and the 2019 European League Against Rheumatism and European Renal Association–European Dialysis and Transplant Association LN guideline do not specifically address the critical clinical question regarding renal biopsy during pregnancy (22,23). In a 2020 review article by Farrington et al, the study authors determined that patients who develop evidence of new or progressive renal disease in the partum period should be considered for renal biopsy if less invasive methods of diagnosis have been exhausted and/or a renal biopsy will change management (24).

Chen et al proposed the following criteria to perform renal biopsy in patients with SLE during the partum period: 1) proteinuria 500 mg/gram creatinine and abnormal serology results (positive anti-dsDNA, low complement, aPL antibodies); 2) confirmation of fetal viability on ultrasound before and after biopsy (including fetal heart rate) at 23 weeks of gestation or greater; 3) minimization of bleeding risk by monitoring coagulation factors and platelet count, holding aspirin and fish oil for at least 1 week prior to the procedure and avoiding biopsy if risks of stopping anticoagulation are too high (i.e., a history of arterial thrombosis); 4) avoiding biopsy beyond 28 weeks gestation; and 5) if biopsy is not possible, considering starting empiric therapy and performing a biopsy during the postpartum period (25). However, this study had a small sample size of 11 women with SLE who underwent renal biopsy during pregnancy, and thus larger studies and further validation are necessary to construct consensus guidelines regarding the safety and timing of renal biopsy in pregnancy.

Dilemma three: use of G-CSF in patients with SLE and severe neutropenia

In a 2006 case series by Vasilui et al, 2 patients with SLE were found to experience severe disease flare in association with the use of G-CSF therapy (26). The case series noted that both patients experienced rapid and irreversible decline in renal function following use of G-CSF, indicating a strong temporal relationship and raising concern regarding the safety of use and need for careful clinical judgement (26).

In a 1999 study of 4 patients with SLE and an ANC level below 1,000/μl refractory to glucocorticoid therapy and other forms of immunosuppression, G-CSF was administered for prevention of infectious complications and/or as supportive therapy during infection (27). The study authors found that G-CSF can be an effective and a generally well-tolerated treatment for neutropenia in patients with SLE with rare instances of SLE flare. If G-CSF is administered, it should be used at the lowest effective

dose and duration, with close monitoring, discontinuing use once the ANC level increases to 1,000/μl (27).

FINAL DIAGNOSIS

In summary, our patient illustrated the complexities of active SLE/LN and development of obstetric-APS during her first pregnancy as she faced disease flare, preeclampsia, and preterm birth at 27 weeks. Patients with prepregnancy hypocomplementemia and elevated dsDNA antibody serologic markers, as seen in our case, have only an approximate 28% chance of carrying a normal pregnancy, as we learned from our expert panelist Dr. Petri. Thus, preconception counseling and pregnancy planning remain imperative.

Further, testing for aPL antibody panel positivity should be conducted in all patients with SLE considering pregnancy, with specific attention drawn to lupus anticoagulant positivity, as it strongly predicts the risk of poor pregnancy outcomes (13,14). These results can guide clinicians to add prophylactic therapy in patients with a history of thrombotic events or a previous pregnancy complication to decrease the risk of obstetric-APS and preeclampsia.

Patients with evidence of active renal disease in the partum period who met prosed criteria according to Chen et al (25) should consider undergoing renal biopsy during the first and second trimester to further elucidate the correct diagnosis and guide treatment decision-making, as kidney biopsy has been demonstrated to be safe and efficacious.

Finally, G-CSF should be used with clinical caution in patients with known SLE, as it can precipitate flare and irreversible renal function decline. If G-CSF is used, the lowest effective dose and duration should be given, with close monitoring of laboratory values and discontinuing use once the ANC level reaches 1,000/μl (27).

ACKNOWLEDGMENTS

The authors would like to thank the patient for allowing us to publish her case. We would also like to thank the expert panelists: Dr. Alberta Hoi, Dr. Guillermo Pons-Estel, and Dr. Michelle Petri for their review of our case during the American College of Rheumatology Global Meet the Experts Webinar on Lupus Nephritis in March of 2022.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Donohue, Garg.
Acquisition of data. Donohue, Garg.
Analysis and interpretation of data. Donohue, Gomez, Singh, Garg.

REFERENCES

1. American College of Obstetricians and Gynecologists. Gestational hypertension and preeclampsia: ACOG practice bulletin, number 222. Obstet Gynecol 2020;135:e237–60.
2. Bouma W, Klinkenberg TJ, van der Horst IC, et al. Mitral valve surgery for mitral regurgitation caused by Libman-Sacks endocarditis: a report of four cases and a systematic review of the literature. J Cardiothorac Surg 2010;5:13.
3. Smyth A, Oliveira GH, Lahr BD, et al. A systematic review and meta-analysis of pregnancy outcomes in patients with systemic lupus erythematosus and lupus nephritis. Clin J Am Soc Nephrol 2010;5: 2060–8.
4. Garg S, Singh T, Panzer SE, et al. Multidisciplinary lupus nephritis clinic reduces time to renal biopsy and improves care quality. ACR Open Rheumatol 2022;4:581–6.
5. Clowse ME, Jamison M, Myers E, et al. A national study of the complications of lupus in pregnancy. Am J Obstet Gynecol 2008;199:127. e1–6.
6. Clowse ME, Magder LS, Witter F, et al. The impact of increased lupus activity on obstetric outcomes. Arthritis Rheum 2005;52:514–21.
7. Clark CA, Spitzer KA, Nadler JN, et al. Preterm deliveries in women with systemic lupus erythematosus. J Rheumatol 2003;30:2127–32.
8. Petri M. Epidemiology of the antiphospholipid antibody syndrome. J Autoimmun 2000;15:145–51.
9. Ünlü O, Zuily S, Erkan D. The clinical significance of antiphospholipid antibodies in systemic lupus erythematosus. Eur J Rheumatol 2016; 3:75–84.
10. Petri M. Antiphospholipid syndrome [review]. Transl Res 2020;225: 70–81.
11. Petri M. Pregnancy and systemic lupus erythematosus [review]. Best Pract Res Clin Obstet Gynaecol 2020;64:24–30.
12. Sammaritano LR, Bermas BL, Chakravarty EE, et al. 2020 American College of Rheumatology guideline for the management of reproductive health in rheumatic and musculoskeletal diseases. Arthritis Rheumatol 2020;72:529–56.
13. Lockshin MD, Kim M, Laskin CA, et al. Prediction of adverse pregnancy outcome by the presence of lupus anticoagulant, but not anticardiolipin antibody, in patients with antiphospholipid antibodies. Arthritis Rheum 2012;64:2311–8.
14. Yelnik CM, Laskin CA, Porter TF, et al. Lupus anticoagulant is the main predictor of adverse pregnancy outcomes in aPL-positive patients: validation of PROMISSE study results. Lupus Sci Med 2016;3:e000131.
15. Liu X, Qiu Y, Yu ED, et al. Comparison of therapeutic interventions for recurrent pregnancy loss in association with antiphospholipid syndrome: a systematic review and network meta-analysis. Am J Reprod Immunol 2020;83:e13219.
16. Duley L, Meher S, Hunter KE, et al. Antiplatelet agents for preventing pre-eclampsia and its complications. Cochrane Database Syst Rev 2019;2019:CD004659.
17. Germain S, Nelson-Piercy C. Lupus nephritis and renal disease in pregnancy. Lupus 2006;15:148–155.
18. Chakravarty EF, Colon I, Langen ES, et al. Factors that predict prematurity and preeclampsia in pregnancies that are complicated by systemic lupus erythematosus. Am J Obstet Gynecol 2005;192: 1897–904.
19. Baer AN, Witter FR, Petri M. Lupus and pregnancy [review]. Obstet Gynecol Surv 2011;66:639–53.
20. Sudjai D, Satho P. Relationship between maternal serum uric acid level and preeclampsia with or without severe features. J Obstet Gynaecol 2022;42:2704–8.

21. Sanchez-Ramos L, Sandroni S, Andres FJ, et al. Calcium excretion in preeclampsia. Obstet Gynecol 1991;77:510–3.

22. Hahn BH, McMahon MA, Wilkinson A, et al. American College of Rheumatology guidelines for screening, treatment, and management of lupus nephritis. Arthritis Care Res (Hoboken) 2012;64:797–808.

23. Fanouriakis A, Kostopoulou M, Cheema K, et al. 2019 Update of the joint European League Against Rheumatism and European Renal Association-European Dialysis and Transplant Association (EULAR/ERA-EDTA) recommendations for the management of lupus nephritis. Ann Rheum Dis 2020;79:713–23.

24. Farrington CA. Kidney imaging and biopsy in pregnancy [review]. Adv Chronic Kidney Dis 2020;27:525–30.

25. Chen TK, Gelber AC, Witter FR, et al. Renal biopsy in the management of lupus nephritis during pregnancy. Lupus 2015;24:147–54.

26. Vasiliu IM, Petri MA, Baer AN. Therapy with granulocyte colony-stimulating factor in systemic lupus erythematosus may be associated with severe flares. J Rheumatol 2006;33:1878–80.

27. Hellmich B, Schnabel A, Gross WL. Treatment of severe neutropenia due to Felty's syndrome or systemic lupus erythematosus with granulocyte colony-stimulating factor. Semin Arthritis Rheum 1999;29; 82–99.

Arthritis Care & Research
Vol. 75, No. 10, October 2023, pp 2044–2053
DOI 10.1002/acr.25105

American College of Rheumatology
Empowering Rheumatology Professionals

Socioeconomic, Disease-Related, and Personal Factors Associated With Participation in Remote Follow-Up in Rheumatoid Arthritis: A Cross-Sectional Study

Liv M. V. Schougaard,[1] Line R. Knudsen,[2] Birgith E. Grove,[1] Jesper M. Vestergaard,[1] Niels H. Hjollund,[3] Ellen-Margrethe Hauge,[2] and Annette de Thurah[2]

Objective. To identify socioeconomic, disease-related, and personal factors associated with participation in remote follow-up in patients with rheumatoid arthritis (RA).

Methods. Following the implementation of a patient-reported outcome-based remote follow-up intervention in RA patients in Denmark, a cross-sectional study was conducted among 775 prevalent patients. In 2019, an electronic questionnaire was sent to eligible RA patients, covering health literacy and patient experience regarding involvement and confidence with remote care. Questionnaire data were linked to nationwide registries regarding socioeconomic status, labor market affiliation, and comorbidity level. Associations between registry- and questionnaire-based factors and remote follow-up were analyzed using multiple logistic regression analysis.

Results. All 775 patients were included in the registry-based analyses, but only 394 of 646 (61%) completed the questionnaire. No attachment to the labor market or low household income was associated with lower odds of remote follow-up participation (odds ratio [OR] 0.53 [95% confidence interval (95% CI) 0.34–0.83]) and (OR 0.69 [95% CI 0.48–1.00]). Further, a high level of comorbidity was associated with lower odds of remote follow-up participation compared to a low/medium level of comorbidity (OR 0.53 [95% CI 0.34–0.81]). No association was found between health literacy and remote follow-up, but remote follow-up attendees reported more confidence in remote care (OR 1.33 [95% CI 1.21–1.47]).

Conclusion. Participation in remote follow-up was associated with attachement to the labor market, household income, degree of comorbidity, and confidence with remote care. Additional research is necessary to investigate whether a larger and more divergent group of RA patients should be considered for inclusion in remote follow-up programs.

INTRODUCTION

The prevalence of rheumatoid arthritis (RA) in developed countries has increased by 60% from 1990 to 2010. It is expected to increase further due to the aging population, earlier diagnosis of RA, and lower mortality rate among individuals with systemic diseases (1). At the same time, we are witnessing a lack of rheumatologists and other health care professionals in rheumatology (2). Thus, there is increasing pressure on the health care system to improve efficacy, reduce waits, avoid waste, and provide patient-centered care without compromising patient safety and equity (3). Alternative models of care are needed, such as the use of telehealth. According to recently published EULAR points to consider for remote care (4), one way to implement telehealth is to collect repetitive patient-reported outcome (PRO) measures at home, which is already used in a wide range of different chronic conditions, including RA (5).

If PRO measures are collected electronically followed by response and action from a health care provider, this approach allows for new opportunities, such as real-time monitoring of symptoms with timely clinical intervention if needed and flexible hospital visit scheduling (6). Several randomized controlled trials have evaluated the effect of PRO-based telehealth interventions in different chronic diseases and found equal quality of care in

Supported by the Danish Health Foundation (grant 21-B-0113).
[1]Liv M. V. Schougaard, PhD, Birgith E. Grove, MHSc, Jesper M. Vestergaard, Master of IT: Gødstrup Hospital, Herning, Denmark; [2]Line R. Knudsen, MCN, Ellen-Margrethe Hauge, MD, PhD, Annette de Thurah, PhD: Aarhus University Hospital and Aarhus University, Aarhus, Denmark; [3]Niels H. Hjollund, MD, PhD: Gødstrup Hospital, Herning, Denmark, and Aarhus University Hospital, Aarhus, Denmark.

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.25105&file=acr25105-sup-0001-Disclosureform.pdf.

Address correspondence via email to Liv M. V. Schougaard, PhD, at livschou@rm.dk.

Submitted for publication September 23, 2022; accepted in revised form February 9, 2023.

SIGNIFICANCE & INNOVATIONS

- Remote care among patients with rheumatoid arthritis (RA) is a novel approach in the health care system.
- Only a selected group of RA patients participates in remote care programs (e.g., patients with attachment to the labor market, high household income, and a lower level of comorbidity).
- Future research should focus on investigating whether a larger and more divergent group of RA patients should be considered for inclusion in remote care.

PRO-based telehealth interventions compared to conventional outpatient follow-up. These studies also found that telehealth interventions contribute to lower health care service use (7–9). Furthermore, frequent PRO measure monitoring between in-person visits has been shown to have positive effects on self-efficacy, health-related quality of life, symptom burden, and survival (10–12).

In Denmark, the telehealth RA (TeRA) study has investigated the effect of monitoring disease activity through a PRO-based telehealth platform (7). In this study, the PRO measures were managed using the generic configurable TelePRO system AmbuFlex, a system designed to provide a decision aid in evaluating whether a patient needs an outpatient visit (13,14). The TeRA study was a noninferiority randomized controlled study testing the hypothesis of if there is no difference in disease activity between either TelePRO doctor, TelePRO nurse, or conventional outpatient follow-up. This hypothesis was confirmed (7). Today, the intervention has been implemented into daily clinical practice.

As of August 2022, ~35% of RA patients at the Department of Rheumatology in Aarhus, Denmark, participate in the telehealth intervention, called remote PRO-based follow-up in this study. The referral is based on a clinician's individual clinical judgment and patient preferences. A Danish study among patients with epilepsy found that socioeconomically advantaged patients were more likely to be referred to remote PRO-based follow-up compared to vulnerable patients (15). Further, another Danish qualitative study among patients with epilepsy found that remote PRO-based follow-up gave rise to ambivalence in clinicians, meaning that from the clinician perspective, remote care based on PRO measures could both increase and decrease the quality of follow-up (16).

However, no study has explored factors associated with participation in remote PRO-based follow-up in RA. The majority of existing studies have focused on the views of the patient regarding the usability of telehealth systems and patient satisfaction. Still, a qualitative study from the TeRA project shows that although patients are generally satisfied with this mode of follow-up, 2 different typologies have been identified; the "keen" and "reluctant" patient (17). The 2 typologies represent opposite perspectives and preferences regarding the core value of and approach to remote follow-up. We do not know if this factor will also impact participation in remote follow-up. We are also unaware of the degree to which other variables, such as socioeconomic factors and disease-related factors, play a role.

However, since remote follow-up presumably will become increasingly used in the future, it is important to consider these points to ensure that patients receive individualized follow-up based on their preferences. Thus, in this study, we aimed to identify socioeconomic, disease-related, and personal factors associated with participation in remote PRO-based follow-up in RA. We hypothesized positive associations between a high socioeconomic status, a high level of health literacy, and participation in the remote follow-up program.

PATIENTS AND METHODS

Remote PRO-based follow-up. Since 2016, the Department of Rheumatology at Aarhus University Hospital has offered remote PRO-based follow-up or conventional outpatient follow-up to RA patients. In remote PRO-based follow-up, patients receive a disease-specific questionnaire regularly (e.g., every 6 months). The disease-specific questionnaire includes self-reported aspects of past and present RA disease activity measured using the RA Flare instrument (18,19) and 2 ad hoc items regarding medication adherence and side effects. Based on the answers from the patient, an algorithm scores the response using a traffic light color code, with green indicating no need for clinical attention, yellow indicating that the patient may need clinical attention, and red indicating that the patient needs clinical attention. The questionnaire responses are integrated into electronic health records (EHRs). Thus, the responses and color codes are provided to rheumatology nurses, who assess the data and other relevant patient data from EHRs (e.g., blood tests and radiography).

All patients receive a yearly rheumatologist outpatient consultation. As of August 2022, 298 patients attend remote PRO-based follow-up, corresponding to 35% of the outpatient population. The remaining patients receive conventional outpatient follow-up with in-clinic visits ~1 or 2 times a year with a rheumatologist. Referral of patients into PRO-based follow-up is based on shared decision-making between the rheumatologist and the patient. First, a rheumatologist assesses the clinical profile of the patient. The target population for PRO-based follow-up is RA patients with low disease activity or remission and a disease duration of ≥2 years to ensure patients are familiar with the symptoms of disease activity and medical treatment. Furthermore, patients should be able to speak and understand Danish. Second, if the patient is assessed for eligibility by the rheumatologist, the final decision regarding participation in PRO-based follow-up is based on patient preference. Referral to remote

follow-up is a shared responsibility among all rheumatologists at the outpatient clinic. Further, nurses are trained to conduct PRO-based follow-up.

Study design and population. We designed a cross-sectional study of RA patients at the Department of Rheumatology of Aarhus University Hospital in Denmark. In January 2019, an electronic questionnaire was sent out to all prevalent RA patients diagnosed according to International Statistical Classification of Diseases and Related Health Problems, Tenth Revision codes M05.3, M05.9, M05.8, and M06.9, ≥18 years old. The study population was identified using the Healthcare Business Intelligence registry in the Central Denmark Region (20). The questionnaire was sent via a Danish national secure, digital mailbox "e-Boks." Patients with no access to the digital mailbox were not able to respond to the questionnaire. Nonresponders received a reminder after 2 weeks. In addition, data on all eligible patients were collected from a wide range of national registries in Statistics Denmark (21). The outcome of interest was whether the patients participated in remote PRO-based follow-up in January 2019, and this information was identified through the generic configurable TelePRO system, AmbuFlex (13,14).

The study was conducted in accordance with the Declaration of Helsinki and with the Danish law on data protection by the Danish Data Protection Agency (reference no. 623448). Specific, voluntary, explicit, and informed consent to participate in the study was obtained in accordance with guidelines from the Danish Data Protection Agency. According to Danish law, approval by the regional committee on health research ethics in the Central Denmark Region was not required because no biomedical intervention was performed. The data were stored and analyzed on secure servers located at Statistics Denmark.

Socioeconomic, disease-related, and personal factors. *Questionnaire data.* The questionnaire covered a range of self-reported aspects, including health literacy and patient experience (involvement in health care and confidence with remote care).

Health literacy. Health literacy was measured using the Health Literacy Questionnaire (HLQ) (22). HLQ includes 9 different subscales measuring different aspects of health literacy skills. The questionnaire has demonstrated sufficient psychometric properties in the context of a Danish population (23). We included 3 HLQ subscales: subscale 4 (social support for health), subscale 6 (ability to actively engage with health care providers), and subscale 9 (understanding health information well enough to know what to do). Each subscale consists of 5 items used to calculate a mean score ranging from 1 to 4 (subscale 4) or 1 to 5 (subscales 6 and 9). Higher scores indicate better degrees of health literacy (22). If 1 or 2 items were missing for each subscale, we estimated the mean score based on the other items in the subscale.

Involvement. The patient's perception of involvement in health care was determined by 5 ad hoc items generated based on a regional project regarding indicators of patient involvement (24). The 5 items consist of the following statements: 1) "The health care professionals asked questions about my own experiences with my disease," 2) "I talked to the health care professionals about the questions and concerns that I had," 3) "The health care professionals invited me to ask questions and talk about my concerns," 4) "I was consulted when decisions about my plans were made," and 5) "I talked adequately to the health care professional about how I manage my condition." Each item has the following response categories: not at all, to a lesser extent, to some extent, to a great extent, or to a very high extent, which was scored from 1 to 5.

Confidence with remote care. Confidence with remote care was measured using an ad hoc item developed based on data from individual semistructured interviews regarding the perspective of the "keen" and the "reluctant" patient (17). We included the item: "remote telephone conversation is just as safe as regular in-person consultations in the outpatient clinic if my RA disease is stable." The statement included a degree of agreement stated on a Likert scale ranging from 0 (completely agree) to 10 (completely disagree). The items regarding the perception of involvement and confidence with remote care were tested among a representative number of RA patients (n = 10) using the cognitive interview techniques "think aloud" and cognitive debriefing (25). The patients found the items relevant and understandable.

Registry data. We had access to various Danish registries via Statistics Denmark (21). The following registries were used: the Danish Civil Registration System (CRS) (26), the Danish Education Register (27), the Danish Register on Personal Income and Transfer Payments (28), the Danish Register for Evaluation and Marginalization (DREAM) (29), and the Danish National Patient Registry (DNPR) (30). We combined registry data at the individual level using a unique personal identification number (Central Person Register [CPR] number) given to all Danish citizens at birth and residents upon immigration (26,31). The questionnaire data were linked with registry data at Statistics Denmark.

The CRS registry contains demographic information regarding all Danish citizens and is updated daily (26). We used the registry to retrieve information about cohabitation status. The cohabitation status variable was dichotomized into "not living alone" and "solo living." The Danish Education Register contains information regarding the highest completed education of Danish adults (27); we retrieved this information regarding study participants the year before their inclusion. Education level was dichotomized into high/medium (≥10 years) and low (<10 years) education level according to the International Standard Classification of Education (32). The Danish Register on Personal Income and Transfer Payments provides an overview of the income composition of the Danish population (28). We retrieved information

regarding yearly equivalized disposable household income at the residence of participants the year before their inclusion in the study. Equivalized household income takes into account differences in the composition of the household. Household income was divided into high/medium (66.6 percentile) and low (33.3 percentile) income levels.

The DREAM registry contains information regarding attachment of an individual to the labor market and their temporary and permanent public benefits, and data are updated weekly (29). We retrieved information regarding attachment to the labor market in individuals within 52 weeks before the date of inclusion in the study. Attachment to the labor market was dichotomized into self-supporting and retirement or temporary/permanent public benefits. The DNPR provides information regarding all diagnoses registered with hospital admissions, emergency room visits, and outpatient visits in private and public hospitals in Denmark (30).

We retrieved information regarding comorbidity of participants 10 years before the study enrollment. Comorbidity was measured using the Charlson comorbidity index (33). The participants were dichotomized by score: 0–2 (low/medium level of comorbidity) and >2 (high level of comorbidity).

Outcome variable. The outcome was whether the enrolled patients participated in remote PRO-based follow-up at the Department of Rheumatology at Aarhus University Hospital in January 2019. When patients are referred to remote PRO-based follow-up, this is registered by a health care provider in the AmbuFlex system. We used the date of referral registration to define whether a patient participated in the remote follow-up program. Data were retrieved from the AmbuFlex database (5,13).

Statistical analysis. Categorical data were presented as frequencies and percentages. For normally distributed continuous data, the mean ± SD were presented, and for non-normally distributed data, the median (interquartile range [IQR]) were also presented. Differences in baseline characteristics between questionnaire responders and nonresponders and patients without a digital mailbox were assessed using Person's chi-square test. A 2-sided *P* value less than 0.05 was considered statistically significant. Associations between questionnaire and registry-based data and remote PRO-based follow-up were estimated using multiple logistic regression analysis. Crude and adjusted odd ratios (ORs) were reported with 95% confidence intervals (95% CIs). We included the following confounder variables: age, sex, cohabitation status, education, and comorbidity. The confounder variables were selected a priori based on previous studies regarding referral, adherence, or nonparticipation in telehealth interventions (15,34,35). A sensitivity analysis was conducted by excluding patients with <1 year of remote follow-up experience. We performed analyses using Stata version 16.

RESULTS

Characteristics of the study population. A total of 775 adult patients were registered as having an RA diagnosis at the Department of Rheumatology at Aarhus University Hospital in January 2019 and were eligible for inclusion in the study. We sent an electronic questionnaire to all eligible study participants; however, 129 patients did not have a secure digital mailbox and could not receive the questionnaire. A total of 394 patients (61%) completed the questionnaire. A flow chart of patients included in the study is shown in Figure 1.

The median age of the total study population (n = 775) was 65.2 years (IQR 20.2 years), 71% were women, and 65.8% did not live alone (Table 1). Moreover, 72.4% had no attachment to the labor market, and 20.1% had a high level of comorbidity. Patients without access to a digital mailbox were older, were more likely to live alone, were lower educated, had lower household income, were less likely to be attached to the labor market, and had higher levels of comorbidity ($P < 0.001$ in all variables). Questionnaire nonresponders were more likely to be attached to the labor market than responders ($P = 0.01$), but no differences were found in the other registry variables. Among the 775 patients, 247 were referred to remote PRO-based follow-up, and 528 received a conventional outpatient follow-up with a rheumatologist. The distribution of patients in remote follow-up (n = 247) stratified based on patients without a digital mailbox, nonresponders, and responders were 27, 62, and 158 patients, respectively. The questionnaire response rate was lower among conventional follow-up participants (55.4%) compared to remote follow-up participants (71.8%) ($P < 0.001$).

An overview of self-reported data among the 394 questionnaire responders is shown in Table 2. Overall, patients reported high levels of health literacy. Scores for the HLQ 6 domain and the HLQ 9 domain were mean ± SD of 4.1 ± 0.74 and 4.1 ± 0.63, respectively. The stratified data according to remote or conventional follow-up showed slightly higher levels among remote follow-up participants compared to conventional follow-up participants in the following self-reported aspects: ability to actively engage with health care providers (HLQ 6 domain), talking to the health care professionals about questions and concerns, talking adequately to the health care professional regarding how to manage the condition, and confidence with using remote care. Missing data were <5% for all subscales and items.

Associations between registry-based data and remote PRO-based follow-up. Associations between registry-based data and participating in remote PRO-based follow-up are shown in Table 3. We found that patients with no attachment to the labor market (retired or receiving public benefits) had an adjusted OR of participating in the remote follow-up

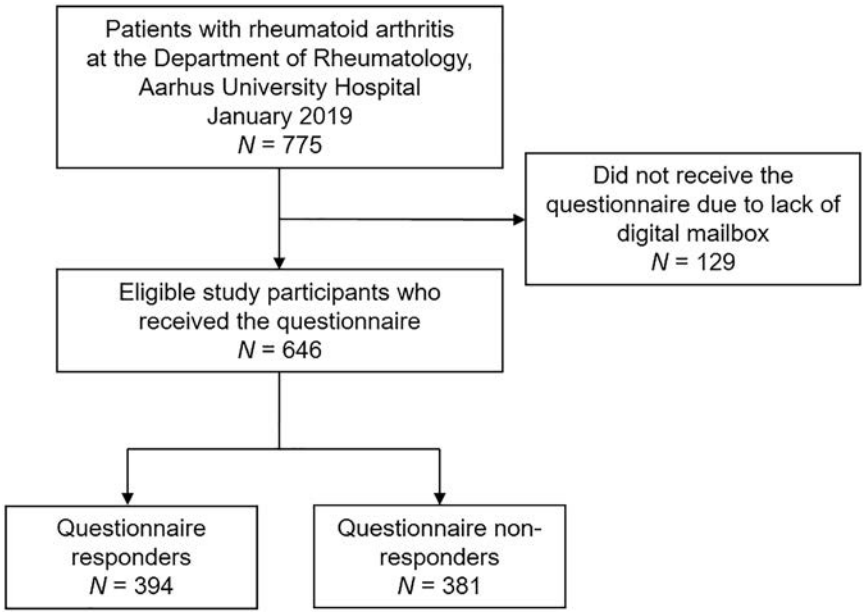

Figure 1. Flow chart of study participants.

of 0.53 (95% CI 0.34 – 0.83) compared to patients with attachment to the labor market (self-supporting patients). A similar result was found in patients with a low household income, as they had an adjusted OR of participating in the remote follow-up of 0.69 (95% CI 0.48 – 1.00) compared to patients with a high/medium household income. Moreover, we found that patients with a high level of comorbidity had an adjusted OR of participating in the remote follow-up of 0.53 (95% CI 0.34 – 0.81) compared to patients with a low/medium level of comorbidity. No statistically significant associations were found in the other socioeconomic factors. A graphical overview of the results is shown in Supplementary Figure 1 (available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25105/abstract).

Table 1. Registry characteristics in 775 patients with rheumatoid arthritis in January 2019, stratified according to whether they were questionnaire responders, nonresponders, and patients with no digital mailbox*

	Total (n = 775)	Respondents (n = 394)	Nonrespondents (n = 252)	No digital mailbox (n = 129)
Age, median (IQR) years	65.2 (53.6–73.8)	63.8 (55.7–71.7)	57.9 (46.0–70.9)	76.8 (70.8–82.3)
Sex				
Female	550 (71.0)	270 (68.5)	182 (72.2)	98 (76.0)
Male	225 (29.0)	124 (31.5)	70 (27.8)	31 (24.0)
Cohabitation status				
Not living alone	510 (65.8)	283 (71.8)	182 (72.2)	45 (34.9)
Solo living	265 (34.2)	111 (28.2)	70 (27.8)	84 (65.1)
Education				
High/medium (≥10 years)	549 (70.8)	295 (74.9)	188 (74.6)	66 (51.2)
Low (<10 years)	205 (26.5)	89 (22.6)	58 (23.0)	58 (45.0)
Missing	21 (2.7)	10 (2.5)	6 (2.4)	5 (3.9)
Household income				
High/medium	519 (67.0)	295 (74.9)	177 (70.2)	47 (36.4)
Low	256 (33.0)	99 (25.1)	75 (29.8)	82 (63.6)
Labor market affiliation				
Self-supporting	214 (27.6)	115 (29.2)	98 (38.9)	NA
Retirement/public benefits	561 (72.4)	279 (70.8)	154 (61.1)	NA
Charlson comorbidity index				
Low/medium (0 − 2)	619 (79.9)	318 (80.7)	213 (84.5)	88 (68.2)
High (>2)	156 (20.1)	76 (19.3)	39 (15.5)	41 (31.8)

* Except where indicated otherwise, values are the number (%) of patients. IQR = interquartile range; NA = not applicable.

Table 2. Self-reported characteristics stratified according to the mode of outpatient follow-up among 394 responding patients with rheumatoid arthritis in January 2019*

	Total (n = 394)	Remote PRO-based follow-up (n = 158)	Conventional follow-up (n = 236)
Social support for health (HLQ 4 domain)			
Mean ± SD	3.2 ± 0.59	3.2 ± 0.58	3.1 ± 0.60
Median (IQR)	3.2 (3.0–3.6)	3.2 (3.0–3.8)	3 (2.8–3.6)
Missing, no. (%)	11 (2.8)	–	–
Ability to actively engage with health care providers (HLQ 6 domain)			
Mean ± SD	4.1 ± 0.74	4.2 ± 0.66	4.0 ± 0.78
Median (IQR)	4 (3.7–4.8)	4 (3.8–4.8)	4 (3.6–4.8)
Missing, no. (%)	8 (2.0)	–	–
Understanding health information well enough to know what to do (HLQ 9 domain)			
Mean ± SD	4.1 ± 0.63	4.1 ± 0.61	4.2 ± 0.63
Median (IQR)	4 (3.8–4.5)	4 (3.8–4.5)	4 (3.8–4.8)
Missing, no. (%)	8 (2.0)	–	–
The health care professionals asked questions about my own experiences with my disease			
Mean ± SD	3.5 ± 1.0	3.6 ± 0.99	3.4 ± 0.99
Missing, no. (%)	12 (3.0)	–	–
I talked to the health care professionals about the questions and concerns that I had			
Mean ± SD	3.7 ± 0.91	3.8 ± 0.82	3.6 ± 0.97
Missing, no. (%)	11 (2.8)	–	–
The health care professionals invited me to ask questions and talk about my concerns			
Mean ± SD	3.1 ± 1.15	3.2 ± 1.12	3.0 ± 1.16
Missing, no. (%)	11 (2.8)	–	–
I was consulted when decisions about my plans were made			
Mean ± SD	3.7 ± 0.99	3.7 ± 0.96	3.7 ± 1.01
Missing, no. (%)	12 (3.0)	–	–
I talked adequately to the health care professional about how I manage my condition			
Mean ± SD	3.6 ± 1.04	3.7 ± 0.96	3.5 ± 1.08
Missing, no. (%)	12 (3.0)	–	–
Confidence in remote care			
Mean ± SD	7.9 ± 2.95	9.0 ± 1.92	7.1 ± 3.26
Missing, no. (%)	9 (2.3)	–	–
RA disease duration, no. (%) of patients			
≥10 years	243 (61.7)	106 (67.1)	137 (58.1)
<10 years	134 (34.0)	46 (29.1)	88 (37.3)
Missing	17 (4.3)	6 (3.8)	11 (4.7)

* HLQ = Health Literacy Questionnaire; IQR = interquartile range; PRO = patient-reported outcome.

Associations between questionnaire-based data and remote PRO-based follow-up. Associations between questionnaire-based data and participating in remote PRO-based follow-up are shown in Table 4. We found that a 1-unit increase in the mean scale score of the items "I talked adequately to the health care professional about how I manage my condition" and "The health care professionals invited me to ask questions and talk about my concerns" had an adjusted ORs of participating in the remote follow-up of 1.28 (95% CI 1.03 – 1.57) and 1.21 (95% CI 1.00 – 1.46), respectively. Similarly, we found that a 1-unit increase in the mean scale score of the item "Remote telephone conversation is just as safe as regular in-person consultations in the outpatient clinic if my RA disease is stable" had an adjusted OR of participating in the remote follow-up of 1.33 (95% CI 1.21 – 1.47). No statistically significant associations were found between the levels of health literacy and remote follow-up. A graphical overview of the results is shown in Supplementary Figure 2 (available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25105/abstract). The sensitivity analysis excluding patients with <1 year of remote follow-up experience (n = 45) did not change the results (data not shown).

DISCUSSION

In this study, we found that RA patients with no attachment to the labor market or RA patients who had a lower level of

Table 3. Odds ratio (OR) of participating in remote PRO-based follow-up in 775 patients with rheumatoid arthritis according to registry-based factors*

	Crude OR (95% CI)	Adjusted OR (95% CI)†
Age, years		
18–54	Ref.	Ref.
55–64	1.42 (0.93–2.17)	1.40 (0.91–2.17)
65–74	1.24 (0.83–1.86)	1.35 (0.88–2.05)
75–99	0.70 (0.44–1.13)	0.82 (0.50–1.35)
Sex		
Female	Ref.	Ref.
Male	1.23 (0.89–1.71)	1.29 (0.92–1.82)
Cohabitation status		
Not living alone	Ref.	Ref.
Solo living	0.94 (0.68–1.29)	0.98 (0.70–1.38)
Education		
High/medium (≥10 years)	Ref.	Ref.
Low (<10 years)	0.83 (0.59–1.18)	0.86 (0.60–1.23)
Household income		
High/medium	Ref.	Ref.
Low	0.67 (0.48–0.93)	0.69 (0.48–1.00)
Labor market affiliation		
Yes (self-supporting)	Ref.	Ref.
No (retirement/public benefits)	0.62 (0.44–0.86)	0.53 (0.34–0.83)
Charlson comorbidity index		
Low/medium 0–2	Ref.	Ref.
High >2	0.56 (0.37–0.84)	0.53 (0.34–0.81)

* 95% CI = 95% confidence intervals; PRO = patient-reported outcome; ref. = reference.
† Adjusted for age, sex, cohabitation status, education, and comorbidity.

Table 4. Odds ratio (OR) of participating in remote PRO-based follow-up in 394 patients with rheumatoid arthritis according to questionnaire-based factors*

	Crude OR (95% CI)	Adjusted OR (95% CI)†
Health literacy		
Social support for health (HLQ 4 domain score)	1.30 (0.91–1.84)	1.27 (0.88–1.84)
Ability to actively engage with health care providers (HLQ 6 domain score)	1.39 (1.04–1.85)	1.34 (0.99–1.82)
Understanding health information well enough to know what to do (HLQ 9 domain score)	0.86 (0.62–1.19)	0.83 (0.59–1.17)
Patient involvement		
The health care professionals asked questions about my own experiences with my disease	1.25 (1.01–1.55)	1.22 (0.98–1.51)
I talked to the health care professionals about the questions and concerns that I had	1.25 (0.99–1.58)	1.24 (0.98–1.58)
The health care professionals invited me to ask questions and talk about my concerns	1.22 (1.02–1.47)	1.21 (1.00–1.46)
I was consulted when decisions about my plans were made	1.09 (0.89–1.35)	1.09 (0.88–1.35)
I talked adequately to the health care professional about how I manage my condition	1.27 (1.04–1.56)	1.28 (1.03–1.57)
Confidence in remote care	1.33 (1.20–1.46)	1.33 (1.21–1.47)

* 95% CI = 95% confidence interval; HLQ = Health Literacy Questionnaire; PRO = patient-reported outcome.
† Adjusted for age, sex, cohabitation status, education, and comorbidity.

household income were less likely to participate in remote PRO-based follow-up. In contrast, other socioeconomic factors did not play a role. Further, patients with a high level of comorbidity were less likely to participate in remote follow-up compared to patients with a low/medium level of comorbidity. No associations between high levels of health literacy and participation in remote follow-up were seen, which did not support our hypothesis. However, some self-reported aspects, such as high confidence with remote care and talking adequately to the health care professional regarding how to manage the condition, were positively associated with participation in remote follow-up.

Only a few studies have used PROs in telehealth to monitor patients remotely in rheumatology (7,36–38). No studies have investigated the characteristics of patients attending remote PRO-based follow-up. Participation in remote follow-up was most pronounced among RA patients affiliated with the labor market, which was consistent with 3 cross-sectional studies identifying flexibility and saved travel time as facilitators of remote care (39–41). The same was found in the qualitative study based on the TeRA study (17). A Danish study among patients with epilepsy has also investigated factors associated with referral to remote PRO-based follow-up (15). Concurrent with our findings, this study also found that patients without attachment to the labor market and low-income levels were less likely to be referred to remote PRO-based follow-up.

Unlike this study, the epilepsy study found that additional socioeconomic factors played a role in referral to remote follow-up (e.g., living alone, low education, or low level of health literacy) (15). The discrepancy in the present findings is most likely caused by the differences in study designs and populations. In the RA study, we saw a ceiling effect on the level of health literacy, similar to the Danish background population (23). The epilepsy study was conducted among incident patients. In contrast, our cross-sectional study was conducted among prevalent patients with a

long disease duration, which may explain why patients in the epilepsy study reported much lower levels of health literacy. Presumably, long experience in ongoing interaction with the health care system may positively impact health literacy levels in patients.

In a recently published systematic review, factors associated with adherence to telemonitoring were investigated using electronic PRO measures in patients with chronic diseases (e.g., RA, heart failure, and chronic pain) (42). In general, the strength of evidence regarding associations between socioeconomic, condition-related, and patient-related factors and adherence was assessed to be inconclusive and inconsistent in all studies. However, the association between sex and adherence was determined to have a moderate level of evidence. It was found that sex was not associated with adherence to PRO-based telemonitoring, including in RA patients (42). This result aligns with findings from the present study, as we did not find an association between sex and participation in remote PRO-based follow-up. Digital nonuse or adherence and referral to telehealth interventions have similarities, but they also refer to different concepts. Mostly because referral to telehealth interventions is decided at the clinical encounter based on a clinical judgment and the preferences of the patient; whereas, for instance, nonuse or adherence to a greater extent is only influenced by patient preferences; thus, a wide range of modifiable and unmodifiable factors could have different relevance.

In the Danish TelePRO system AmbuFlex, several PRO-based telehealth solutions have been implemented in clinical practice (5,13). Only a proportion of the total patient population is included in remote PRO-based follow-up, and the proportion varies between patient groups (e.g., 77% of the patient population was included in remote follow-up in inflammatory bowel disease [43], 50% in epilepsy [15], and 35% in RA [7]). This finding indicates that it is important to continue investigating the characteristics of patients attending remote PRO-based follow-up and focus on whether the use of PRO measures in routine care can be expanded to be used in other ways. Future research should focus on how health care services can be supportive of vulnerable patients to a larger extent. This may include differential use of PRO measures, such as using PRO measures prior to telephone consultations or in-clinic visits or as a proxy solution if patients are unable to complete a questionnaire by themselves. In addition, a future research focus should be investigating whether specific factors individually play a role in the referral procedure. For example, the clinician may be more reluctant to refer patients with a high level of comorbidity.

A key strength of this study is that it investigates a well-established PRO-based telehealth intervention that has been implemented in a real-life clinical setting since 2016. Thus, patients and health care providers have used the intervention for years, and patients are familiar with using PRO measures in remote care. All eligible participants were included in the registry-based analyses, and data were retrieved from several national Danish registries with high-quality data and few missing data. This fact has reduced the risk of bias in the registry-based results.

However, this study has some limitations. Firstly, we used a cross-sectional design, a snapshot with no causal patterns that can be investigated. Patients were already enrolled in remote follow-up when they responded to the questionnaire. We cannot rule out information bias if the patient responses were affected by the mode of outpatient follow-up. For example, patients already attending remote follow-up may have reported higher levels of confidence with remote care if they had positive experiences, which may have overestimated this result. Second, the questionnaire was not sent to patients without access to a digital mailbox. This group of patients was older, lower educated, had a lower income level, and had a higher level of comorbidity than the other patients included (see Table 1). This decision has reduced the questionnaire response rate and resulted in a risk of selection bias. The response rate was lower among patients in conventional follow-up compared to remote follow-up patients. Thus, nonresponse was related to socioeconomic factors and the outcome of interest. Research has shown that low health literacy is associated with low socioeconomic factors, such as low education, income, and unemployment (44). We expect the proportion of patients with lower health literacy levels to be higher among nonresponders (including patients without a digital mailbox) compared to responders, which may have made us underestimate the impact of health literacy. Considering that, we should have sent a paper version of the questionnaire to patients without a digital mailbox. A third limitation is that the study only included patients from a single university hospital unit in 1 Danish region, which may have limited the external validity of the results. Finally, there may be other factors not assessed in this study that could influence participation in remote PRO-based follow-up (e.g., ethnicity/language skills and digital literacy).

In conclusion, patients participating in remote follow-up were more likely to have attachment to the the labor market, had higher income levels, had a lower level of comorbidity, and expressed higher confidence in remote care. A future focus should be on how to support RA patients receiving conventional follow-up to achieve confidence in remote follow-up, as well as increased focus on vulnerable patient groups and the need for differential use of health care services.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Schougaard, Hjollund, Hauge, de Thurah.

Acquisition of data. Schougaard, Knudsen, Hjollund, de Thurah.

Analysis and interpretation of data. Schougaard, Knudsen, Grove, Vestergaard, Hjollund, Hauge, de Thurah.

REFERENCES

1. Safiri S, Kolahi AA, Hoy D, et al. Global, regional and national burden of rheumatoid arthritis 1990–2017: a systematic analysis of the Global Burden of Disease study 2017. Ann Rheum Dis 2019;78:1463–71.
2. Unger J, Putrik P, Buttgereit F, et al. Workforce requirements in rheumatology: a systematic literature review informing the development of a workforce prediction risk of bias tool and the EULAR points to consider. RMD Open 2018;4:e000756.
3. Institute of Medicine Committee on Quality of Health Care in America. Crossing the quality chasm: a new health system for the 21st century [review]. Washington (DC): National Academies Press (US); 2001.
4. De Thurah A, Bosch P, Marques A, et al. 2022 EULAR points to consider for remote care in rheumatic and musculoskeletal diseases. Ann Rheum Dis 2022;81:1065–71.
5. Hjollund NH, Larsen LP, Biering K, et al. Use of Patient-Reported Outcome (PRO) measures at group and patient levels: experiences from the generic integrated PRO system, WestChronic. Interact J Med Res 2014;3:e5.
6. Calvert M, Kyte D, Price G, et al. Maximising the impact of patient reported outcome assessment for patients and society. BMJ 2019; 364:k5267.
7. De Thurah A, Stengaard-Pedersen K, Axelsen M, et al. Tele-health followup strategy for tight control of disease activity in rheumatoid arthritis: results of a randomized controlled trial. Arthritis Care Res (Hoboken) 2018;70:353–60.
8. De Jong MJ, van der Meulen-de Jong AE, Romberg-Camps MJ, et al. Telemedicine for management of inflammatory bowel disease (myIBDcoach): a pragmatic, multicentre, randomised controlled trial. Lancet 2017;390:959–68.
9. Laurberg T, Schougaard LM, Hjollund NH, et al. Randomized controlled study to evaluate the impact of flexible patient-controlled visits in people with type 1 diabetes: the DiabetesFlex Trial. Diabet Med 2022;39:e14791.
10. Absolom K, Warrington L, Hudson E, et al. Phase III randomized controlled trial of eRAPID: eHealth intervention during chemotherapy. J Clin Oncol 2021;39:734–47.
11. Basch E, Deal AM, Kris MG, et al. Symptom monitoring with patient-reported outcomes during routine cancer treatment: a randomized controlled trial. J Clin Oncol 2016;34:557–65.
12. Denis F, Lethrosne C, Pourel N, et al. Randomized trial comparing a web-mediated follow-up with routine surveillance in lung cancer patients. J National Cancer Inst 2017;109.
13. Schougaard LM, Larsen LP, Jessen A, et al. AmbuFlex: tele-patient-reported outcomes (telePRO) as the basis for follow-up in chronic and malignant diseases. Qual Life Res 2016;25:525–34.
14. Hjollund NH. Fifteen years' use of patient-reported outcome measures at the group and patient levels: trend analysis. J Medical Internet Res 2019;21:e15856.
15. Schougaard LM, de Thurah A, Christensen J, et al. Sociodemographic, personal, and disease-related determinants of referral to patient-reported outcome-based follow-up of remote outpatients: a prospective cohort study. Qual Life Res 2020;29:1335–47.
16. Mejdahl CT, Schougaard LM, Hjollund NH, et al. PRO-based follow-up as a means of self-management support: an interpretive description of the patient perspective. J Patient Rep Outcomes 2018;2:38.
17. Knudsen LR, de Thurah A, Lomborg K. Experiences with telehealth followup in patients with rheumatoid arthritis: a qualitative interview study. Arthritis Care Res (Hoboken) 2018;70:1366–72.
18. De Thurah A, Maribo T, Stengaard-Pedersen K. Patient self-assessment of flare in rheumatoid arthritis: criterion and concurrent validity of the Flare instrument. Clin Rheumatol 2016;35:467–71.
19. Maribo T, de Thurah A, Stengaard-Pedersen K. Patient-self assessment of flare in rheumatoid arthritis: translation and reliability of the Flare instrument. Clin Rheumatol 2016;35:1053–8.
20. Business Intelligence in Central Denmark Region (In Danish). URL: https://www.rm.dk/om-os/organisation/it/business-intelligence/.
21. Danmarks Statistik. Brug af forskermaskiner. URL: https://www.dst.dk/da/TilSalg/Forskningsservice/brug-af-forskermaskiner.
22. Osborne RH, Batterham RW, Elsworth GR, et al. The grounded psychometric development and initial validation of the Health Literacy Questionnaire (HLQ). BMC Public Health 2013;13:658.
23. Maindal HT, Kayser L, Norgaard O, et al. Cultural adaptation and validation of the Health Literacy Questionnaire (HLQ): robust nine-dimension Danish language confirmatory factor model. Springerplus 2016;5:1232.
24. DEFACTUM®. Udvælgelse af inddragelsesspørgsmål til patienter. 2016.URL: https://www.defactum.dk/siteassets/defactum/3-projektsite/indikatorer-for-patientinddragelse/delrapport-3--regionsvalidering_vigtighed_relevans.pdf.
25. Beatty PC, Willis GB. Research synthesis: the practice of cognitive interviewing. Public Opin Q 2007;71:287–311.
26. Pedersen CB. The Danish Civil Registration System. Scand J Public Health 2011;39 Suppl:22–5.
27. Jensen VM, Rasmussen AW. Danish Education Registers. Scand J Public Health 2011;39 Suppl:91–4.
28. Baadsgaard M, Quitzau J. Danish registers on personal income and transfer payments. Scand J Public Health 2011;39 Suppl:103–5.
29. Hjollund NH, Larsen FB, Andersen JH. Register-based follow-up of social benefits and other transfer payments: accuracy and degree of completeness in a Danish interdepartmental administrative database compared with a population-based survey. Scand J Public Health 2007;35:497–502.
30. Lynge E, Sandegaard JL, Rebolj M. The Danish National Patient Register. Scand J Public Health 2011;39 Suppl:30–3.
31. Frank L. Epidemiology. When an entire country is a cohort. Science 2000;287:2398–9.
32. United Nations Educational, Scientific and Cultural Organization. International Standard Classification of Education. 1998. URL: http://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-1997-en_0.pdf.
33. Charlson ME, Pompei P, Ales KL, et al. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40:373–83.
34. Meyerowitz-Katz G, Ravi S, Arnolda L, et al. Rates of attrition and dropout in app-based interventions for chronic disease: systematic review and meta-analysis. J Med Internet Res 2020;22:e20283.
35. Nielsen AS, Kidholm K, Kayser L. Patients' reasons for non-use of digital patient-reported outcome concepts: a scoping review. Health Informatics J 2020;26:2811–33.
36. El Miedany Y, El Gaafary M, Youssef S, et al. Toward electronic health recording: evaluation of electronic patient-reported outcome measures system for remote monitoring of early rheumatoid arthritis. J Rheumatol 2016;43:2106–12.
37. Gossec L, Cantagrel A, Soubrier M, et al. An e-health interactive self-assessment website (Sanoia®) in rheumatoid arthritis. A 12-month randomized controlled trial in 320 patients. Joint Bone Spine 2018; 85:709–14.
38. Salaffi F, Ciapetti A, Gasparini S, et al. Web/Internet-based telemonitoring of a randomized controlled trial evaluating the time-integrated effects of a 24-week multicomponent intervention on key health

outcomes in patients with fibromyalgia. Clin Exp Rheumatol. 2015;33 Suppl 88:S93–101.

39. Lawford BJ, Bennell KL, Hinman RS. Consumer perceptions of and willingness to use remotely delivered service models for exercise management of knee and hip osteoarthritis: a cross-sectional survey. Arthritis Care Res (Hoboken) 2017;69:667–76.
40. Ferwerda M, van Beugen S, van Burik A, et al. What patients think about E-health: patients' perspective on internet-based cognitive behavioral treatment for patients with rheumatoid arthritis and psoriasis. Clin Rheumatol 2013;32:869–73.
41. Bullock DR, Vehe RK, Zhang L, et al. Telemedicine and other care models in pediatric rheumatology: an exploratory study of parents' perceptions of barriers to care and care preferences. Pediatr Rheumatol Online J 2017;15:55.
42. Wiegel J, Seppen B, van der Leeden M, et al. Adherence to telemonitoring by electronic patient-reported outcome measures in patients with chronic diseases: a systematic review. Int J Environ Res Public Health 2021;18:10161.
43. Appel CW, Pedersen SC, Nielsen AS, et al. Telemedicine based on patient-reported outcomes in management of patients with inflammatory bowel disease in a real-life setting - a before and after cohort study. Scand J Gastroenterol 2022;57:825–31.
44. Stormacq C, Van den Broucke S, Wosinski J. Does health literacy mediate the relationship between socioeconomic status and health disparities? Integrative review. Health Promot Int 2019;34:e1–17.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 75, No. 10, October 2023, pp 2054–2062
DOI 10.1002/acr.25106

Clinical Visit Frequencies in Rheumatology: A Systematic Literature Review

Yuxuan Jiang,[1] Robert S. Rudin,[2] and Daniel H. Solomon[3]

Objective. Clinical visits are a fundamental aspect of rheumatic disease care, but recommendations for appropriate visit frequencies are largely absent from guidelines, scarcely studied, and inconsistently reported. The aim of this systematic review was to summarize the evidence pertaining to visit frequencies for major rheumatic diseases.

Methods. This systematic review was conducted according to Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. Title/abstract screening, full-text screening, and extraction were carried out by 2 independent authors. Annual visit frequencies were either extracted or calculated and stratified by disease type and country of study. Weighted mean annual visit frequencies were calculated.

Results. A total of 273 relevant manuscript records were screened, and 28 were included after applying selection criteria. The included studies were equally divided between US and non-US and were published between 1985 and 2021. Most (n = 16) focused on rheumatoid arthritis (RA), systemic lupus erythematosus (SLE; n = 5), and fibromyalgia (FM; n = 4). For RA, the average annual visit frequencies were 5.25 for US rheumatologists, 4.80 for US non-rheumatologists, 3.29 for non-US rheumatologists, and 2.74 for non-US non-rheumatologists. For SLE, annual visit frequencies for non-rheumatologists were much higher than for US rheumatologists (12.3 versus 3.24). For FM, annual visit frequencies were 1.80 for US rheumatologists and 0.40 for non-US rheumatologists. There was a decreasing trend of visit frequency to rheumatologists from 1982 to 2019.

Conclusion. Evidence for rheumatology clinical visits was limited and heterogeneous on a global scale. However, general trends suggest more frequent visits in the US and less frequent visits in recent years.

INTRODUCTION

Rheumatology is mainly an outpatient specialty, with patients' clinical visits often comprising a major aspect of care (1). Timely initiation and sustained maintenance of medical interventions, which typically occur at clinical visits, are associated with improved prognosis and symptomatic amelioration in many rheumatic diseases (2). However, the appropriate frequency to visit a physician for rheumatic disease follow-up has not been a major focus of study (3), resulting in a lack of guidance for rheumatic disease follow-up for common conditions like rheumatoid arthritis (RA) (4) and osteoarthritis (OA) (5).

Competing and sometimes contradictory evidence abounds in rheumatology health service literature with regard to optimal visit frequencies. A US-based study from the late 1990s reported that more frequent visits to rheumatologists were associated with greater improvements in pain and functional disability in RA (6), while a 2021 Danish randomized control trial found no difference in patient prognosis between usual care and an experimental arm that featured fewer visits (7). An analysis of a US national patient database revealed significant regional variation regarding rheumatology visits and referral patterns in the US; more frequent referrals and visits to a rheumatologist were not associated with a patient's overall satisfaction or perceived health status (8). After the advent of the COVID-19 pandemic, questions have been raised for both the traditional interval of patient care and the proper role of virtual care moving forward (9). As the global burden of rheumatic diseases continues to rise, substantial gaps in the rheumatic workforce, medical best practices, and accessibility to rheumatic care highlight the need for robust and clinically applicable evidence regarding visit frequencies to optimize

Supported by the NIH (grant AR-P30-072577) and the Rheumatology Research Foundation.

[1]Yuxuan Jiang, BSc: McMaster University, Hamilton, Ontario, Canada; [2]Robert S. Rudin, PhD: RAND Corporation, Boston, Massachusetts; [3]Daniel H. Solomon, MD, MPH: Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts.

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.25106&file=acr25106-sup-0001-Disclosureform.pdf.

Address correspondence via email to Daniel H. Solomon, MD, MPH, at dsolomon@bwh.harvard.edu.

Submitted for publication December 7, 2022; accepted in revised form February 16, 2023.

SIGNIFICANCE & INNOVATIONS

- Clinical visits are a fundamental aspect of rheumatic disease care, but recommendations for appropriate visit frequencies are largely absent from guidelines, scarcely studied, and inconsistently reported.
- We included 28 relevant manuscripts in this systematic literature review; the included studies were equally divided between US and non-US and were published between 1985 and 2021.
- For rheumatoid arthritis, the average annual visit frequencies were 5.25 for US rheumatologists, 4.80 for US non-rheumatologists, 3.29 for non-US rheumatologists, and 2.74 for non-US non-rheumatologists.
- Trends suggest more frequent visits in the US and less frequent visits in recent years.

patient outcome, inform policy planning, and project future service needs and health care expenditures (10).

With the goal of better understanding visit frequency in rheumatic diseases and how visit frequency might differ across countries, we conducted a systematic literature review to identify studies that have estimated visit frequency (the mean number of annual patient visits to rheumatologists and other specialists) for major rheumatic diseases. The review included literature from 1946 to the present and systematically summarized and analyzed existing practices of clinical visit frequency with resolution specific to country/state and disease. Where possible, we created summary statistics, but heterogeneity in the included literature precluded a formal meta-analysis.

MATERIALS AND METHODS

Study design and search strategy. A predefined study protocol (https://www.crd.york.ac.uk/PROSPERO/display_record.php?RecordID=306299) was developed and deposited with PROSPERO. Most systematic literature reviews begin with developing a PICO (population, intervention, comparator, outcomes) table. However, we demanded no intervention or comparison, thus a PICO table was unnecessary.

The search strategy included all primary articles, with the exception of systematic reviews (including randomized trials, cohort studies, case–control studies, and cross-sectional studies) that report a measure of clinical visit frequency in rheumatology or general practices. To be as inclusive as possible, no setting or type of rheumatologic diagnosis was set a priori. All studies published prior to February 2, 2022, the date on which the search was last run, were considered eligible.

Due to the highly challenging nature of precisely defining concepts of a clinical visit in the literature, multiple iterations of search terms were generated, with the help of Yale MeSH analyzer (11), and calibrated with a suite of 11 key papers deemed by the authors to adequately capture the breadth of this field. Using the search terms generated from this strategy, one author (YJ) searched the following databases: OVID Medline, Cumulative Index to Nursing and Allied Health Literature (CINAHL), and World Health Organization Global Index Medicus since their inception, with the input of a medical librarian. See Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25106, for the final list of search terms; in brief, it includes the following concepts: house calls, home visits, office visits, clinic visits, frequency (weekly, monthly, yearly), and rheumatology.

Study selection and data extraction. Two reviewers (YJ and DHS) independently scanned all abstracts and potentially eligible full-text articles to determine eligibility for inclusion. Discrepancies in judgements between the reviewers were discussed until consensus was reached. Using a standard data extraction sheet, information was extracted on study and patient characteristics, including the following: year of publication; study period start and end dates; sources of data (i.e., clinical record, administrative database); geographic location and setting of rheumatology care; rheumatic disease diagnosis; and provider type. The outcome of interest was clinical visit frequency in mean number of visits per patient per year. For a small group of studies, the outcome data were not presented in detail, and authors were contacted to help with relevant numbers.

The validity of individual trials was assessed using the Risk of Bias 2.0 instrument, endorsed by the Cochrane Collaboration, and risk of bias for observational studies was assessed using a modified version of the Newcastle-Ottawa tool (see Supplementary Figure 1 and Supplementary Table 1, respectively, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25106). The tool was modified to include only the outcome assessment measures; as no comparison group was required, the rest of the tool was not applicable.

Statistical analyses. We categorized studies by rheumatic disease and by country. The preponderance of studies were from the US; we separately categorized studies from the US and studies from non-US settings. Annual visit frequencies from each study were combined using a weighted mean, weighted on the number of patients included. Some studies gave SDs or medians and interquartile ranges, but not all of them did, thus SEs could not be estimated across studies.

RESULTS

Study eligibility and selection. As shown in Supplementary Figure 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25106, a total of 273 records were identified (7 were duplicates), of which 267 were examined for title and abstract eligibility. Of those, 65 were retrieved for full-text screening and eligibility assessment. Of the 65 studies, 22 did not contain disease-specific data on visit

frequency, 1 was only available as an abstract, and 14 studies could not be confirmed regarding the availability of potential unpublished data for calculating outcomes of interest due to our inability to make contact, despite our inquiries. In the end, 28 studies were included for data extraction.

Study characteristics. The 28 studies included for clinical visit frequency were equally divided by countries of origin into US and non-US, with the majority (85%) from North America and Europe (Tables 1 and 2). The studies were published between 1985 and 2021. Rheumatoid arthritis (RA) was the most frequently reported condition (16 mentions), followed by systemic lupus erythematosus (SLE) (5 mentions) and fibromyalgia (FM) (4 mentions). The mean study period was 4.5 years, with a maximum of 29 years and a minimum of <1 year. In 21 (75.0%) of studies, data on clinical visit frequency were derived from clinic and hospital records, 5 (17.9%) from public health care registries, and the remaining 2 (7.1%) from private insurance claims. One-half of included studies reported data from a community rheumatology practice setting, 7 did not specify the setting, 6 reported hospital-based outpatient practices, and the remaining 1 study reported "primary and secondary care" in the Swedish context (which is not easily convertible into the above categories). Seventy-four percent of all studies reported data on visit frequency for rheumatologists only, with the remainder including a mixture of general practitioners (GPs), internists, and rheumatologists. No noticeable difference in study characteristics was observed between those focusing on RA and the non-RA studies or between the US and non-US studies.

Clinical visit frequencies across the disease spectrum. A wide range of mean clinical visit frequencies was

Table 1. Characteristics of included studies with a reported diagnosis that included but was not limited to rheumatoid arthritis (RA), grouped by location and type of provider*

Author, year (ref.)	Study period	Data source	Setting and location	Provider
US studies				
Yelin et al, 1985 (25)	1982–1983	Clinic record†	Community clinics; Northern California	Rheumatologists, internist
Yelin et al, 1996 (26)	1982–1989	Clinic record	Community clinics; Northern California	Rheumatologists
Criswell et al, 1997 (27)	1982–1983	UCSF longitudinal RA panel	Community clinics; Northern California	Rheumatologists
Ward 1997 (6)	1979–1981	Stanford Outcomes in RA study	Community clinic; California	Rheumatologists
Katz et al, 1998 (28)‡	1993–1993	Medicare claims (excluding HMOs)	Community clinic; Colorado, Massachusetts, and Virginia	Rheumatologists
Gabriel et al, 2001 (29)	1987–1994	Rochester Epidemiology Project	Hospital-based outpatient clinic; Minnesota	Rheumatologists
Ethgen et al, 2002 (30)§	1996–1998	Hospital record	Hospital-based outpatient clinic; Kansas	Rheumatologists
Bartels et al, 2011 (31)	2004–2006	Medicare claims	Rheumatologist visits; any location, random US national sample	Rheumatologists
Accortt et al, 2017 (32)	2010–2013	Truven Health Analytics MarketScan database	Rheumatologist visits; any location, US national sample	NA
Non-US studies				
Chan et al, 2008 (33)	2001–2006	Rheumatology service database	Outpatient clinic; New Zealand	Rheumatologists
Hagel et al, 2013 (34)	2001–2011	Skåne Healthcare Register	All settings; Sweden	Physician
Bengtson et al, 2016 (35)	2006–2010	Hospital record	Primary care; Sweden	PCP and secondary care
McBain et al, 2016 (36)¶	NA	Hospital record	Hospital-based outpatient clinic; London, UK	Rheumatologists
Barnabe et al, 2017 (37)#	1993–2011	Public claim database (AHCIP)	Outpatient clinics; Canada	PCP
Kim et al, 2020 (38)	2019	Clinical record data	Hospital-based outpatient clinic; South Korea	Rheumatologists
Muskens et al, 2021 (39)	2014–2019	Hospital record	Hospital-based outpatient clinic; The Netherlands	Rheumatologists
Poggenborg et al, 2021 (7)	2015–2017	Hospital record	Hospital-based outpatient clinic; Denmark	Rheumatologists

* AHCIP = Alberta Health Care Insurance Plan; HMO = health maintenance organization; NA = not available; PCP = primary care provider; ref. = reference; UCSF = University of California San Francisco.
† Clinical records obtained via telephone survey.
‡ Osteoarthritis (any site), mechanical spinal disorders, bursitis, tendinitis, fibromyalgia, polymyalgia rheumatica, systemic lupus erythematosus, all rheumatologic diagnoses.
§ Osteoarthritis.
¶ Psoriatic arthritis.
# Ankylosing spondylitis, psoriatic diseases, crystal-related arthritis.

Table 2. Characteristics of included studies with a reported diagnosis that did not include rheumatoid arthritis (RA), grouped by location and type of provider*

Author, year (ref.)	Study period	Diagnosis	Data source	Setting and location	Provider
US studies					
Kremers et al, 2005 (40)	1970–1999	PMR	Truven Health Analytics MarketScan database	All settings; Olmsted County, Minnesota	Rheumatologists, generalists
Molina et al, 2008 (41)	2003–2003	SLE	Triple-S insurance record	All settings; Puerto Rico	Rheumatologists
Julian et al, 2009 (42)	2002–2004	SLE	UCSF Lupus Outcomes Study	All settings; US national sample	Rheumatologists
Singh et al, 2011 (43)	NA	Crystal-related arthritis	Clinic and hospital record	Outpatient clinic, Hospital outpatient, and VA clinics; San Diego, Cincinnati, and Minneapolis	Rheumatologists
Chandran et al, 2012 (44)	2018–2019	FM	Clinic record (convenience sample)	Outpatient clinic; US national sample	GPs, rheumatologists, neurologists, or psychiatrists
Non-US studies					
Badley et al, 2015 (45)	2007–2008	Inflammatory arthritis	Ontario Health Insurance Plan	NA; Canada	Rheumatologists
Elek et al, 2015 (46)	2008–2012	General rheumatic disease	GYEMSZI database	Outpatient clinics; Hungary	All physicians, rheumatologists
Andrés et al, 2016 (47)	2009–2010	SpA	Hospital record	Outpatient clinics; Spain	Rheumatologists, internists
Valent et al, 2020 (48)	2009–2020	GCA	Hospital record	Outpatient clinics; Italy	Rheumatologists, internists
Gendelman et al, 2021 (12)	1998–2014	FM	HMO group record	Outpatient clinics; Israel	Rheumatologists
Winkelmann et al, 2011 (13)	2011–2011	FM	Clinic record	Outpatient clinics; France and Germany	GPs and rheumatologists

* FM = fibromyalgia; GCA = giant cell arteritis; GP = general practitioner; GYEMSZI = National Institute for the Quality and Organizational Development in Healthcare and Medicines (of Hungary); HMO = health maintenance organization; NA = not available; PMR = polymyalgia rheumatica; ref. = reference; SLE = systemic lupus erythematosus; SpA = spondyloarthritis; VA = Veterans Affairs.

reported across rheumatic diseases and geographic locations. For RA (Table 3), the highest weighted mean visit frequency was reported for US-based rheumatologists at a rate of 5.25 visits per patient-year (this unit will apply hereafter unless stated otherwise), followed by US-based non-rheumatologists (4.80), non-US rheumatologists (3.29), and non-US non-rheumatologists (2.74). A temporal decrease in visit frequencies for RA can generally be seen among US-based rheumatologists, from 13.1 in 1985 (at fee-for-service practices) to ∼2.0 in the first decade of the twenty-first century. On the contrary, the reverse trend was observed in non-US–based rheumatologists, from 1.9 in New Zealand in 2008 to 3.4 in South Korea in 2022 to close to 4.0 in Western and Northern Europe in 2021. The highest frequency of RA clinical visits was reported in a US study from 1985 at 13.1, whereas the lowest was reported in 2001 in a hospital-based outpatient study in Minnesota at a rate of 1.3.

For FM (Table 4), the weighted mean visit frequency for US-based rheumatologists of 1.8 was substantially higher than that for non-US–based rheumatologists at 0.40. However, if analyzed with the exclusion of the Israeli study (12), Western European visit frequencies were higher than in the US. A 2012 study using a large US national sample demonstrated a positive association between visit frequency and FM severity measured with the Fibromyalgia Impact Questionnaire score: patients with mild FM met with their rheumatologists a mean of 2.6 times per year, whereas those with severe FM had 6.6 visits per year, the highest among all included FM studies. This trend was replicated in Germany and France in the 2011 European study (13). On the opposite end of the spectrum, the 2021 Israeli study reported the lowest FM annual visit frequency at 0.32.

For SLE (Table 5), no non-US studies met the eligibility criteria. There was a dramatic difference between visit frequencies to non-rheumatologists (12.3) and those to rheumatologists (3.2). Compared to medication-adherent patients, nonadherent patients were reported to have visited rheumatologists (4.34 versus 3.24) and other physicians (22.43 versus 18.52) more

Table 3. Visit frequencies for rheumatoid arthritis (RA)*

Author, year (ref.)	Sample size, no. of patients	Mean annual visit frequency
US studies, rheumatologists		
Yelin et al, 1985 (25)		
Fee for service	626	13.16
HMOs	186	10.15
Yelin et al, 1996 (26)		
Fee for service	798	9.91
HMOs	227	9.83
Criswell et al, 1997 (27)	310	6.56
Ward 1997 (6)	127	7.20
Katz et al, 1998 (28)	8,027	5.40
Gabriel et al, 2001 (29)	249	1.18
Ethgen et al, 2002 (30)	642	1.97
Bartels et al, 2011 (31)†	3,298	2.40
Total, weighted	14,490	5.25
US studies, non-rheumatologists		
Accortt et al, 2017 (32)‡	6,737	4.8
Non-US studies, rheumatologists		
Chan et al, 2008 (33)		
New RA	26	2.60
Flaring with new DMARD	177	2.55
Stable taking bDMARD	170	1.48
Stable taking NSAID or csDMARD	75	1.22
Kim et al, 2020 (38)	378	3.40
Muskens et al, 2021 (39)§	1,059	3.8
Poggenborg et al, 2021 (7)	117	3.5
Total, weighted	2,002	3.29
Non-US–based non-rheumatologists		
Barnabe et al, 2017 (37)	93,490	1.00
Hagel et al, 2013 (34)	3,977	8.3
Bengtsson et al, 2016 (35)		
Primary care provider	7,712	8.0
Secondary care	7,712	13
Total, weighted	105,179	2.74

* bDMARD = biologic disease-modifying antirheumatic drug; csDMARD = conventional synthetic DMARD; HMO = health maintenance organization; NSAID = nonsteroidal antiinflammatory drug; ref. = reference.
† Represents the total visits for any reason to a rheumatologist among patients meeting RA definitions using International Classification of Diseases algorithms.
‡ Physician specialty not available.
§ Interrupted time series (calculated using baseline data prior to intervention).

frequently and to have visited GPs less frequently (6.27 versus 6.43); adherent and nonadherent patients were defined based on reporting that forgetting medications was never a problem versus reporting that it was a problem at least some of the time. The lowest of all reported SLE annual visit frequencies was in Puerto Rico, at 1.5 visits per patient-year.

Due to the paucity of studies pertaining to each individual condition, visit frequencies for all non-RA–arthritis diagnoses were grouped into Table 6. Only 2 US studies reported on clinical visit frequencies for OA, with a weighted average of 2.15. Striking differences can be seen in the rates of clinical visits between US regions and Alberta, Canada, for crystal-related arthritis, with the former reporting 2.0 visits to rheumatologists per patient-year, while the latter reported 0.034 visits per patient-year, including both rheumatologists and internists. Visit frequencies to rheumatologists and non-rheumatologists were similar for psoriatic diseases, at ~0.25 per patient per year. This was not the case for the spondyloarthritides, which showed a 4-fold difference between rheumatologist and non-rheumatologist visits.

DISCUSSION

This systematic review identified and summarized estimated visit frequencies for major rheumatic diseases reported in the literature from 1946 to the present. This exercise helps shed light on the sparse data concerning rheumatology visit frequencies. We have attempted to compare visit frequencies across time, across disease categories, and across geographic locations. Of note, our comprehensive results spanned 5 of 7 World Bank region groups (14) (South America and sub-Saharan Africa sources were absent). Of the 28 included studies from a total of 272 unique records examined, the great majority were from North America and European centers, involved RA, featured community rheumatology practices, and consisted solely of rheumatologists. Weighted mean visit frequencies for RA were 5.25, 4.80, 3.29, and 2.7 per year for US rheumatologists, US non-rheumatologists, non-US rheumatologists, and non-US non-rheumatologists, respectively. FM patients made 1.80 and 0.40 annual visits to US and non-US rheumatologists, respectively. For US rheumatologists and non-rheumatologists, visit frequency for SLE was 3.2 and 12.3, respectively.

Two general trends observed in our data warrant examination. First, visit frequencies were higher in US studies as compared to non-US studies for all conditions in which such comparisons were possible, especially for RA and FM (Tables 3 and 4). The US has one of the highest per capita rheumatology workforces, at 1.78 per 100,000 people. The American College of Rheumatology (ACR) workforce report describes a deficit, but the US rheumatology workforce appears to have increased from 4,049 US rheumatologists according to 2005 data (15) to 5,602 in 2022 (16). When it comes to per capita rheumatologists, focusing on the countries included in Table 3, the US, with 1.78 rheumatologists per 100,000 people (10), exceeds New Zealand (0.59) (17), South Korea (0.60) (18), and The Netherlands (0.61) (19). Differing guideline recommendations could also play a role in the differing data on visit frequency; while both the ACR and EULAR recommend treat-to-target strategies for RA management, the ACR gave a loose, minimum monitoring frequency of at least every 6 months (20), whereas EULAR specified an interval of 1–3 months during active disease (21), potentiating variable visit frequencies due to individual provider preferences. Other guidelines on disease treatment from the ACR and EULAR provide no clear recommendations on visit frequency.

Table 4. Visit frequencies for fibromyalgia (FM)*

Author, year (ref.)	Sample size, no.	Mean annual visit frequency, no. per patient-years
US studies, rheumatologists		
Katz et al, 1998 (28)	2,971	1.8
US studies, non-rheumatologists		
Chandran et al, 2012 (44)†		
Mild FM	21	2.7
Moderate FM	49	5.2
Severe FM	133	6.9
Non-US studies, rheumatologists		
Gendelman et al, 2021 (12)	14,269	0.32
Winkelmann et al, 2011 (13)		
France total	88	2.9
Mild	17	3.0
Moderate	33	2.5
Severe	38	3.4
Germany total	211	4.9
Mild	52	4.1
Moderate	66	4.6
Severe	93	5.5
Total, weighted	14,568	0.401

* Ref. = reference.
† Severity-level classification was based on Fibromyalgia Impact Questionnaire score (0 to <39 = mild; 39 to <59 = moderate; 59 to 100 = severe). Data are for primary care physicians, rheumatologists, neurologists, and psychiatrists.

Second, it is evident from Table 3 that a clear trend exists for reduction in rheumatologist visits for RA over time, from 13.1 visits per year in 1985 to 2.4 visits per year in 2011. The introduction of conventional synthetic disease-modifying antirheumatic drugs (csDMARDs) such as methotrexate fundamentally changed RA management and patient outcomes (22), and the clinical approval of tumor necrosis factor inhibitors by the Food and Drug Administration in 1999 heralded the age of biologic DMARDs in the treatment of RA. Despite the need for intensive monitoring at the early stage of initiation, advanced therapies allow for sustained long-term low disease activity or remission that would gradually reduce the need for frequent rheumatologist visits, which is reflected by the temporal trend in Table 3.

Our systematic review has several strengths. For one, it is the first and only study of which we are aware that provides a global, comprehensive picture pertaining to clinical visit frequencies across the spectrum of common rheumatic conditions, covering a consequential 4-decade period in the history of rheumatology, from 1985 to 2021. An additional strength of this review is that search terms were engineered using Yale MeSH analyzer and calibrated with key citations. Finally, a third strength of our review, beyond its strict adherence to the PRISMA workflow, is its broad coverage of major medical literatures (Medline), allied health literature (CINAHL), and global gray literature (WHO Global Index Medicus).

Our review has several limitations. Due to the nature and heterogeneity of the literature on clinical visit frequency, most if not all of the included studies reported study visit frequency as a secondary outcome. Not being the main focus of their respective reports, data on visit frequency tended to lack details, such as precise provider types (including advanced practice providers), measurements of statistical variance, disease severity, settings, and context such as clinical visits versus laboratory visits, although the latter may diminish in the future with the emergence of novel care models such as at-home blood sampling (23) and one-stop clinics (24). Future studies in rheumatology visit frequencies should strive to include many if not all of the aforementioned details to allow for a more granular understanding of the factors influencing the manner by which care is being utilized. It is also possible that visit frequency will continue to change as technology enables us to communicate with patients in previously unobserved ways. Furthermore, we were not able to assess reports written in languages other than English. Finally, the lack of existing and validated quality assessment tools for studies in qualitative health service research compelled us to adopt and modify the Ottawa-Newcastle scale, which may not be as useful in the context of this review.

In conclusion, we found the literature on clinical visit frequency to be sparse and methodologically heterogeneous, focusing mainly on RA and published mainly in the US. Studies reported more frequent visits to US versus non-US rheumatologists and universally showed a decreasing temporal trend of visit frequencies in the US. Future studies are strongly encouraged to focus on visit frequencies across time as the primary outcome, to provide rigorous definitions of the nature of reported visits, to clarify in detail the setting of visits and type of providers, and, for

Table 5. Visit frequencies to rheumatologists for systemic lupus erythematosus*

Author, year (ref.)	Sample size, no.	Mean annual visit frequency, no. per patient-years
US rheumatologists		
Katz et al, 1998 (28)	783	2.7
Julian et al, 2009 (42)		
Adherent	454	3.24
Nonadherent	380	4.34
Molina et al, 2008 (41)	No data	3.0
Total, weighted	1,617	3.24
US non-rheumatologists		
Julian et al, 2009 (42)		
Adherent, GP visits	454	3.80
Adherent, other HCP	454	18.52
Nonadherent, GP visits	380	4.78
Nonadherent, other HCP	380	22.43
Molina et al, 2008 (41)	No data	1.5
Total, weighted	1,668	12.3

* GP = general practitioner; HCP = health care provider; ref. = reference.

Table 6. Visit frequencies for other arthritis*

Author, year (ref.) and diagnosis	Sample size, no.	Visit frequency, no. per patient-years	Comments
US studies			
Singh et al, 2011 (43)			
Gout	285	1.85	
Katz et al, 1998 (28)			
OA	15,715	2.2	OA for any site; median = 2
Mechanical spinal disorders	7,334	2.1	Median = 1
Bursitis, tendinitis		1.9	Median = 1
PMR		3.3	Median = 2
All rheumatologic diagnoses		3.8	Median = 3
Ethgen et al, 2002 (30)			
OA	395	0.097	
Kremers et al, 2005 (40)			
PMR	364	0.84	Rheumatologist
PMR		3.8	Generalist
Elek et al, 2015 (46)			
General rheumatic disease	430,000	0.021	Extensive margin fixed-effects logit equation
General rheumatic disease		0.0068	Intensive margin fixed-effects truncated Poisson equation
General rheumatic disease		0.19	Pooled zero-inflated negative binomial model
Valent et al, 2020 (48)			
GCA	208	0.71	Rheumatologist
GCA		0.65	Internal medicine
Non-US studies			
McBain et al, 2016 (36)			
PsA mixed with RA	48	2.15	
Barnabe et al, 2017 (37)			
Psoriasis and PsA	6,040	0.27	Rheumatologist
Psoriasis and PsA	6,040	0.25	Non-rheumatologist
Crystal-related arthritis	44,221	0.51	PCP; crystal-related arthritis
Crystal-related arthritis		0.034	Rheumatologists or internists; crystal-related arthritis
AS	7,685	0.37	Rheumatologist
AS		0.22	Non-rheumatologist
Andrés et al, 2016 (47)			
SpA	1,168	2.0	Rheumatologist
SpA	1,168	0.50	Non-rheumatologist
Badley et al, 2015 (45)			
Inflammatory arthritis	No data	0.0069	All physicians
All arthritis	No data	0.012	All physicians

* AS = ankylosing spondylitis; GCA = giant cell arteritis; OA = osteoarthritis; PCP = primary care provider; PMR = polymyalgia rheumatica; PsA = psoriatic arthritis; RA = rheumatoid arthritis; ref. = reference; SpA = spondyloarthritis.

any quantitative data, to report on measures of statistical variance. To advance the field of research in clinical visit frequency, there is a need to develop validated methods to link patient disease activity scores (or prognostic data) with visit frequencies. Doing so would shed light on the implications of clinical visits for disease progression (or the lack thereof), thereby allowing for the determination of the ideal visit frequency for each common rheumatic disease by professional societies in their guidelines, potentially eliminating unnecessary visits, and hence health care expenditure, and optimizing resource utilization.

ACKNOWLEDGMENT

We thank Ms. Susanna Galbraith, MLIS, virtual services librarian, Health Sciences Library, McMaster University, for her assistance in the construction of the search terms.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. Dr. Solomon had full access to all

of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.
Study conception and design. Jiang, Solomon.
Acquisition of data. Jiang, Solomon.
Analysis and interpretation of data. Jiang, Rudin, Solomon.

ADDITIONAL DISCLOSURES

Author Rudin is an employee of the RAND Corporation.

REFERENCES

1. American College of Rheumatology. What is a rheumatologist? URL: https://www.rheumatology.org/I-Am-A/Patient-Caregiver/Health-Care-Team/What-is-a-Rheumatologist.
2. Duncan R, Cheng L, Law MR, et al. The impact of introducing multidisciplinary care assessments on access to rheumatology care in British Columbia: an interrupted time series analysis. BMC Health Serv Res 2022;22:327.
3. Ganguli I, Wasfy JH, Ferris TG. What is the right number of clinic appointments? Visit frequency and the accountable care organization. JAMA 2015;313:1905–6.
4. National Guideline Centre (UK). Frequency of monitoring: rheumatoid arthritis in adults: diagnosis and management: evidence review E. National Institute for Health and Care Excellence (NICE); 2018. URL: http://www.ncbi.nlm.nih.gov/books/NBK577121/.
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;72:149–62.
6. Ward MM. Rheumatology visit frequency and changes in functional disability and pain in patients with rheumatoid arthritis. J Rheumatol 1997;24:35–42.
7. Poggenborg RP, Madsen OR, Sweeney AT, et al. Patient-controlled outpatient follow-up on demand for patients with rheumatoid arthritis: a 2-year randomized controlled trial. Clin Rheumatol 2021;40:3599–604.
8. Clough JD, Patel K, Shrank WH. Variation in specialty outpatient care patterns in the Medicare population. J Gen Intern Med 2016;31:1278–86.
9. Zachrison KS, Yan Z, Schwamm LH. Changes in virtual and in-person health care utilization in a large health system during the COVID-19 pandemic. JAMA Netw Open 2021;4:e2129973.
10. Al Maini M, Adelowo F, Al Saleh J, et al. The global challenges and opportunities in the practice of rheumatology: white paper by the World Forum on Rheumatic and Musculoskeletal Diseases. Clin Rheumatol 2015;34:819–29.
11. Yale MeSH Analyzer. URL: https://mesh.med.yale.edu/.
12. Gendelman O, Shapira R, Tiosano S, et al. Utilisation of healthcare services and drug consumption in fibromyalgia: a cross-sectional analysis of the Clalit Health Service database. Int J Clin Pract 2021;75:e14729.
13. Winkelmann A, Perrot S, Schaefer C, et al. Impact of fibromyalgia severity on health economic costs: results from a European cross-sectional study. Appl Health Econ Health Policy 2011;9:125–36.
14. The World Bank. World Bank country and lending groups. World Bank Data Help Desk. URL: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups\.
15. Hogan PF, Bouchery E. Workforce study of rheumatologists: final report. American College of Rheumatologists; 2006. URL: https://www.rheumatology.org/Portals/0/Files/LewinReport.pdf.
16. Battafarano DF, Ditmyer M, Bolster MB, et al. 2015 American College of Rheumatology workforce study: supply and demand projections of adult rheumatology workforce, 2015–2030. Arthritis Care Res (Hoboken) 2018;70:617–26.
17. Harrison AA, Tugnet N, Taylor WJ. A survey of the New Zealand rheumatology workforce. N Z Med J 2019;132:70–6.
18. Lee CU, Kim JN, Kim JW, et al. Korean rheumatology workforce from 1992 to 2015: current status and future demand. Korean J Intern Med 2019;34:660–8.
19. Miedema HS, van der Linden SM, Rasker JJ, et al. National database of patients visiting rheumatologists in The Netherlands: the standard diagnosis register of rheumatic diseases. A report and preliminary analysis. Br J Rheumatol. 1998;37:555–61.
20. 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;73:924–39.
21. Smolen JS, Landewé RB, Bijlsma JW, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2019 update. Ann Rheum Dis 2020;79:685–99.
22. Case JP. Old and new drugs used in rheumatoid arthritis: a historical perspective. Part 1: the older drugs. Am J Ther 2001;8:123–43.
23. Muehlensiepen F, May S, Zarbl J, et al. At-home blood self-sampling in rheumatology: a qualitative study with patients and health care professionals. BMC Health Serv Res 2022;22:1470.
24. Väre P, Nikiphorou E, Hannonen P, et al. Delivering a one-stop, integrated, and patient-centered service for patients with rheumatic diseases. SAGE Open Med 2016;4:2050312116654404.
25. Yelin EH, Henke CJ, Kramer JS, et al. A comparison of the treatment of rheumatoid arthritis in health maintenance organizations and fee-for-service practices. New Engl J Med 1985;312:962–7.
26. Yelin EH, Criswell LA, Feigenbaum PG. Health care utilization and outcomes among persons with rheumatoid arthritis in fee-for-service and prepaid group practice settings. JAMA 1996;276:1048–53.
27. Criswell LA, Such CL, Neuhaus JM, et al. Variation among rheumatologists in clinical outcomes and frequency of office visits for rheumatoid arthritis. J Rheumatol 1997;24:1266–71.
28. Katz JN, Barrett J, Liang MH, et al. Utilization of rheumatology physician services by the elderly. Am J Med 1998;105:312–8.
29. Gabriel SE, Wagner JL, Zinsmeister AR, et al. Is rheumatoid arthritis care more costly when provided by rheumatologists compared with generalists? Arthritis Rheum 2001;44:1504–14.
30. Ethgen O, Kahler KH, Kong SX, et al. The effect of health related quality of life on reported use of health care resources in patients with osteoarthritis and rheumatoid arthritis: a longitudinal analysis. J Rheumatol 2002;29:1147–55.
31. Bartels CM, Kind AJ, Everett C, et al. Low frequency of primary lipid screening among Medicare patients with rheumatoid arthritis. Arthritis Rheum 2011;63:1221–30.
32. Accortt NA, Schenfeld J, Chang E, et al. Changes in healthcare utilization after etanercept initiation in patients with rheumatoid arthritis: a retrospective claims analysis. Adv Ther 2017;34:2093–103.
33. Chan G, Goh F, Hodgson T, et al. Outpatient follow-up for patients with rheumatoid arthritis in relation to New Zealand Rheumatology Association guidelines at Dunedin Hospital. N Z Med J 2008;121:34–41.
34. Hagel S, Petersson IF, Bremander A, et al. Trends in the first decade of 21st century healthcare utilisation in a rheumatoid arthritis cohort compared with the general population. Ann Rheum Dis 2013;72:1212–6.
35. Bengtsson K, Jacobsson LT, Rydberg B, et al. Comparisons between comorbid conditions and health care consumption in rheumatoid arthritis patients with or without biological disease-modifying

anti-rheumatic drugs: a register-based study. BMC Musculoskelet Dis 2016;17:499.

36. McBain H, Shipley M, Olaleye A, et al. A patient-initiated DMARD self-monitoring service for people with rheumatoid or psoriatic arthritis on methotrexate: a randomised controlled trial. Ann Rheum Dis 2016; 75:1343–9.
37. Barnabe C, Jones CA, Bernatsky S, et al. Inflammatory arthritis prevalence and health services use in the First Nations and non-First Nations populations of Alberta, Canada. Arthritis Care Res (Hoboken) 2017;69:467–74.
38. Kim Y, Ahn E, Lee S, et al. Changing patterns of medical visits and factors associated with no-show in patients with rheumatoid arthritis during COVID-19 pandemic. J Korean Med Sci 2020;35:e423.
39. Muskens WD, Rongen-van Dartel SA, Vogel C, et al. Telemedicine in the management of rheumatoid arthritis: maintaining disease control with less health-care utilization. Rheumatology Adv Pract 2021;5: rkaa079.
40. Kremers HM, Reinalda MS, Crowson CS, et al. Use of physician services in a population-based cohort of patients with polymyalgia rheumatica over the course of their disease. Arthritis Rheum 2005;53:395–403.
41. Molina MJ, Mayor AM, Franco AE, et al. Utilization of health services and prescription patterns among lupus patients followed by primary care physicians and rheumatologists in Puerto Rico. Ethn Dis 2008; 18 Suppl 2:S2–205–10.
42. Julian LJ, Yelin E, Yazdany J, et al. Depression, medication adherence, and service utilization in systemic lupus erythematosus. Arthritis Rheum 2009;61:240–6.
43. Singh JA, Sarkin A, Shieh M, et al. Health care utilization in patients with gout. Sem Arthritis Rheum 2011;40:501–11.
44. Chandran A, Schaefer C, Ryan K, et al. The comparative economic burden of mild, moderate, and severe fibromyalgia: results from a retrospective chart review and cross-sectional survey of working-age U.S. adults. J Manag Care Pharm 2012;18:415–26.
45. Badley EM, Canizares M, Gunz AC, et al. Visits to rheumatologists for arthritis: the role of access to primary care physicians, geographic availability of rheumatologists, and socioeconomic status. Arthritis Care Res (Hoboken) 2015;67:230–9.
46. Elek P, Varadi B, Varga M. Effects of geographical accessibility on the use of outpatient care services: quasi-experimental evidence from panel count data. Health Econ 2015;24:1131–46.
47. Andrés M, Sivera F, Perez-Vicente S, et al. Centre characteristics determine ambulatory care and referrals in patients with spondyloarthritis. Rheumatol Int 2016;36:1515–23.
48. Valent F, Bond M, Cavallaro E, et al. Data linkage analysis of giant cell arteritis in Italy: healthcare burden and cost of illness in the Italian region of Friuli Venezia Giulia (2001–2017). Vasc Med 2020;25:150–6.

Arthritis Care & Research
Vol. 75, No. 10, October 2023, pp 2063–2072
DOI 10.1002/acr.25129

American College of Rheumatology
Empowering Rheumatology Professionals

Incidence and Risk Factors for Eosinophilia and Lung Disease in Biologic-Exposed Children With Systemic Juvenile Idiopathic Arthritis

Holly Wobma,[1] Sage R. Arvila,[1] Maria L. Taylor,[1] Ki Pui Lam,[1] Marina Ohashi,[2] Catherine Gebhart,[2] Helene Powers,[1] Siobhan Case,[3] Mia T. Chandler,[1] Margaret H. Chang,[1] Ezra Cohen,[1] Megan Day-Lewis,[1] Martha P. Fishman,[1] Olha Halyabar,[1] Jonathan S. Hausmann,[1] Melissa M. Hazen,[1] Pui Y. Lee,[1] Mindy S. Lo,[1] Esra Meidan,[1] Jordan E. Roberts,[1] Mary Beth F. Son,[1] Robert P. Sundel,[1] Fatma Dedeoğlu,[1] Peter A. Nigrovic,[3] Alicia Casey,[1] Joyce Chang,[1] and Lauren A. Henderson[1]

Objective. Although interleukin-1 (IL-1)/IL-6 inhibitors are effective therapies for systemic juvenile idiopathic arthritis (JIA), some patients develop eosinophilia and lung disease during treatment. This study was undertaken to retrospectively evaluate incidence and risk factors for eosinophilia and describe lung disease outcomes in IL-1/IL-6 inhibitor–exposed patients with systemic JIA.

Methods. Among JIA patients at our institution exposed to interleukin-1 (IL-1)/IL-6 inhibitors (1995–2022), we compared incidence rate of eosinophilia in systemic JIA compared to other JIA, stratified by medication class (IL-1/IL-6 inhibitors, other cytokine inhibitors, methotrexate). We used Cox models to identify predictors of eosinophilia during IL-1/IL-6 inhibitor use and summarized treatment changes and outcomes after eosinophilia, including lung disease. HLA typing was performed on a clinical or research basis.

Results. There were 264 new medication exposures in 75 patients with systemic JIA and 41 patients with other JIA. A total of 49% of patients with systemic JIA with HLA typing (n = 45) were positive for HLA–DRB1*15 alleles. Eosinophilia was common during IL-1/IL-6 inhibitor use and did not differ by systemic JIA compared to other JIA (0.08 and 0.07 per person-year, respectively; $P = 0.30$). Among systemic JIA patients, pretreatment macrophage activation syndrome (MAS) was associated with a higher rate of subsequent eosinophilia on biologic therapy (unadjusted hazard ratio 3.2 [95% confidence interval 1.2–8.3]). A total of 4 of 5 patients who switched therapy within 10 weeks of eosinophilia experienced disease flare compared to none of the patients who continued the original therapy. A total of 8 of 25 patients with pulmonary evaluations had lung disease, and all had severe manifestations of systemic JIA (MAS, intensive care unit stay). One death was attributed to systemic JIA–lung disease.

Conclusion. Eosinophilia is common in JIA patients using IL-1/IL-6 inhibitors. Severe disease may be associated with eosinophilia and lung disease in systemic JIA.

INTRODUCTION

Systemic juvenile idiopathic arthritis (JIA) is a category of JIA characterized by quotidian fever, evanescent rash, and arthritis, and is frequently complicated by macrophage activation syndrome (MAS) (1–3). Historically, outcomes in children with systemic JIA who receive treatment with conventional arthritis therapies (methotrexate [MTX] and/or tumor necrosis factor

Dr. Wobma's work was supported by the NIH/National Institute of Allergy and Infectious Diseases (grant T32-AI-00751236). Dr. Roberts's work was supported by the Childhood Arthritis and Rheumatology Research Alliance, the NIH/National Institute of Allergy and Infectious Diseases (grant T32-AI-00751235), and the Lupus Foundation of America (Career Development award). Dr. Nigrovic's work was supported by the NIH/National Institute of Arthritis and Musculoskeletal and Skin Diseases (grant P30-AR-070253). Dr. Chang's work was supported by the NIH/National Heart, Lung, and Blood Institute (grant K23-HL-148539). Dr. Henderson's work was supported by the NIH/National Institute of Arthritis and Musculoskeletal and Skin Diseases (grants K08-AR-073339 and P30-AR-070253), the Rheumatology Research Foundation (Career Development Bridge funding), and the Arthritis National Research Foundation (All Arthritis grant).

[1]Holly Wobma, MD, PhD, Sage R. Arvila, BSc, Maria L. Taylor, BSc, Ki Pui Lam, PhD, Helene Powers, RN, MSN, FNP, Mia T. Chandler, MD, MPH, Margaret H. Chang, MD, PhD, Ezra Cohen, MD, Megan Day-Lewis, RN, MSN, CPNP, Martha P. Fishman, MD, Olha Halyabar, MD, Jonathan S. Hausmann, MD, Melissa M. Hazen, MD, Pui Y. Lee, MD, PhD, Mindy S. Lo, MD, PhD, Esra Meidan, MD, Jordan E. Roberts, MD, MPH, Mary Beth F. Son, MD, Robert P. Sundel, MD, Fatma Dedeoğlu, MD, Alicia Casey, MD, Joyce Chang, MD, MSCE, Lauren A. Henderson, MD, MMSc: Boston Children's Hospital, Boston, Massachusetts; [2]Marina Ohashi, BSc, Catherine Gebhart, PhD: HLA-Lab, American Red Cross, Dedham, Massachusetts; [3]Siobhan Case, MD, MHS, Peter A. Nigrovic, MD: Boston Children's Hospital and Brigham & Women's Hospital, Boston, Massachusetts.

Drs. Chang and Henderson contributed equally to this work.

SIGNIFICANCE & INNOVATIONS

- Eosinophilia is common in juvenile idiopathic arthritis (JIA) patients during treatment with interleukin-1 (IL-1) and IL-6 inhibitors, including systemic JIA patients.
- Pretreatment macrophage activation syndrome may be a risk factor for subsequent eosinophilia during IL-1/IL-6 inhibitor therapy in systemic JIA.
- Features of severe disease were prominent in systemic JIA patients who later developed lung disease.
- There was a high proportion of systemic JIA patients in our cohort who were positive for HLA–DRB1*15 alleles, including children who did not develop eosinophilia or lung disease after biologic drug exposure.

[TNF] inhibitors) were poor, since erosive arthritis was common and often treatment refractory (4,5).

The cytokine profile of systemic JIA is distinct from other types of JIA and centers around interleukin-1 (IL-1), IL-18, and IL-6 (6–8). Supporting the clinical importance of these cytokines, first-line therapy with IL-1 or IL-6 targeted biologic disease-modifying antirheumatic drugs (bDMARDs) improves inflammation and prevents the development of erosive arthritis (9–12). Initial treatment with these bDMARDs is now the standard of care, as reflected in consensus treatment guidelines (13,14).

Concurrent with the shift toward first-line bDMARD therapy, there have been increased reports of pulmonary complications in systemic JIA, including pulmonary hypertension and interstitial lung disease with unique features such as endogenous lipoid pneumonia and pulmonary alveolar proteinosis (15–18). Systemic JIA–related lung disease can be fatal, provoking widespread concern and intense interest in understanding its pathogenesis. Epidemiologic risk factors in recent cohorts include a history of MAS, young age at diagnosis, high IL-18 levels, trisomy 21, and adverse drug reaction to bDMARDs (16,18). In addition, several studies have noted an association between bDMARD exposure and systemic JIA–lung disease, albeit confounded by disease severity (15,16,18).

Compared to patients with systemic JIA without lung disease, those with pulmonary complications are more likely to have received immunomodulatory therapies, particularly IL-1/IL-6 inhibitors (15,16). Saper et al observed that some patients who received treatment with IL-1/IL-6 inhibitors also develop eosinophilia and nonevanescent, pruritic rashes, particularly among those positive for the HLA haplotype DRB1*15:01-HLA–DRB5*01:01-DQA1*01:02-DQB1*06:02 (HLA–DRB1*15:01 used as proxy) (19). There is concern that these findings represent a delayed type hypersensitivity-like reaction (DTH) that develops in a subset of patients with systemic JIA and may increase the risk of lung disease. In some patients, these DTHs have been described as drug reaction with eosinophilia and systemic symptoms (DRESS) (19). An implication is that the same drugs that have dramatically improved outcomes in children with systemic JIA may be responsible for the potentially fatal complication of systemic JIA–lung disease.

There are still areas of uncertainty. Systemic JIA–lung disease occurs in patients not exposed to IL-1/IL-6 inhibitors (15,16). Systemic JIA–lung disease improves in some patients with the addition of bDMARD therapy, making it unclear if bDMARD therapy should be discontinued (18). Last, the rates of eosinophilia in children with systemic JIA, as well as in children with nonsystemic forms of JIA exposed to these biologic drugs, are unknown, making it unclear how to interpret and respond to eosinophilia in clinical practice (19).

To address these questions, we conducted a single-center retrospective cohort study to 1) compare the incidence of eosinophilia among new users of bDMARDs with systemic JIA compared to those with other forms of JIA, 2) identify risk factors for and outcomes of eosinophilia among children with systemic JIA, and 3) describe characteristics in children who developed systemic JIA–lung disease.

PATIENTS & METHODS

Study design. The Boston Children's Hospital Institutional Review Board granted an exemption and waiver of informed consent (approval no. IRB-P00040299) for retrospective data collection and secondary use of biosamples previously collected under an approved protocol, in which subjects provided written informed consent to use their biospecimens for future research (approval nos. IRB-P00005723 and IRB-07-09-0375).

Study population. Children were considered for inclusion if they received a billing diagnosis code for systemic JIA or any other category of JIA in the period from 1995 to 2022 (International Classification of Diseases, Ninth Revision, Clinical Modification [ICD-9-CM] codes: 714.0-3, 716.5, 719.00-09, 720.00-02, 720.8, and 720.9; ICD-10-CM codes: M08.0-9 and L40.54) were <18 years old, had a medication administration or prescription record for an IL-1/IL-6 inhibitor at our center, and had ≥2 rheumatology clinic visits. The charts of identified patients were manually reviewed to confirm a diagnosis of systemic JIA/JIA by the treating physician, and each case was separately adjudicated by a pediatric rheumatologist (HW) as most likely

Author disclosures are available online at https://onlinelibrary.wiley.com/doi/10.1002/acr.25129.

Address correspondence via email to Lauren A. Henderson, MD, MMSc, at lauren.henderson@childrens.harvard.edu.

Submitted for publication December 9, 2022; accepted in revised form April 4, 2023.

systemic JIA/JIA and not an alternative diagnosis regardless of whether the patient met International League of Associations for Rheumatology classification criteria. Medication exposures for which the eosinophilia outcome was not evaluable were excluded. Data were abstracted from medical records and entered into a REDCap database (20).

Study measures. *Medication exposures.* In patients who received an IL-1/IL-6 inhibitor, we considered incident exposures to each medication in the following classes 1) IL-1/IL-6 inhibitors, including anakinra, canakinumab, and tocilizumab; 2) TNF inhibitors or other bDMARDs; and 3) MTX monotherapy. Medication exposure periods were defined from the start date of each new medication (index date) until the outcome of interest, date of medication discontinuation with or without switch to a different cytokine inhibitor, or date of the last follow-up visit, whichever occurred first. Individuals could contribute to multiple exposure windows (including drugs within the same class) if each medication exposure was the first exposure to that medication (Figure 1).

Outcome measures. The primary outcome was eosinophilia, defined as 2 new consecutive absolute eosinophil counts (AECs) >500 cells/μl during medication exposure, or, if the preexposure AEC was already >500 cells/μl, a doubling of the preexposure AEC. Preexposure AEC was defined as the peak AEC in the 3 months prior to starting a new medication. "Time to eosinophilia" was based on the time from starting a medication to the first date contributing to fulfilling the above criteria. As a secondary outcome, we also evaluated the presence of rash that was not a typical evanescent systemic JIA rash at the time of meeting eosinophilia criteria. In patients who developed eosinophilia, changes to medication regimens within 10 weeks of fulfilling eosinophilia criteria were documented regardless of the motivation for the change. Subsequent changes in AEC and clinical status, such as flare (defined as an increase in joint count or levels of inflammatory markers) or MAS at the next clinic visit were recorded.

We described 2 additional secondary outcomes of systemic JIA–lung disease and systemic JIA–related death. Charts were reviewed for signs and symptoms of pulmonary disease (dyspnea, cough, hypoxia, and clubbing) and diagnostic testing (abnormal echocardiogram, computed tomography [CT] of the chest, or pulmonary function tests [PFTs]). Pulmonary evaluation for systemic JIA–lung disease was defined as completion of PFTs and/or CT of the chest. Patients with a history of respiratory symptoms or abnormal pulmonary evaluation were adjudicated for systemic JIA–lung disease and disease severity by a pediatric pulmonologist (AC). Systemic JIA–lung disease severity was defined as mild (no dyspnea or hypoxia and stable/improving serial CT scans of the chest) or severe–progressive (respiratory distress and worsening serial chest imaging).

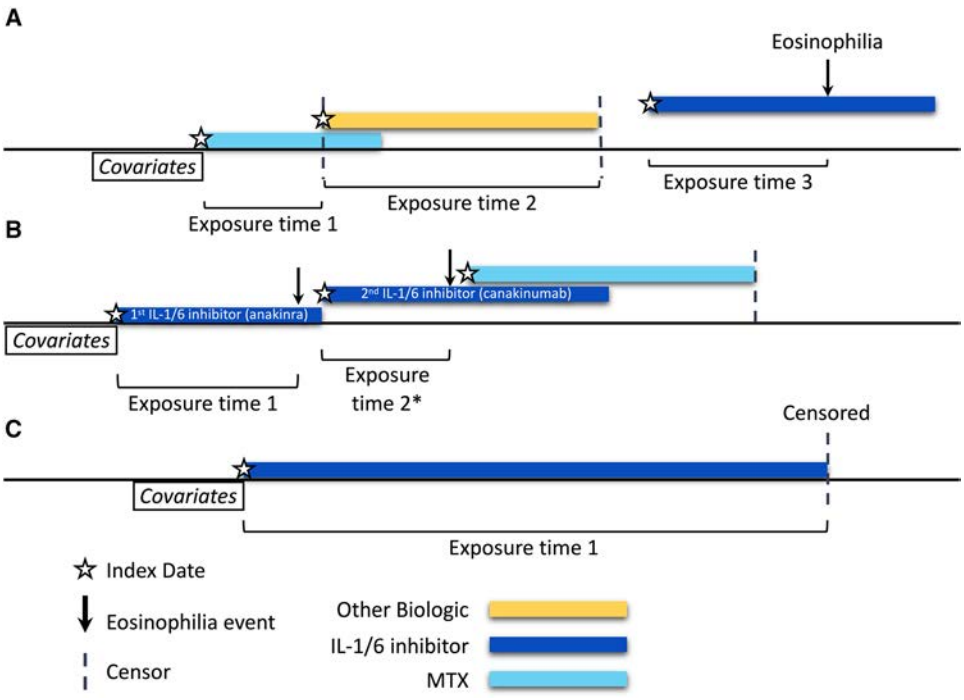

Figure 1. Line diagram illustrating exposure and outcome assessment periods. Both single event (eosinophilia) per subject (incidence) and multiple events per subject (Cox analysis) analyses were performed. **A**, Exposure 1 (time on methotrexate [MTX]) is censored when another biologic drug was added to capture only MTX monotherapy; exposure 2 is censored at therapy discontinuation; and exposure 3 is censored at time of outcome of interest (eosinophilia). **B**, Exposures 1 and 2 are censored at drug discontinuation and the time of outcome, respectively. **C**, Exposure 1 is censored at the time of drug discontinuation. * = only included in multiple events analysis.

Those with suspicious symptoms but no pulmonary evaluation were classified as having "probable systemic JIA–lung disease" and included in our final systemic JIA–lung disease count. The date of systemic JIA–lung disease diagnosis was defined as the first date a provider-documented systemic JIA–lung disease. Deaths were adjudicated as systemic JIA–related or systemic JIA–lung disease-related by 2 rheumatologists and a pulmonologist (HW, LAH, and AC, respectively).

Covariates. Demographic features included age at first physician diagnosis of JIA, disease duration at each index date, calendar year of diagnosis, biologic sex, as well as race and ethnicity as documented in the medical record by patient self-report obtained via open-ended response at registration. Baseline clinical features included history of atopy and AEC prior to the first exposure to biologic therapy or MTX. HLA typing was available in 13 patients on a clinical basis and 32 patients on a research basis through our institutional biorepository. For research-based HLA typing, DNA was extracted from either whole blood or peripheral blood mononuclear cells (Qiagen). High-resolution sequencing of 11 HLA loci was performed by the HLA Research Testing Services of the American Red Cross. To account for non-European ancestries, we assessed HLA–DRB1*15:XX instead of 15:01 alone. We also assessed several systemic JIA–related severity indicators, including pretreatment MAS (at initial presentation, prior to any biologic or MTX) per the 2016 classification criteria for MAS in systemic JIA (21), the number of systemic JIA–related hospitalizations and episodes of MAS in the window 2 months prior and 6 months after systemic JIA diagnosis, and systemic JIA–related intensive care unit (ICU) admissions occurring prior to systemic JIA–lung disease diagnosis.

Statistical analysis. Standard descriptive statistics, including Student's *t*-test or Wilcoxon's rank sum test for continuous variables, and chi-square test or Fisher's exact test for categorical variables were used to compare the characteristics of patients with systemic JIA to characteristics of individuals with other forms of JIA (other JIA) at baseline, and patients with systemic JIA at the time of eosinophilia compared to patients with systemic JIA without eosinophilia (at the median time to eosinophilia in the former patients).

We calculated incidence rates (IRs) for first occurrences of eosinophilia (single event per patient) using person-years of exposure time as the offset. Mid-p exact tests were used to compare incidence rates between patients with systemic JIA compared to those with other forms of JIA (other JIA) by medication class, and to compare incidence rates between medication classes. For IL-1/IL-6 inhibitor exposure times among patients with systemic JIA, we also considered the possibility of eosinophilia events occurring on multiple different medications in the same patient. Univariable and multivariable Cox proportional hazards regression models (multiple events per patient) adjusted for clustering within-subject were used to identify factors associated with time to eosinophilia. HLA type was included a priori in the multivariable model, and additional covariates were tested in the model using forward selection and retained if the *P* value was less than 0.25 or other coefficients changed by >15%. Schoenfeld residuals were calculated to ensure the proportional hazards assumption was satisfied for all models. All statistical tests were conducted using STATA version 16.0 at a significance level of 0.05 using 2-sided tests, unless otherwise specified.

RESULTS

Patient cohort. A total of 116 patients were included in this analysis, of whom 75 had systemic JIA and 41 had other forms of JIA (other JIA) (see Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25129/abstract). The median age at diagnosis in children with systemic JIA was 5 years (interquartile range [IQR] 2–9 years) compared to 7 years (IQR 3–11 years) in other JIA. A greater proportion of patients with other JIA were female. Both groups were predominantly of White race and non-Hispanic ethnicity. Eosinophilia was present in 20% of patients with systemic JIA patients prior to initiation of the first biologic medication (14 of 69) or MTX compared to 8% of patients with other JIA (3 of 37). Children with systemic JIA had a median 4.2 years (IQR 2.1–8.7 years) of follow-up.

A total of 60% of the patients with systemic JIA were HLA typed, of whom 49% were positive for HLA–DRB1*15:XX alleles (see Supplementary Tables 1 and 2, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25129/abstract). A total of 38% percent of the patients with HLA typing had the HLA–DRB1*15:01 allele. There were no significant differences in baseline characteristics between patients who were HLA typed compared to those without typing (Supplementary Table 1). HLA–DRB1*15 alleles were common in all racial and ethnic groups (Table 1 and Supplementary Table 1).

Incidence of eosinophilia in patients with systemic JIA and other JIA. There were 276 incident medication exposures, of which 264 were evaluable and included in the analysis (see Supplementary Figure 1, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25129/abstract). We identified 32 eosinophilia events across all medication classes, of which 27 events were first occurrences of eosinophilia in an individual patient (19 systemic JIA, 8 other JIA), and 5 events were recurrences of eosinophilia on a different medication (4 systemic JIA, 1 other JIA). Eosinophilia was observed during exposure to all medication classes evaluated. Considering only the first occurrence of eosinophilia per patient, the crude incidence rate of eosinophilia for any medication class was 0.09 per person-year of exposure in patients with systemic JIA (95% confidence interval [95% CI] 0.06–0.14) and 0.06 per person-year (95% CI 0.03–0.11) in patients with other JIA. There was a higher incidence of eosinophilia events on IL-1/IL-6 inhibitors compared to other medication classes (IR 0.10/person-year [95% CI

Table 1. Clinical characteristics of patients with JIA exposed to IL-1/IL-6 inhibitors*

Characteristics	Systemic JIA (n = 75)	Other JIA (n = 41)	*P*
Age at diagnosis, median (IQR) years	5 (2–9)	7 (3–11)	0.09
Female sex	44 (59)	33 (80)	0.02
Race			0.69
American Indian	0 (0)	1 (2)	
Asian	4 (5)	2 (5)	
Black	6 (8)	1 (2)	
Other	7 (9)	4 (10)	
Undisclosed	10 (13)	7 (17)	
White	48 (64)	26 (63)	
Hispanic ethnicity	9 (12)	7 (17)	0.45
Calendar year at diagnosis,			0.03
1995–2010	25 (33)	22 (54)	
2010–2022	50 (67)	19 (46)	
History of atopy†	28 (37)	9 (22)	0.09
Follow-up, median (IQR) years	4.2 (2.1–8.7)	9.5 (6.0–12.5)	<0.01
Preexposure AEC >500 cells/μl‡	14 (20)	3 (8)	0.16
Pretreatment MAS§	17 (23)	–	–
HLA–DRB1*15:XX status¶		–	–
Positive	22 (29)		
Negative	23 (31)		
Not HLA typed	30 (40)		

* Except where indicate otherwise, values are the number (%) of patients. Categorical variables were compared using Chi-square or Fisher's exact tests (for n < 5) and continuous variables were compared using Wilcoxon's rank sum tests.
† Reported food allergy, drug allergy, or asthma.
‡ Preexposure absolute eosinophil count (AEC) is prior to the first medication (biologic therapies or methotrexate [MTX]) exposure. Preexposure AEC was available for 69 patients with systemic juvenile idiopathic arthritis (JIA) and 37 patients with other JIA.
§ Pretreatment macrophage activation syndrome (MAS) is MAS at initial presentation of systemic JIA and prior to treatment with biologic therapies or MTX.
¶ A total of 45 of 75 patients with systemic JIA had HLA typing, of whom 49% were positive for DRB1*15:XX; self-reported racial/ethnic backgrounds were: White 14 of 31, Black 1 of 3, Asian 2 of 2, Hispanic 3 of 5, other/undisclosed 2 of 4.
Interleukin-1 = IL-1.

0.06–0.15] compared to 0.04/person-year [95% CI 0.02–0.10]; *P* = 0.04, 1-sided). However, the overall incidence of eosinophilia did not differ significantly between systemic JIA (IR 0.08/person-year [95% CI 0.05–0.13]) compared to other JIA (IR 0.07/person-year [95% CI 0.03–0.15]; *P* = 0.30) (Table 2).

Factors associated with eosinophilia during IL-1/IL-6 inhibitor exposure in systemic JIA. On univariable analysis, a history of pretreatment MAS and shorter disease duration at the start date of the new medication were both significantly associated with an ∼3-fold higher rate of eosinophilia during IL-1/IL-6 inhibitor exposure in patients with systemic JIA (Table 3). Among those who were HLA typed, the presence of HLA–DRB1*15:XX was similarly associated with a higher rate of eosinophilia, although this was not statistically significant. Only these 3 features satisfied criteria for selection in the final multivariable model. While no single feature indicated as being a statistically significant independent predictor of eosinophilia, the magnitudes of their effect sizes were similar (Table 3), and there was evidence of a correlation between HLA–DRB1*15:XX positivity and pretreatment MAS (pretreatment MAS occurred in 73% of patients with the allele compared to 27%

Table 2. Incidence rates of eosinophilia by JIA type during exposure to IL-1/IL-6 inhibitors and other medication classes*

	Systemic JIA				Other JIA				
Medication	Person-years	Events	Incidence rate	95% CI	Person-years	Events	Incidence rate	95% CI	*P*†
IL-1/IL-6 inhibitor	213	18	0.08	0.05–0.13	92	6	0.07	0.03–0.15	0.30
Other bDMARD	47	1	0.02	0.00–0.15	103	3	0.03	0.01–0.09	0.42
MTX alone	13	1	0.08	0.01–0.56	28	0	0.00	–	0.15

* 95% CI = 95% confidence interval; bDMARD = biologic disease-modifying antirheumatic drug; MTX = methotrexate. See Table 1 for other definitions.
† Midpoint *P* values comparing incidence rates between systemic JIA versus non–systemic JIA for each medication class.

Table 3. Predictors of eosinophilia in patients with systemic JIA who were exposed to IL-1/IL-6 inhibitors*

	Unadjusted			Adjusted		
	HR	95% CI	*P*	HR	95% CI	*P*
Age <2 years at diagnosis	1.53	0.56–4.21	0.41	–	–	–
Female sex	0.68	0.27–1.74	0.43	–	–	–
Calendar year at diagnosis >2010	2.60	0.86–7.90	0.09	–	–	–
Disease duration <1 month†	2.62	1.11–6.22	0.03	1.94	0.70–5.41	0.21
Pretreatment MAS‡	3.16	1.20–8.29	0.02	2.18	0.71–6.76	0.18
History of atopy§	1.20	0.47–3.08	0.70	–	–	–
Preexposure AEC >500 cells/μl¶	1.71	0.79–3.73	0.17	–	–	–
HLA type						
DRB1*15:XX negative	–	Ref.	–	–	–	–
DRB1*15:XX positive	2.91	0.87–9.67	0.08	2.26	0.73–7.02	0.16
DRB1*15:XX unknown	1.74	0.52–5.80	0.37	1.52	0.43–5.36	0.52
Prednisone dose at index date						
None	–	Ref.	–	–	–	–
<1mg/kg/day	0.90	0.34–2.40	0.83	–	–	–
≥1mg/kg/day	1.05	0.36–3.03	0.93	–	–	–
Concurrent non-bDMARD#	0.47	0.18–1.26	0.13	–	–	–

* 95% CI = 95% confidence interval; bDMARD = biologic disease-modifying antirheumatic drug; HR = hazard ratio; ref. = reference.
† At index date.
‡ MAS at initial presentation of systemic JIA and prior to treatment with biologic therapies or methotrexate.
§ Reported food allergy, drug allergy, or asthma.
¶ Preexposure AEC is prior to the first exposure to biologic therapies or methotrexate.
# Time-varying.

of patients without). Of note, 7 of 22 HLA–DRB1*15:XX–positive patients (32%) developed eosinophilia during IL-1/IL-6 inhibitor exposure compared to 4 of 23 patients negative for the allele (17%) ($P = 0.31$).

Clinical characteristics at the time of eosinophilia during IL-1/IL-6 inhibitor exposure in systemic JIA. A total of 21 eosinophilia events during 115 IL-1/IL-6 inhibitor exposures occurred in 18 patients with systemic JIA (24% of the systemic JIA cohort). The median time from each index date to eosinophilia was 37 days (IQR 16–171 days; range 2–609 days), and the median peak AEC at any time during IL-1/IL-6 inhibitor exposure was 1,270 cells/μl (IQR 840–2,770 cells/μl). Eosinophilia events were associated with a higher rate of atypical rash (7 of 21 [33%]) than IL-1/IL-6 inhibitor exposures without eosinophilia (5 of 94 [5%]; $P < 0.01$). Eosinophilia events were also more frequently associated with a history of pretreatment MAS (43% compared to 21%; $P = 0.04$). There was no MAS at the time of fulfillment of eosinophilia criteria, and median ferritin and aspartate aminotransferase levels at the onset eosinophilia were in the normal range (see Supplementary Table 3, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25129/abstract). A total of 6 of 7 patients with both eosinophilia and atypical rash were HLA typed and all 6 were positive for the HLA–DRB1*15:XX allele (compared to 1 of 5 typed patients with eosinophilia without rash).

Response to eosinophilia on IL-1/IL-6 inhibitors in systemic JIA. In 8 of 21 instances, no medication change was made within 10 weeks of the eosinophilia event, in 7 of 21 instances a non-bDMARD was added, in 5 of 21 instances patients were switched to a different IL-1/IL-6 inhibitor, and in one instance the medication was discontinued without adding or switching. As shown in Figure 2, subsequent disease flare or MAS was seen in 4 of 5 patients (80%) whose medications were switched in this period. In contrast, the disease status of those who remained on their established bDMARD with or without addition of a new non-bDMARD remained stable, and eosinophilia resolved by the subsequent clinic visit in 10 of 15 instances (67%).

Incidence and associated factors for systemic JIA–lung disease. In the systemic JIA cohort, 25 of 75 children (33%) underwent a pulmonary evaluation, of whom 7 (9%) were confirmed to have systemic JIA–lung disease, which is comparable to prior reports (18). A total of 2 patients with systemic JIA–lung disease were reported in a prior cohort (16). One additional patient had probable systemic JIA–lung disease based on respiratory symptoms. The median time from systemic JIA diagnosis to systemic JIA–lung disease diagnosis was 1.0 year (IQR 0.7–3.9 years, range 0.4–6.2 years).

Of the children with confirmed or probable systemic JIA–lung disease, 3 of 8 had mild disease, and 3 of 8 had severe progressive disease. One patient remained asymptomatic but did not have serial imaging to determine progression. The final patient with suspected systemic JIA–lung disease was lost to follow up. All-cause mortality was high in patients with systemic JIA–lung disease (3 of 8—all severe-progressive [38%]), although only 1 of these deaths was attributable to systemic JIA–lung disease (see Supplementary Table 4, available at http://onlinelibrary.wiley.com/doi/10.1002/

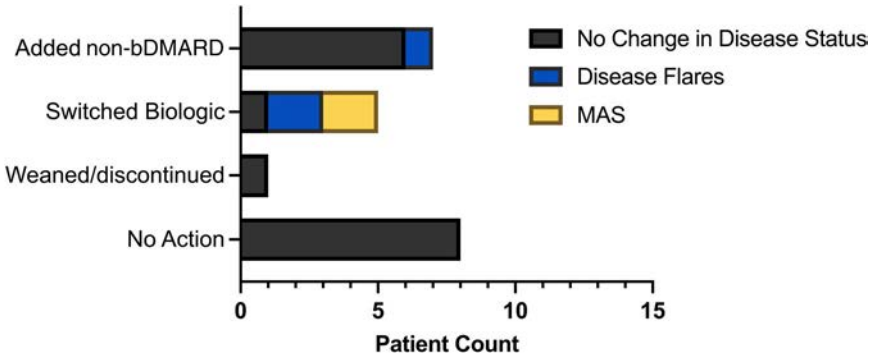

Figure 2. Disease activity following changes in therapy in patients with systemic juvenile idiopathic arthritis (JIA) who have eosinophilia. The bar graph reports clinical outcomes in systemic JIA patients with medications changes that were made within 10 weeks of eosinophilia. Eosinophilia was defined as 2 consecutive absolute eosinophil counts (AECs) >500 cells/μl or doubling of preexposure AEC. While we recorded switching to any biologic drug, all patients switched to another interleukin-1 (IL-1) or IL-6 inhibitor. MAS = macrophage activation syndrome; non-bDMARD = non-biologic disease-modifying antirheumatic drug. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25129/abstract.

acr.25129/abstract). Death directly attributed to lung disease in the entire systemic JIA cohort was 1.3% (1 of 75).

Patients with systemic JIA–lung disease frequently had a history of eosinophilia events with any medication exposure (6 of 8 patients) as well as preexposure eosinophilia prior to first biologic therapy or MTX (4 of 8 patients). Eosinophilia events were also more common in patients evaluated for systemic JIA–lung disease (13 of 25 patients, including those with a normal evaluation) compared to those who were not evaluated (Figure 3). In addition, age <24 months at systemic JIA diagnosis, positivity for HLA–DRB1*15:XX, atypical rash with eosinophilia, ≥1 episodes of MAS within the first 6 months of disease, and history of clubbing were all more common but not exclusive to patients who later developed systemic JIA–lung disease. Of the 22 patients exposed to IL-1/IL-6 inhibitors who were positive for HLA–DRB1*15:XX, 16 (73%) did not have confirmed or suspected systemic JIA–lung disease during a median follow-up duration of 3.4 years (IQR 0.9–8.5 years). The single feature most common in those with systemic JIA–lung disease (7 of 8 patients) and least common in those with a negative systemic JIA–lung disease evaluation (2 of 17 patients) was a systemic JIA–related ICU stay prior to lung disease diagnosis. One patient developed systemic JIA–lung disease prior to exposure to an IL-1/IL-6 inhibitor.

DISCUSSION

Our findings have implications for monitoring children with systemic JIA in the context of recent concerns regarding DTH and systemic JIA–lung disease. First, eosinophilia is common during exposure to IL-1/IL-6 inhibitors in both children with systemic JIA and those with other forms of JIA. Eosinophilia may be isolated in these cases without features of DTH. Second, higher disease severity may be associated with a higher incidence of eosinophilia as well as a higher likelihood of developing systemic JIA–lung disease. Lastly, there was a high prevalence of the HLA–DRB1*15 alleles in our systemic JIA cohort, even among patients who received treatment with IL-1/IL-6 inhibitors who did not develop eosinophilia or systemic JIA–lung disease over extended follow-up.

The incidence of eosinophilia in our cohort was higher in patients who received treatment with IL-1/IL-6 inhibitors than with other immunomodulators used in JIA. However, there was no significant difference in eosinophilia incidence among IL-1/IL-6 inhibitor–exposed patients with systemic JIA compared to other forms of JIA in which DTH and lung disease are infrequently reported. This finding suggests that eosinophilia associated with IL-1/IL-6 inhibition is not unique to systemic JIA physiology. Indeed, increases in eosinophil count on IL-1/IL-6 inhibitors have also been reported in rheumatoid arthritis, Kawasaki disease, and even myocardial infarction (19,22,23). A total of 20% of patients with systemic JIA had eosinophilia prior to exposure to biologic drugs or MTX. Nearly 25% of IL-1/IL-6 inhibitor–exposed patients with systemic JIA in our cohort developed eosinophilia. Overall, the high frequency of preexposure eosinophilia in patients with systemic JIA and eosinophilia events during IL-1/IL-6 inhibitor use in all subtypes of JIA, indicates that eosinophilia is relatively common in this population of patients requiring treatment with bDMARDs.

Pretreatment MAS was associated with a 3-fold increased rate of eosinophilia on univariable analysis, and MAS before initiation of therapy was also more common in those with HLA–DRB1*15:XX. MAS before drug exposure did not remain a significant independent predictor of eosinophilia in our multivariable model; however, we may not have had sufficient power to detect this association. While pretreatment MAS may be a risk factor for eosinophilia, laboratory markers of systemic inflammation were typically normal by the time eosinophilia events were observed on drugs. There were no episodes of MAS at or within 10 weeks of incident eosinophilia events in patients who remained on IL-1/IL-6 inhibitors. In addition, eosinophilia in patients with systemic JIA on

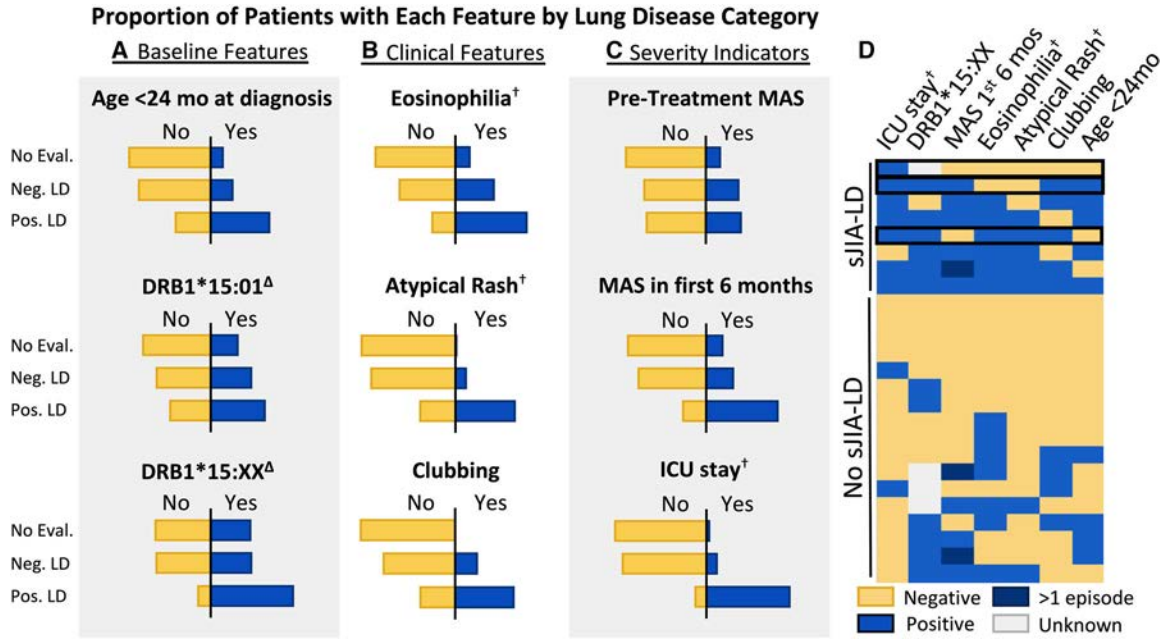

Figure 3. Clinical and disease features in patients with systemic JIA (sJIA) patients with and those without lung disease (LD). **A–C**, The proportion of patients with systemic JIA with the given baseline features prior to medication exposure (**A**), clinical features during treatment (**B**), and severity indicators (**C**) are shown. The bar graphs depict 3 different patient populations. The top bar shows patients with systemic JIA who did not have a pulmonary evaluation (no evaluation; n = 50), the middle bar represents patients who underwent pulmonary evaluation and did not have lung disease (negative lung disease; n = 17), and the bottom bar depicts patients with systemic JIA who were adjudicated as having definite or probable lung disease (positive lung disease; n = 8). Eosinophilia refers to eosinophilia events on medication. Atypical rash refers to the presence of this feature at the time of eosinophilia diagnosis. Clubbing refers to whether this was ever documented in the medical history of the patient. Pretreatment MAS refers to MAS at initial disease presentation, prior to treatment with biologic drugs or methotrexate. **D**) Heatmap depicting clinical features in individual patients with systemic JIA who had lung disease and patients with systemic JIA who underwent a pulmonary evaluation and did not have lung disease. Black outlines indicate individual patients who died (only 1 of whom died of systemic JIA–lung disease). See Figure 2 for other definitions. Δ = Sample size for HLA-typed cohort: n = 24 for no pulmonary evaluation (eval.), n = 14 for negative lung disease, and n = 7 for positive lung disease. † = Prior to diagnosis of systemic JIA–lung disease. ICU = intensive care unit; mos = months.

IL-1/IL-6 inhibitors was most commonly an isolated finding. Approximately 33% of eosinophilia events were associated with an atypical rash, but elevated liver function tests were uncommon.

Prior reports have suggested that subsequent MAS was more common in IL-1/IL-6 inhibitor–exposed patients with systemic JIA with eosinophilia that met the Registry of Severe Cutaneous Adverse Reaction classification criteria DRESS than in "drug-tolerant" controls, defined as those with stable disease on biologic therapy and low doses of glucocorticoids for >1 year (19). Our results are not directly comparable due to differences in cohort selection, study design, and outcome definitions. We evaluated eosinophilia with or without atypical rash in all drug-exposed systemic JIA patients regardless of underlying disease activity and did not specifically classify DRESS. In addition, we chose an incident user cohort design to avoid potential bias from depletion of susceptibles (24). In our cohort, changing therapy after eosinophilia was followed by disease flare or MAS in 4 of 5 patients. As most patients did not have systemic inflammation or MAS during eosinophilia events, routine switching for isolated eosinophilia may not be without potential harm. Unmeasured underlying disease severity could also have contributed to this observation. Investigation into the relationship between eosinophilia and disease trajectory in systemic JIA is warranted.

Severe disease features were common in children with systemic JIA–lung disease. A greater proportion of patients with systemic JIA–lung disease also had eosinophilia prior to initiation of biologic therapies or MTX (50%) compared to the entire systemic JIA cohort (20%). As with systemic JIA, patients without lung disease who developed eosinophilia, occurrence of MAS early in disease course was frequent in children with systemic JIA–lung disease. While the small number of patients with systemic JIA–lung disease precluded statistical comparisons with patients with systemic JIA without lung disease, it was notable that 7 of 8 patients had a systemic JIA–related ICU stay preceding lung disease diagnosis. These results suggest that systemic JIA disease severity may be associated with both eosinophilia and lung disease in patients receiving treatment with IL-1/IL-6 inhibitors. Our findings are consistent with prior case–control studies that identified disease activity and MAS as potential risk factors for systemic JIA–lung disease as well as a mouse model that demonstrated recurrent MAS flares resulted in increased lung inflammation and alterations in alveolar macrophages (15,17,18). Further

work is needed to understand the role of disease activity and MAS in driving both eosinophilia and systemic JIA–lung disease.

The clinical course of lung disease varied substantially, with some patients (3 of 8 patients) developing progressive disease while others (4 of 8 patients) had mild symptoms (1 patient lost to follow-up). There were patients with systemic JIA–lung disease who did not have prior eosinophilia (2 of 8 patients), and 1 patient developed pulmonary involvement prior to treatment with IL-1/IL-6 inhibitors. The resulting overall mortality rate in our patients with systemic JIA–lung disease was 38% (3 of 8 patients), with systemic JIA–lung disease–related death at 12.5% (1 of 8 patients), which is substantially lower than prior estimates ranging from 58% to 68% (15,16). The difference in survival rates may indicate that the most severe pulmonary presentations were reported in the initial case series describing this phenomenon. With greater awareness of systemic JIA–lung disease, more frequent screening may identify children with milder pulmonary involvement.

Nearly half (49%) of children with systemic JIA who were typed in our cohort were positive for the HLA–DRB1*15:XX allele, including 42% of those type on a research basis and not for clinical suspicion of lung disease. The majority of patients with HLA–DRB1*15:XX did not develop eosinophilic events or clinical systemic JIA–lung disease following IL-1/IL-6 inhibitor exposure. HLA–DRB1*15:01 has been linked with DRESS-like reactions in patients with systemic JIA who are of European ancestry, and similar features have been reported in patients with systemic JIA who are of non-European ancestry with other alleles in the DRB1*15 family (19). While we do not have access to information regarding ancestry, HLA–DRB1*15:XX positivity was observed across all self-reported racial and ethnic groups in our study. While the true proportion of patients with HLA–DRB1*15 alleles may be slightly lower in our total systemic JIA cohort, this HLA–DRB1*15:XX positivity rate exceeds the expected carriage rates based on allele frequencies of 16.5%, 15.2%, and 14.8% in healthy Asian, White, and Black individuals, respectively, in the National Marrow Donor Program database (25). The proportion of patients with HLA–DRB1*15 alleles in our cohort is also higher than what was reported in a large genome-wide association study (International Childhood Arthritis Genetics Consortium) in which 26% of individuals with systemic JIA who were of European ancestry carried these alleles (19,26).

Comparison to the INCHARGE data is somewhat limited because we were unable to determine ancestry in our study. Yet this degree of difference also raises the possibility that HLA–DRB1*15:XX could be a disease-associated allele, as previously reported in Japanese and Korean cohorts of adult-onset Still's disease (27–29), or potentially associated with severe disease features such as MAS and systemic JIA–lung disease. It is also possible that the requirement for serial rheumatology visits in our inclusion criteria may have enriched our systemic JIA population for patients with chronic, treatment-refractory disease over those with monophasic courses. The relationship between HLA–DRB1*15 alleles and features of systemic JIA disease severity need to be evaluated in future studies.

Last, although a relationship between HLA–DRB1*15-associated DTH and IL-1/IL-6 inhibitors has been posited, alternative hypotheses have been advanced to explain how these alleles may be linked to eosinophilia and atypical rash. In the cytokine plasticity hypothesis, it is proposed that IL-1/IL-6 blockade coupled with the cytokine milieu and environmental exposures may polarize T cells in patients with systemic JIA to a pathogenic phenotype (30). We cannot exclude that a small subset of patients in our cohort developed drug reactions to IL-1/IL-6 inhibitors. However, the relatively high frequency of eosinophilia in our systemic JIA cohort, as well as the many children with HLA–DRB1*15:XX who continued treatment with IL-1/IL-6 inhibitors without eosinophilic events or lung disease, suggests that the presence of this haplotype may not necessarily preclude use of these drugs in this patient population. Our data also suggest that disease activity and MAS may play a role in predisposing patients to both eosinophilia and lung disease, an association that needs to be studied in larger, multicenter cohorts.

Strengths of our study include the long follow-up period, use of disease control, and the incident user design. By including an unselected population of IL-1/IL-6 inhibitor–exposed patients with systemic JIA, we could estimate incidence rates, which is not possible with case-control designs. Using an established biorepository enabled us to HLA type most of our cohort. Our study also has several limitations. As a retrospective single-center study, our results may have limited generalizability to other centers. We studied IL-1 and IL-6 inhibitors together since both medication classes have been associated with eosinophilia in systemic JIA; however, fewer patients with other JIA received IL-1 inhibitors. Only 33% of patients underwent a pulmonary evaluation for systemic JIA–lung disease, and therefore rates of subclinical lung disease may be underestimated. Pulmonary evaluation was nonrandom and performed more frequently in patients with features such as young age at disease onset, eosinophilia, clubbing, and positivity for HLA–DRB1*15 alleles. However, the long median follow-up period of 4 years, relative to the median time to systemic JIA–lung disease diagnosis of 1 year likely reduced misclassification of clinically significant systemic JIA–lung disease cases. We had limited power to evaluate individual risk factors for eosinophilia or address multicollinearity and did not evaluate trajectories of eosinophilia over time. Last, we were unable to incorporate contemporary metrics of disease activity like IL-18 and CXCL9 that were only obtained in recently evaluated patients.

Eosinophilia during treatment with IL-1/IL-6 inhibitors is common in patients with all types of JIA and usually occurs without other features of DTH or coinciding MAS. High systemic JIA disease severity, particularly MAS, may be associated with later development of both eosinophilia and systemic JIA–lung disease. Thus, the importance of maintaining disease control should be considered prior to changing therapy in response to isolated eosinophilia, since this may risk disease flare. HLA–DRB1*15

alleles were frequent in our cohort, and many children exposed to IL-1/IL-6 inhibitors with these alleles were without clinical symptoms of lung disease during follow-up, suggesting that HLA type alone is insufficient to determine the safety of IL-1/IL-6 inhibitor use in systemic JIA.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Wobma, Casey, Chang, Henderson.
Acquisition of data. Wobma, Powers, Case, Chandler, Chang, Cohen, Day-Lewis, Fishman, Halyabar, Hausmann, Hazen, Lee, Lo, Meidan, Roberts, Son, Sundel, Dedeoğlu, Nigrovic, Casey, Chang, Henderson.
Analysis and interpretation of data. Wobma, Arvila, Taylor, Lam, Ohashi, Gebhart, Casey, Chang, Henderson.

REFERENCES

1. Behrens EM, Beukelman T, Gallo L, et al. Evaluation of the presentation of systemic onset juvenile rheumatoid arthritis: data from the Pennsylvania Systemic Onset Juvenile Arthritis Registry (PASOJAR). J Rheumatol 2008;35:343–8.
2. Petty RE, Southwood TR, Manners P, et al. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. J Rheumatol 2004;31:390–2.
3. Minoia F, Davì S, Horne A, et al. Clinical features, treatment, and outcome of macrophage activation syndrome complicating systemic juvenile idiopathic arthritis: a multinational, multicenter study of 362 patients. Arthritis Rheumatol 2014;66:3160–9.
4. Lomater C, Gerloni V, Gattinara M, et al. Systemic onset juvenile idiopathic arthritis: a retrospective study of 80 consecutive patients followed for 10 years. J Rheumatol 2000;27:491–6.
5. Packham JC, Hall MA. Long-term follow-up of 246 adults with juvenile idiopathic arthritis: functional outcome. Rheumatology (Oxford) 2002;41:1428–35.
6. Pascual V, Allantaz F, Arce E, et al. Role of interleukin-1 (IL-1) in the pathogenesis of systemic onset juvenile idiopathic arthritis and clinical response to IL-1 blockade. J Exp Med 2005;201:1479–86.
7. Yasin S, Fall N, Brown RA, et al. IL-18 as a biomarker linking systemic juvenile idiopathic arthritis and macrophage activation syndrome. Rheumatology (Oxford) 2020;59:361–6.
8. De Benedetti F, Massa M, Robbioni P, et al. Correlation of serum interleukin-6 levels with joint involvement and thrombocytosis in systemic juvenile rheumatoid arthritis. Arthritis Rheum 1991;34:1158–63.
9. Ter Haar NM, van Dijkhuizen EH, Swart JF, et al. Treatment to target using recombinant interleukin-1 receptor antagonist as first-line monotherapy in new-onset systemic juvenile idiopathic arthritis: results from a five-year follow-up study. Arthritis Rheumatol 2019;71:1163–73.
10. Nigrovic PA, Mannion M, Prince FH, et al. Anakinra as first-line disease-modifying therapy in systemic juvenile idiopathic arthritis: report of forty-six patients from an international multicenter series. Arthritis Rheum 2011;63:545–55.
11. Yokota S, Imagawa T, Mori M, et al. Efficacy and safety of tocilizumab in patients with systemic-onset juvenile idiopathic arthritis: a randomised, double-blind, placebo-controlled, withdrawal phase III trial. Lancet 2008;371:998–1006.
12. Nigrovic PA. Review: is there a window of opportunity for treatment of systemic juvenile idiopathic arthritis [review]? Arthritis Rheumatol 2014;66:1405–13.
13. Onel KB, Horton DB, Lovell DJ, et al. 2021 American College of Rheumatology guideline for the treatment of juvenile idiopathic arthritis: therapeutic approaches for oligoarthritis, temporomandibular joint arthritis, and systemic juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2022;74:521–37.
14. Dewitt EM, Kimura Y, Beukelman T, et al. Consensus treatment plans for new-onset systemic juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2012;64:1001–10.
15. Kimura Y, Weiss JE, Haroldson KL, et al. Pulmonary hypertension and other potentially fatal pulmonary complications in systemic juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2013;65:745–52.
16. Saper VE, Chen G, Deutsch GH, et al. Emergent high fatality lung disease in systemic juvenile arthritis. Ann Rheum Dis 2019;78:1722–31.
17. Gao DK, Salomonis N, Henderlight M, et al. IFN-γ is essential for alveolar macrophage-driven pulmonary inflammation in macrophage activation syndrome. JCI Insight 2021;6:e147593.
18. Schulert GS, Yasin S, Carey B, et al. Systemic juvenile idiopathic arthritis–associated lung disease: characterization and risk factors. Arthritis Rheumatol 2019;71:1943–54.
19. Saper VE, Ombrello MJ, Tremoulet AH, et al. Severe delayed hypersensitivity reactions to IL-1 and IL-6 inhibitors link to common HLA-DRB1*15 alleles. Ann Rheum Dis 2022;81:406–15.
20. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019;95:103208.
21. Ravelli A, Minoia F, Davì S, et al. 2016 classification criteria for macrophage activation syndrome complicating systemic juvenile idiopathic arthritis: a European League Against Rheumatism/American College of Rheumatology/Paediatric Rheumatology international trials organisation collaborative initiative. Arthritis Rheumatol 2016;68:566–76.
22. Morrisroe K, Wong M. Drug-induced hypereosinophilia related to tocilizumab therapy for rheumatoid arthritis. Rheumatology (Oxford) 2015;54:2113–4.
23. Del Buono MG, Damonte JI, Trankle CR, et al. Effect of interleukin-1 blockade with anakinra on leukocyte count in patients with ST-segment elevation acute myocardial infarction. Sci Rep 2022;12:1254.
24. Moride Y, Abenhaim L. Evidence of the depletion of susceptibles effect in non-experimental pharmacoepidemiologic research. J Clin Epidemiol 1994;47:731–7.
25. Maiers M, Gragert L, Klitz W. High-resolution HLA alleles and haplotypes in the United States population. Hum Immunol 2007;68:779–88.
26. Ombrello MJ, Arthur VL, Remmers EF, et al. Genetic architecture distinguishes systemic juvenile idiopathic arthritis from other forms of juvenile idiopathic arthritis: clinical and therapeutic implications. Ann Rheum Dis 2017;76:906–13.
27. Affleck AG, Littlewood SM. Adult-onset Still's disease with atypical cutaneous features. J Eur Acad Dermatol Venereol 2005;19:360–63.
28. Joung CI, Lee HS, Lee SW, et al. Association between HLA-DRB1 and clinical features of adult onset Still's disease in Korea. Clin Exp Rheumatol 2003;21:489–92.
29. Asano T, Furukawa H, Sato S, et al. Effects of HLA-DRB1 alleles on susceptibility and clinical manifestations in Japanese patients with adult onset Still's disease. Arthritis Res Ther 2017;19:199.
30. Binstadt BA, Nigrovic PA. The conundrum of lung disease and drug hypersensitivity-like reactions in systemic juvenile idiopathic arthritis. Arthritis Rheumatol 2022;74:1122–31.

Arthritis Care & Research
Vol. 75, No. 10, October 2023, pp 2073–2081
DOI 10.1002/acr.25121

American College of Rheumatology
Empowering Rheumatology Professionals

Socioeconomic Factors Are Associated With Severity of Hospitalization in Pediatric Lupus: An Analysis of the 2016 Kids' Inpatient Database

William Daniel Soulsby, Erica Lawson, Megumi Okumura, and Matthew S. Pantell

Objective. Health disparities in adult lupus, including higher disease severity and activity among those in poverty, have been identified. Similar associations in pediatric lupus have not been clearly established. This study was undertaken to investigate the relationship of income level and other socioeconomic factors with length of stay (LOS) in the hospital and severe lupus features using the 2016 Kids' Inpatient Database (KID).

Methods. Lupus hospitalizations were identified in children ages 2–20 years in the 2016 KID using International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10) codes (M32). Univariate and multivariate negative binomial regression analyses were used to analyze the association of income level, race and ethnicity, and insurance status with LOS in the hospital. Univariate and multivariate logistic regression analyses were used to analyze the association of the same predictors with the presence of severe lupus features, defined using ICD-10 codes associated with lupus sequelae (e.g., lupus nephritis).

Results. A total of 3,367 unweighted (4,650 weighted) lupus hospitalizations were identified. Income level was found to be a statistically significant predictor of increased LOS in the hospital for those in the lowest income quartile (adjusted incidence rate ratio 1.12 [95% confidence interval (95% CI) 1.02–1.23]). Black race, "other" race, and public insurance were also associated with severe lupus features (adjusted odds ratio [OR_{adj}] 1.51 [95% CI 1.11–2.06]; OR_{adj} 1.61 [95% CI 1.01–2.55]; and OR_{adj} 1.51 [95% CI 1.17–2.55], respectively).

Conclusion. Using a nationally representative data set, income level was found to be a statistically significant predictor of LOS in the hospital among those with the lowest reported income, highlighting a potential target population for intervention. Additionally, Black race and public insurance were associated with severe lupus features.

INTRODUCTION

Systemic lupus erythematosus (SLE) is a chronic autoimmune disease characterized by episodic, multiorgan inflammation leading to rash, arthritis, renal disease, neuropsychiatric symptoms, and other symptoms. It primarily affects women of childbearing age with ∼20% of cases presenting in childhood or adolescence (1). Health disparities are prevalent across the spectrum of chronic disease, including lupus (2,3). Several adult studies, including the seminal Lupus in Minorities: Nature Versus Nurture trial, have demonstrated significant health disparities, such as higher disease severity and activity among Latinx and African American patients and significantly higher mortality rates among those living in poverty (4–10).

Similar trends have been identified in pediatric lupus, including increased death and length of stay (LOS) in the hospital in underrepresented populations, particularly Latinx individuals (11,12). However, other study cohorts have not demonstrated an effect of race and ethnicity on long-term outcomes in pediatric patients with SLE (13). While these findings of racial disparities are important, they do not specifically address the impact of socioeconomic factors that may be more easily modifiable, such as income level, social support structures, or access to care (e.g., insurance status). Understanding the contribution of modifiable risk factors to health disparities in pediatric lupus will be critical to improve outcomes.

The possibility of ameliorating health disparities in adult lupus was highlighted in a recent analysis of the Lupus

Dr. Pantell's work was supported by the NIH loan repayment program (grant 9L40-HD-106442-02A1), the Agency for Healthcare Research and Quality (grant K12-HS-026383), and the National Center for Advancing Translational Sciences (grant KL2-TR-001870).

William Daniel Soulsby, MD, Erica Lawson, MD, Megumi Okumura, MD, MAS, Matthew S. Pantell, MD, MS: University of California, San Francisco.

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.25121&file=acr25121-sup-0001-Disclosureform.pdf.

Address correspondence via email to William Daniel Soulsby, MD, at daniel.soulsby@ucsf.edu.

Submitted for publication September 30, 2022; accepted in revised form March 23, 2023.

SIGNIFICANCE & INNOVATIONS

- Among a nationally representative data set of pediatric discharges across the US, low income was associated with a prolonged length of hospital stay in systemic lupus erythematosus (SLE).
- Racial disparities were also identified, including Black race and "other" race being associated with severe features of pediatric lupus, primarily driven by renal manifestations of the disease.
- The racial and ethnic breakdown of this nationally representative population of hospitalized pediatric patients with lupus was markedly different compared to the data reported for existing US-based patient registries of children with rheumatic diseases and included higher proportions of patients from underrepresented groups.
- Public insurance was also identified as a statistically significant predictor of severe features associated with pediatric SLE.

Outcomes Study. In this cohort, those living below 125% of the federal poverty limit demonstrated higher disease activity and accrued disease damage. However, those who were able to leave poverty during the follow-up period had similar disease activity scores compared to those who were never in poverty, suggesting potential reversibility (14). It is unclear if similar trends are identifiable among pediatric patients. If associations between income and lupus outcomes are found in pediatric patients, investigating societal programs, such as universal basic income (15,16), or more proximal programs, such as case management support (17,18), could be pursued to as a potential approach to decrease disease activity and damage.

In this cross-sectional analysis, our primary aim was to investigate associations of income level (as a proxy for socioeconomic status [19]) and other socioeconomic factors, including insurance status and race and ethnicity, with LOS in the hospital (as a proxy for disease severity and poor outcomes [20–23]) in pediatric lupus. We hypothesized that the lowest income levels were associated with longer LOS compared to less impoverished groups. Data were derived from the largest nationally representative inpatient pediatric sample available in the US, the 2016 Kids' Inpatient Database (KID).

PATIENTS AND METHODS

Study design and analytic sample. We cross-sectionally analyzed hospitalizations in children with SLE using data from the 2016 KID, Healthcare Cost and Utilization Project (HCUP), Agency for Healthcare Research and Quality (24). At the time of statistical analysis, this was the newest version of this large administrative data set. The KID is the largest all-payer and publicly available pediatric inpatient care database, sampled from >4,200 community hospitals (including children's hospitals) in the US. It comprises data from >3 million pediatric discharges in individuals <21 years of age and estimates a weighted >7 million annual pediatric discharges. It includes discharge-level data regarding inpatient discharges from participating hospitals and states sampling 10% of uncomplicated births and 80% of complicated births and all other pediatric discharges. Sampling weights, based on hospital characteristics, can be used to yield national estimates (24). Approval for exemption was obtained from the University of California, San Francisco Institutional Review Board.

Identification of study population. A flow diagram of the identification of the study population is shown (Figure 1). SLE hospitalizations were identified in children ages 2–20 years old with a primary or supplemental diagnosis of lupus (with 30 possible diagnoses listed per hospitalization in the KID) coded using an International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10) code of M32 (including subcodes, excluding M32.8, which may be used for neonatal lupus). Only those observations for nonelective admissions were included in the analysis (unweighted n = 3,417). Those with missing income, insurance, or race/ethnicity data were excluded from final analysis (total unweighted n = 3,136).

Primary predictor. Income level was defined as median annual income for a zip code divided into quartiles per data source. In the 2016 KID, this equates to $1–42,999 in quartile 1; $43,000–53,999 in quartile 2; $54,000–70,999 in quartile 3; and ≥$71,000 in quartile 4.

Primary outcome. LOS in the hospital was queried. The median (interquartile range [IQR]) and the mean ± SD by income quartile were calculated.

Study covariates. Race and ethnicity (defined as White, Black, Hispanic, Asian/Pacific Islander, or other/not available) and insurance status (public, private, or other) were separately analyzed as primary predictors in the statistical analysis. Study covariates included age, sex, geographic location (Northeast, South, Midwest, and West), and teaching status of hospital (rural, urban non-teaching, and urban teaching). Severity of illness (minor, moderate, major, or extreme loss of function) was controlled for using All Patient Refined Diagnosis Related Groups (APR DRG) severity of illness to adjust for potential confounding in multivariate modeling, since we sought to understand the impact of socioeconomic factors on LOS in the hospital, regardless of disease severity.

Measures for secondary outcomes. Our secondary aim was to investigate whether income level and other socioeconomic factors were associated with the presence of severe lupus features during hospitalization. Adult lupus studies using claims data

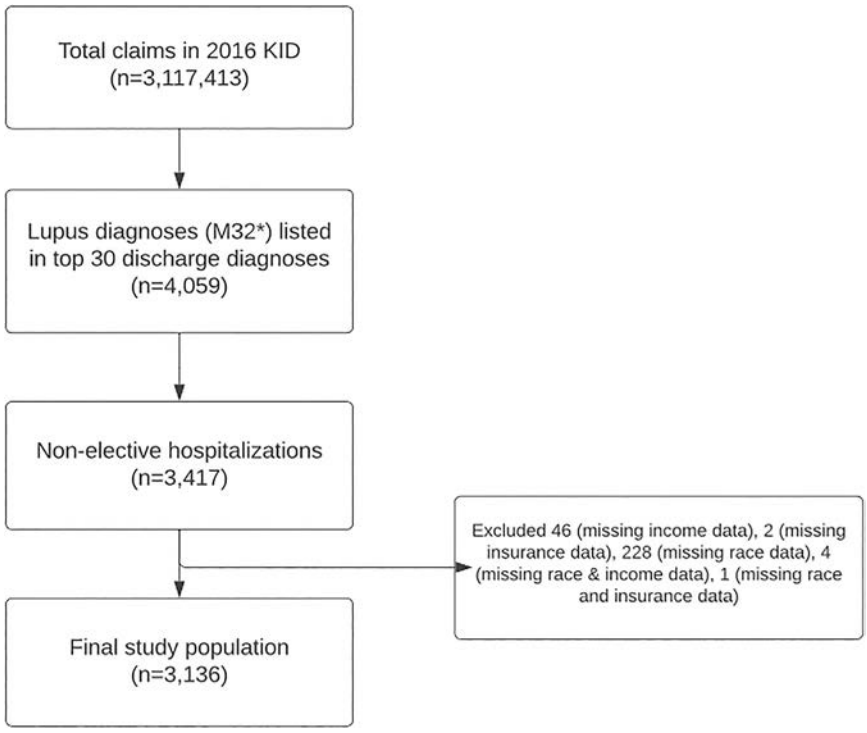

Figure 1. Flow diagram showing identification of the study population. KID = Kids' Inpatient Database.

utilized the Ward Index as a surrogate for disease severity; however, many of its parameters are not relevant to an exclusively pediatric population (25). Due to the lack of a validated pediatric measure to identify severe lupus, we created a dichotomous outcome measure indicating the presence of severe lupus features. ICD-10 codes were used (Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25121/abstract) to capture a breadth of severe lupus manifestations according to various organ systems, including renal, hematologic, neurologic, and cardiopulmonary manifestations, using diagnosis codes listed in prior study (26).

Statistical analysis. We compared baseline characteristics by income quartile using descriptive statistics. A chi-square test was used for categorical variables, while a Kruskal-Wallis test was used for continuous versus categorical predictors, given significant skew in the underlying data set. All data shown below are weighted unless otherwise specified, with analyses performed using the survey weights provided by HCUP (27).

Analysis of the distribution of LOS in the cohort was performed. Based on significant right skew and an outcome measure using count data (LOS), diagnostics were performed between negative binomial regression versus Poisson regression for choice of model for the primary outcome. Negative binomial regression was found to be most appropriate to accommodate longer LOS and was used to calculate incidence rate ratios (IRRs) to estimate the effect of income level on LOS unadjusted, then adjusted for sex and age, and finally adjusted for other covariates of interest (race and ethnicity, insurance status) and confounders (APR DRG severity of hospitalization, and location/teaching status of hospital).

Additionally, analogous univariate and multivariable logistic regression models were created for the primary aim of investigating the effect of income level on prolonged LOS in the hospital (defined as a median of ≥4 days) to facilitate the interpretability of our findings. Prior to model creation, collinearity was tested among all predictors using linear regression analysis and calculation of variance inflation factor—all found to be <10 (1.39). Using the final adjusted model, a marginal mean LOS was calculated at each income quartile. This method was repeated for calculation of unadjusted and adjusted IRR for insurance status and race and ethnicity separately as the primary predictor of LOS in the hospital. Interactions between income level and race were tested in the regression analyses for the primary and secondary outcomes. Since the interaction terms were either not statistically significant or did not qualitatively change the interpretation of the analysis, they were excluded from the final models. All analyses were performed using Stata version 16.1 (28).

RESULTS

Demographic characteristics. A total of 3,136 unweighted (4,326 weighted) lupus hospitalizations were

Table 1. Demographic and clinical characteristics of 3,136 unweighted (4,326 weighted) lupus hospitalizations in the 2016 Kids' Inpatient Database according to income level percentile*

Variable	$0–42,999 (n = 1,235)	$43,000–53,999 (n = 756)	$54,000–70,999 (n = 623)	≥$71,000 (n = 522)
Age at admission, years	18 (15–19)	18 (15–19)	18 (15–19)	18 (15–19)
Female sex	85.4 (82.1–88.1)	81.8 (77.4–85.4)	87.2 (83.3–90.3)	82.7 (78.6–86.1)
Race and ethnicity				
White	11.2 (9.1–13.8)	17.7 (14.4–21.5)	23.3 (19.0–28.3)	30.0 (24.5–36.2)
Black	53.0 (47.5–58.5)	32.8 (27.3–39.0)	31.3 (25.1–37.1)	28.6 (23.8–33.9)
Hispanic	27.6 (22.5–33.3)	38.2 (32.9–43.7)	30.4 (25.2–36.0)	16.8 (13.3–21.0)
Asian/Pacific Islander	3.0 (2.0–4.4)	4.8 (3.1–7.3)	8.5 (6.1–11.7)	16.6 (12.2–22.3)
Other/not available	5.3 (3.6–7.5)	6.6 (4.1–10.3)	6.5 (4.1–10.3)	8.0 (5.8–10.9)
Insurance				
Public	68.5 (64.6–72.1)	60.7 (55.4–65.8)	49.0 (44.1–53.9)	34.4 (28.8–40.5)
Private	24.1 (20.7–27.8)	29.5 (25.0–34.4)	40.0 (34.0–44.3)	57.1 (50.9–63.0)
Other (including self-pay and no charge)	7.4 (5.7–9.6)	9.8 (7.4–12.8)	12.0 (8.2–17.4)	8.6 (5.6–13.0)
Region of hospital				
Northeast	15.3 (11.1–20.7)	13.4 (9.8–17.9)	18.0 (13.3–23.9)	18.4 (13.0–25.4)
Midwest	18.6 (14.1–24.1)	15.4 (11.3–20.6)	14.6 (10.2–20.5)	14.1 (8.1–23.4)
South	50.2 (43.1–57.3)	41.9 (35.0–49.0)	40.4 (33.2–48.0)	31.3 (22.6–41.7)
West	15.9 (11.1–22.4)	29.4 (23.2–36.4)	27.1 (20.1–35.5)	36.2 (26.6–47.0)
Location/teaching status of hospital				
Rural	3.1 (2.0–4.7)	1.3 (0.01–2.4)	0.0 (0.0–0.01)	0.0 (0.00–0.00)
Urban non-teaching	10.1 (6.9–14.4)	10.8 (7.7–14.9)	14.5 (9.6–21.4)	11.4 (7.8–16.3)
Urban teaching	86.9 (82.5–90.3)	87.9 (83.7–91.1)	85.2 (78.3–90.2)	88.6 (83.7–92.2)
Length of stay				
Median (IQR)	4 (2–7)	4 (2–7)	4 (2–6)	4 (2–6)
Mean	7 (10)	6 (8)	6 (7)	6 (9)
Total charges, mean (IQR)	$32,921 ($17,564–66,171)	$36,055 ($18,628–74,353)	$31,298 ($17,817–60,962)	$34,290 ($19,603–74,114)
Severity of illness (APR DRG)				
Minor loss of function	6.1 (4.9–7.7)	6.3 (4.7–8.3)	7.5 (5.6–10.0)	6.1 (4.5–8.4)
Moderate loss of function	37.1 (33.7–40.7)	36.5 (32.9–40.3)	37.0 (33.0–41.3)	41.1 (36.7–45.7)
Major loss of function	41.6 (38.5–44.7)	44.2 (40.6–48.0)	45.0 (41.0–49.2)	40.6 (36.0–45.3)
Extreme loss of function	15.2 (13.0–17.7)	13.0 (10.1–16.5)	10.4 (8.2–13.1)	12.2 (9.2–16.0)
Any severe lupus features	42.3 (38.6–46.1)	42.9 (38.4–47.5)	38.8 (34.4–43.4)	39.0 (33.9–44.4)
Any severe renal lupus features	34.1 (30.4–37.9)	33.1 (28.6–37.9)	30.1 (25.9–34.6)	26.9 (22.2–32.1)
Any severe nonrenal lupus features	15.7 (13.6–18.1)	15.2 (12.4–18.5)	14.9 (12.0–18.4)	17.1 (13.5–21.4)

* Values are the percentage (95% confidence interval) for categorical variables and the median (interquartile range [IQR]) for continuous variables. APR DRG = All Patient Refined Diagnosis Related Groups.

identified in the 2016 KID data set. Table 1 summarizes the main demographic features by income quartile. The median age at admission was 18 years (IQR 15–19 years), and the subjects were predominantly female (82–87% across all income quartiles). Racial diversity was observed, with statistically higher representation of those from underrepresented populations in the lowest income quartile (53% Black and 28% Hispanic versus 11% White). Additionally, significantly higher proportions of discharges in the lowest income quartile reported public insurance utilization, compared to the highest income quartile (69% versus 34%). The majority of hospitalizations associated with the lowest income level were geographically from the South (50%). Hospitalizations were predominantly in urban teaching hospitals across all income quartiles (85–89%). Higher proportions of hospitalizations had severe lupus features in the lowest income quartile group compared to highest income quartile group (42% versus 39%), though these differences were not significantly different. Of severe lupus features, renal features were more common among those in lower income levels (e.g., 34% in the lowest quartile versus 27% in the highest) though the proportions of nonrenal features were roughly equivalent among all groups (15–17%).

Primary aim. Income level was found to be a statistically significant predictor of increased LOS in the hospital for those in the lowest income quartile (making between $1–42,999 per year) compared to those in the highest income quartile (making ≥$71,000 per year) in adjusted analysis (Table 2). In the adjusted analysis, those in the lowest income quartile had an IRR of 1.12, meaning that they had a 12% increase in LOS compared to individuals in the highest income quartile (95% confidence interval [95% CI] 1.02–1.23). Using the final adjusted negative binomial regression model, marginal adjusted mean LOS according to income quartile was calculated, with the longest LOS found in the lowest income quartile (Figure 2).

Income level was also found to be a statistically significant predictor of prolonged LOS in the hospital (≥4 days) in the

Table 2. Negative binomial regression model to investigate the effect of income level, race and ethnicity, and insurance status on hospital length of stay among children nonelectively admitted (n = 3,136 unweighted; n = 4,326 weighted) with a diagnosis of systemic lupus erythematosus in the 2016 Kids' Inpatient Database*

	Unadjusted			Age-/sex-adjusted model			Final adjusted model†		
	IRR	95% CI	*P*	IRR	95% CI	*P*	IRR	95% CI	*P*
Income level percentile									
$1–42,999	1.09	0.92–1.29	0.34	1.09	0.92–1.29	0.33	1.12‡	1.02–1.23‡	0.02‡
$43,000–53,999	1.01	0.85–1.20	0.90	1.01	0.85–1.20	0.92	1.08	0.97–1.20	0.18
$54,000–70,999	0.94	0.79–1.11	0.46	0.94	0.80–1.12	0.50	1.02	0.92–1.14	0.65
≥$71,000	–	–	–	–	–	–	–	–	–
Race and ethnicity									
White	–	–	–	–	–	–	–	–	–
Black	1.05	0.88–1.24	0.60	1.05	0.89–1.25	0.54	1.01	0.90–1.13	0.91
Hispanic	1.08	0.92–1.26	0.37	1.06	0.91–1.24	0.46	0.99	0.87–1.12	0.81
Asian/Pacific Islander	1.28‡	1.02–1.59‡	0.03‡	1.25	1.00–1.55	0.05	1.16	0.98–1.37	0.08
Other	1.04	0.82–1.30	0.77	1.02	0.81–1.29	0.86	0.97	0.82–1.14	0.71
Insurance									
Public	0.97	0.86–1.10	0.62	0.95	0.85–1.07	0.44	0.92‡	0.85–0.99‡	0.03‡
Private	–	–	–	–	–	–	–	–	–
Other (including self-pay and no charge)	0.86	0.73–1.01	0.07	0.86	0.73–1.01	0.06	0.91	0.80–1.03	0.14

* 95% CI = 95% confidence interval; IRR = incidence rate ratio.
† Adjusted for age, sex, hospital region, location/teaching status of hospital, All Patient Refined Diagnosis Related Groups severity index, and insurance status/income level and/or race and ethnicity.
‡ Value was statistically significant.

2 lowest income quartiles in the adjusted analysis, with an odds ratio OR of 1.28 (95% CI 1.01–1.62) in the lowest income quartile and OR of 1.34 (95% CI 1.05–1.73) in the second lowest income quartile (Supplementary Table 2, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25121/abstract). Although Asian/Pacific Islander race was associated with an increase in LOS in the unadjusted analysis (IRR 1.28 [95% CI 1.02–1.59]), this was no longer statistically significant upon adjusting for age, sex, hospital region, location, and teaching status of hospital, severity of hospitalization (APR DRG), income level, or insurance status. Asian/Pacific Islander race was a statistically significant predictor of prolonged LOS in the hospital (≥4 days) with an OR of 1.94 (95% CI: 1.30–2.89) in the adjusted analysis (Supplementary Table 2). Public insurance was found to be associated with an 8% lower LOS in the hospital (IRR 0.92 [95% CI: 0.85–0.99]), which was only statistically significant in the adjusted analysis. A sensitivity analysis in only those with a primary diagnosis of lupus (to improve the specificity of a lupus diagnosis relating to the need for hospitalization) did not significantly change the findings of the primary outcome.

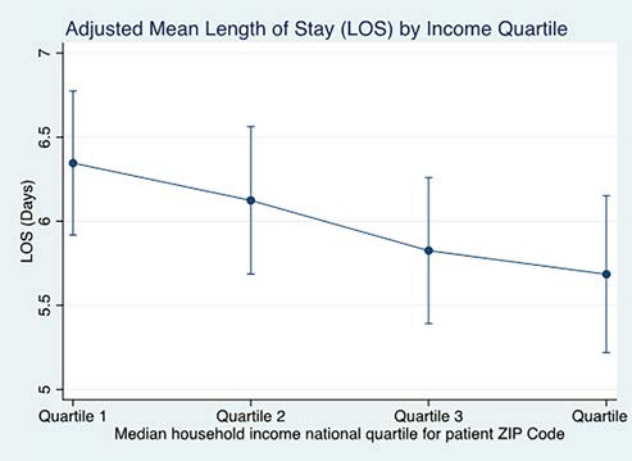

Figure 2. Multivariate negative binomial regression analysis was used to estimate the adjusted mean LOS according to income quartile in children nonelectively admitted to the hospital with a diagnosis of systemic lupus erythematosus in the 2016 Kids' Inpatient Database. The model was adjusted for age, sex, hospital region, location/teaching status of hospital, All Patient Refined Diagnosis Related Groups severity index, presence of severe lupus features, and insurance status/income level and/or race and ethnicity.

Secondary aim. Income level was not found to be a statistically significant predictor of severe lupus features in unadjusted or adjusted analyses (Table 3). Hispanic, Black, and other race were identified as statistically significant predictors of severe lupus features in unadjusted analyses, with Black and other race remaining statistically significant upon final adjustment for age, sex, and other confounders (Black race adjusted OR [OR_{adj}] 1.51 [95% CI 1.11–2.06]; other race OR_{adj} 1.61 [95% CI: 1.01–2.55]). Public insurance was also identified as a statistically significant predictor of severe lupus features in unadjusted and adjusted analyses. In the adjusted analysis, individuals who were publicly insured had a 51% increase in the odds of severe lupus features compared to those with private insurance (OR_{adj} 1.51 [95% CI 1.17–1.94]).

To understand if there were differences according to income level and other predictors between severe renal versus nonrenal

Table 3. Logistic regression model to investigate the effect of income level, race and ethnicity, and insurance status on presence of severe lupus features among children nonelectively admitted (n = 3,136 unweighted; n = 4,326 weighted) with a diagnosis of systemic lupus erythematosus in the 2016 Kids' Inpatient Database*

	Unadjusted			Age-/sex-adjusted			Final adjusted†		
	OR	95% CI	*P*	OR	95% CI	*P*	OR	95% CI	*P*
Income level percentile									
$1–42,999	1.15	0.87–1.51	0.32	1.14	0.87–1.50	0.33	0.82	0.58–1.14	0.24
$43,000–53,999	1.17	0.89–1.55	0.26	1.15	0.87–1.51	0.34	0.93	0.67–1.29	0.66
$54,000–70,999	0.99	0.76–1.30	0.95	0.99	0.76–1.30	0.96	0.90	0.65–1.24	0.51
≥$71,000	–	–	–	–	–	–	–	–	–
Race and ethnicity									
White	–	–	–	–	–	–	–	–	–
Black	1.59‡	1.20–2.10‡	0.001‡	1.57‡	1.19–2.07‡	0.002‡	1.51‡	1.11–2.06‡	0.01‡
Hispanic	1.39‡	1.08–1.79	0.01‡	1.42‡	1.11–1.83‡	0.01‡	1.26	0.91–1.75	0.16
Asian/Pacific Islander	1.33	0.92–1.93	0.13	1.41	0.98–2.03	0.07	1.13	0.74–1.73	0.56
Other/not available	1.56‡	1.03–2.36‡	0.04‡	1.66‡	1.09–2.53‡	0.02‡	1.61‡	1.01–2.55‡	0.045‡
Insurance									
Public	1.47‡	1.20–1.81‡	<0.001‡	1.55‡	1.26–1.91‡	<0.001‡	1.51‡	1.17–1.94‡	0.002‡
Private	–	–	–	–	–	–	–	–	–
Other (including self-pay and no charge)	1.18	0.87–1.60	0.29	1.14	0.83–1.57	0.42	1.19	0.80–1.76	0.40

* Severe lupus features were defined by International Statistical Classification of Diseases and Related Health Problems, Tenth Revision diagnosis codes per Supplementary Table 1 (available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25121/abstract). 95% CI = 95% confidence interval; OR = odds ratio.
† Adjusted for age, sex, hospital region, location/teaching status of hospital, All Patient Refined Diagnosis Related Groups severity index, and insurance status/income level and/or race and ethnicity.
‡ Value was statistically significant.

lupus features, a stratified analysis was performed using logistic regression models to investigate severe renal (identified in n = 1,001 observations) versus nonrenal (identified in n = 493 observations) features (Table 4). Similar trends were identified in the prediction of severe lupus features with Black race (OR_{adj} 1.89 [95% CI 1.33–2.68]) and public insurance (OR_{adj} 1.55 [95% CI 1.17–2.04]), identified as statistically significant predictors of severe renal lupus features. Neither income level nor insurance status were identified as statistically significant predictors of nonrenal severe lupus manifestations in unadjusted or adjusted analyses (Table 4B). While "other" race was associated with severe nonrenal features in the unadjusted analysis (OR 1.76 [95% CI 1.12–2.76]), it was no longer statistically significant upon adjusting for confounders.

DISCUSSION

Using a nationally representative data set of pediatric discharges from across the US, this analysis highlights racial and socioeconomic health disparities impacting pediatric lupus hospitalizations. Among pediatric hospitalizations including a diagnosis code for lupus, those with the lowest levels of income experienced prolonged LOS in the hospital compared to individuals in the highest income level. Additionally, Black and "other" race and public insurance were identified as statistically significant predictors of severe lupus features.

Income level represents a potentially modifiable risk factor for increased LOS in the hospital. Prolonged LOS in the hospital is an important health care metric that has been associated with illness severity and death (20,23) along with poor care coordination (21) in the setting of complex medical plans (22). While difficult to quantify, prolonged hospitalizations may also impact the social, emotional, and economic welfare of hospitalized children and their families, and as such, may disproportionately impact those who are the most socioeconomically disadvantaged (22). While the magnitude of the effect of income on LOS in the hospital in this study was not qualitatively large, this may be a worthwhile focus for future interventional studies targeting health disparities in pediatric lupus. Data from the Lupus Outcomes Study demonstrated potential reversibility of the impact of poverty on disease activity and damage among those who left poverty (14). While programs providing financial support, such as universal basic income, could then theoretically reverse the impact of socioeconomic status on pediatric lupus outcomes (29,30), this is unlikely to be implemented for the purposes of addressing health disparities in pediatric lupus.

Additionally, such programs may not address other factors associated with poverty that may contribute to severe disease and prolonged LOS, including limited access to outpatient care or medications, housing insecurity, or lack of transportation, among other factors. However, other interventions that may alleviate these consequences of poverty and resultant barriers to care could be applied in a pediatric lupus population. One such program is the Novel Interventions in Children's Healthcare program, which is a home- and community-based program that pairs patients and families with an interventionist aiming to ameliorate

Table 4. Regression model used to investigate the effect of income level, race and ethnicity, and insurance status on presence of severe renal lupus features or severe nonrenal lupus features among children nonelectively admitted (n = 3,136 unweighted; n = 4,326 weighted) with a diagnosis of systemic lupus erythematosus in the 2016 Kids' Inpatient Database for a subgroup analysis of the secondary outcome measure*

	Unadjusted			Age- and sex-adjusted			Final adjusted†		
Features	OR	95% CI	*P*	OR	95% CI	*P*	OR	95% CI	*P*
Severe renal lupus									
Income level percentile									
$1–42,999	1.41‡	1.04–1.91‡	0.03‡	1.40	1.03–1.91	0.03	0.99	0.68–1.44	0.96
$43,000–53,999	1.35	0.97–1.87	0.08	1.31	0.94–1.83	0.11	1.09	0.75–1.58	0.66
$54,000–70,999	1.17	0.87–1.58	0.31	1.17	0.87–1.58	0.30	1.09	0.78–1.54	0.61
≥$71,000	–	–	–	–	–	–	–	–	–
Race and ethnicity									
White	–	–	–	–	–	–	–	–	–
Black	2.01‡	1.48–2.73‡	<0.001‡	1.99‡	1.46–2.71‡	<0.001‡	1.89‡	1.33–2.68‡	<0.001‡
Hispanic	1.46‡	1.10–1.95‡	0.01‡	1.51‡	1.13–2.01‡	0.01‡	1.29	0.89–1.87	0.18
Asian/Pacific Islander	1.55	0.99–2.44	0.06	1.67‡	1.07–2.61‡	0.03‡	1.50	0.91–2.47	0.11
Other/not available	1.39	0.84–2.29	0.20	1.50	0.91–2.48	0.12	1.37	0.80–2.34	0.25
Insurance									
Public	1.57‡	1.26–1.97‡	<0.001‡	1.67‡	1.33–2.11‡	<0.001‡	1.55‡	1.17–2.04‡	0.002‡
Private	–	–	–	–	–	–	–	–	–
Other (including self-pay and no charge)	1.19	0.86–1.65	0.30	1.15	0.82–1.61	0.43	1.13	0.75–1.71	0.55
Severe nonrenal lupus									
Income level percentile									
$1–42,999			0.54	0.91			0.85		0.42
$43,000–53,999	0.90	0.66–1.24	0.44	0.87	0.66–1.24	0.54	0.81	0.58–1.25	0.31
$54,000–70,999	0.87	0.61–1.24	0.40	0.85	0.61–1.23	0.42	0.86	0.54–1.22	0.51
≥$71,000	0.85	0.58–1.24	–	–	0.58–1.25	0.42	–	0.56–1.33	–
Race and ethnicity	–	–			–	–		–	
White	–	–	–	–	–	–	–	–	–
Black	0.97	0.71–1.32	0.85	0.96	0.71–1.31	0.80	0.92	0.66–1.28	0.62
Hispanic	1.24	0.90–1.70	0.19	1.24	0.90–1.70	0.19	1.12	0.79–1.58	0.53
Asian/Pacific Islander	1.18	0.75–1.87	0.48	1.19	0.75–1.88	0.45	0.84	0.51–1.40	0.51
Other/not available	1.76‡	1.12–2.76‡	0.01‡	1.78‡	1.13–2.80‡	0.01‡	1.56	0.94–2.59	0.08
Insurance									
Public	1.02	0.81–1.29	0.89	1.03	0.82–1.30	0.79	0.98	0.77–1.26	0.90
Private	–	–	–	–	–	–	–	–	–
Other (including self-pay and no charge)	1.03	0.67–1.59	0.90	1.02	0.66–1.57	0.94	1.08	0.68–1.74	0.74

* Severe lupus features were defined using International Statistical Classification of Diseases and Related Health Problems, Tenth Revision diagnosis codes per Supplementary Table 1 (available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25121/abstract). 95% CI = 95% confidence interval; OR = odds ratio.
† Adjusted for age, sex, hospital region, location/teaching status of hospital, All Patient Refined Diagnosis Related Groups severity index, and insurance status/income level and/or race and ethnicity.
‡ Value was statistically significant.

barriers to care in children with complex, chronic illness (31). This program has been shown to improve care and lower health care costs in children with type I diabetes and chronic pain (32,33). This program could be implemented to lower disparities related to socioeconomic factors in pediatric lupus and should be one avenue of future intervention and research.

Our analysis also identified racial disparities among Black and other non-White youth, who had a statistically significant increased odds of severe lupus features compared to White youth, not previously identified in pediatric lupus literature. Additionally, Asian/Pacific Islander race was associated with prolonged LOS in the hospital in adjusted analysis. In analyzing severe renal versus nonrenal features separately, our analysis suggests that racial disparities may be primarily driven by renal manifestations of the disease, including chronic kidney disease (CKD) and renal failure. These findings build on an existing body of literature, including a recent analysis of the Childhood Arthritis and Rheumatology Research Alliance Registry demonstrating worse short-term renal outcomes and a 4-fold increase in the rate of CKD stages 3, 4, and 5 in Black patients compared to individuals of non-White race (34). The foundation of these disparities will be critical to understand to achieve racial equity for patients with lupus.

While prior studies have proposed genetic ancestry as the etiology between race and ethnicity and poor outcomes (35), including increased death and end-stage renal disease (ESRD) among Black and Asian/Pacific Islander patients with lupus, race and ethnicity alone cannot predict genetic ancestry due to significant genetic admixture within a single racial group (36). Rather, it must be recognized that race and ethnicity are social

constructs rather than biologic phenomena implicated in childhood health (36–38). More focus is needed on social underpinnings that may underlie these relationships rather than biologic factors as the justification for these findings.

Additionally, our population-based results highlight racial disparities among those with pediatric lupus who require hospitalization. Recent data demonstrate that existing US-based patient registries of children with rheumatic diseases include a population that is 70–90% White (39,40). However, in our study, those identifying as White only comprised between 11% and 30% of the population at each income level, while those identifying as Black comprised 29–53% of the population at each income level. Importantly, as we did not have individual-level data in our data set, we cannot exclude the possibility that recurrent hospitalizations in a smaller subset of patients may skew demographic data. Likewise, the true underlying prevalence of pediatric lupus according to racial group is unknown in the general population. Nevertheless, our data suggest that Black patients and other individuals of minority racial groups may be more likely to be admitted to the hospital for lupus. Further study of the underlying epidemiology of pediatric lupus and increasing efforts to recruit diverse patient populations in pediatric rheumatology research will be critical to future work in mitigating health disparities in pediatric lupus.

Finally, our study suggests worse outcomes among those with public insurance, an emerging disparity in the pediatric lupus literature. In a recent analysis of patients transitioning from pediatric to adult care, public insurance status was identified as a statistically significant predictor of ESRD and death (41). Reviewing our demographic data, individuals who make less money are more likely to report using public insurance compared to private (~69% in the lowest income quartile compared to ~34% in the highest). Even when controlling for income level, however, public insurance still yields a statistically significant increase in the odds of severe lupus features (OR_{adj} 1.51 [95% CI 1.17–1.94]), suggesting an impact of insurance type on lupus outcomes beyond the effect of socioeconomic status. Although the association identified was present upon controlling for a severity of illness indicator (APR DRG), there is the small possibility of confounding by disease indication given that individuals with more severe disease may qualify for supplemental public insurance programs compared to those with less severe disease. Whether there are differential outcomes by state or geographic region regarding public insurance programs is a key area of future study, since successes with public insurance programs in some states could be reasonably applied to others to mitigate this disparity.

Our study has several limitations. The KID provides discharge-level data, and therefore we do not know how many unique individuals represent the number of discharges within the data set. As the population is relatively small, a small group of patients may skew the demographic and clinical characteristics in our analysis. There may be misclassification error in diagnosis, since lupus may not necessarily be coded for an admission for a lupus-related illness, such as renal failure or infection from treatment. Many of our variables also serve as proxy measures for socioeconomic status and other social determinants of health. For example, income level was derived from the median income for zip codes and does not reflect individual income (nor can it address other components that impact wealth, such as debt). Additionally, our secondary outcome measure was not validated for the identification of severe lupus features and likely was not all-encompassing, though it does capture severe sequelae of lupus across many organ systems.

This analysis may support the idea that income level may not fully capture complex social factors encapsulating poverty, which also includes access to education, personal wealth, and neighborhood safety. Further studies using geocoded-derived indices of economic deprivation, such as the Area Deprivation Index (42), would be beneficial to study disparities among lupus patients and will be the subject of future work. Last, we did not include variables such as family support, family structure, or indices of continuity of care that may also influence LOS.

In conclusion, our analysis of a large cohort of pediatric SLE hospitalizations demonstrated that income level, as a proxy for poverty, is a statistically significant predictor of LOS in the hospital, with those in the lowest income levels reporting the longest LOS. Black and other non-White race and public insurance were also shown to be independent predictors of severe lupus features, suggesting the existence of racial and insurance-based health disparities in pediatric lupus outcomes. Future work is required to better understand the role of economic deprivation and poverty and its possible intersection with identified racial disparities in outcomes of pediatric rheumatic diseases such as lupus. Importantly, we can then focus equity research on generating interventions to ameliorate the worse health outcomes found in minority races and those living in poverty. Future research should also include dedicated analyses between LOS in the hospital and health care costs, as the identification of inefficient health care expenditure may incentivize interventions by hospital systems that could also mitigate health disparities in pediatric lupus. Though the median cost of hospitalization in our study yielded no statistically significant differences between groups, such analyses may identify disparities. Finally, the impact of income and socioeconomic factors on other metrics of hospitalization, including hospital readmission, should be the subject of future study.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Soulsby, Lawson, Okumura, Pantell.

Acquisition of data. Soulsby, Pantell.
Analysis and interpretation of data. Soulsby, Lawson, Okumura, Pantell.

REFERENCES

1. Harry O, Yasin S, Brunner H. Childhood-onset systemic lupus erythematosus: a review and update. J Pediatr 2018;196:22–30.e2.
2. Sundquist J, Johansson SE. The influence of socioeconomic status, ethnicity and lifestyle on body mass index in a longitudinal study. Int J Epidemiol 1998;27:57–63.
3. Demas KL, Costenbader KH. Disparities in lupus care and outcomes. Curr Opin Rheumatol 2009;21:102–9.
4. Chen E, Martin AD, Matthews KA. Understanding health disparities: the role of race and socioeconomic status in children's health. Am J Public Health 2006;96:702–8.
5. Li S, Gong T, Peng Y, et al. Prevalence and incidence of systemic lupus erythematosus and associated outcomes in the 2009–2016 US Medicare population. Lupus 2020;29:15–26.
6. Yelin E, Yazdany J, Trupin L. Relationship between poverty and mortality in systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2018;70:1101–6.
7. Crosslin KL, Wiginton KL. The impact of race and ethnicity on disease severity in systemic lupus erythematosus. Eth Dis 2009;19:301–7.
8. Sule S, Fivush B, Neu A, et al. Increased risk of death in African American patients with end-stage renal disease secondary to lupus. Clin Kidney J 2014;7:40–4.
9. Krishnan E. Hospitalization and mortality of patients with systemic lupus erythematosus. J Rheumatol 2006;33:1770–4.
10. Alarcón GS, Bastian HM, Beasley TM, et al. Systemic lupus erythematosus in a multi-ethnic cohort (LUMINA) XXXII: [corrected] contributions of admixture and socioeconomic status to renal involvement. Lupus 2006;15:26–31.
11. Knight AM, Weiss PF, Morales KH, et al. National trends in pediatric lupus hospitalizations in the United States: 2000–2009. J Rheumatol 2014;41:539–46.
12. Son MB, Johnson VM, Hersh AO, et al. Outcomes in hospitalized pediatric patients with systemic lupus erythematosus. Pediatrics 2014;133:e106–13.
13. Miettunen PM, Ortiz-Alvarez O, Petty RE, et al. Gender and ethnic origin have no effect on longterm outcome of childhood-onset systemic lupus erythematosus. J Rheumatol 2004;31:1650–4.
14. Yelin E, Trupin L, Yazdany J. A prospective study of the impact of current poverty, history of poverty, and exiting poverty on accumulation of disease damage in systemic lupus erythematosus. Arthritis Rheumatol 2017;69:1612–22.
15. Ruckert A, Huynh C, Labonté R. Reducing health inequities: is universal basic income the way forward? J Public Health (Oxf) 2018;40:3–7.
16. Painter A. A universal basic income: the answer to poverty, insecurity, and health inequality? BMJ 2016;355:i6473.
17. Bazzoli GJ, Carcaise-Edinboro P, Sabik LM, et al. Integrated case management: does it reduce health service disparities across African American and White Medicaid beneficiaries? Med Care Res Rev 2017;74:486–501.
18. Anandarajah A. Designing an intervention to improve management of high-risk lupus patients through care coordination. Rheum Dis Clin North Am 2020;46:723–34.
19. Thomas AJ, Eberly LE, Davey Smith G, et al. ZIP-code-based versus tract-based income measures as long-term risk-adjusted mortality predictors. Am J Epidemiol 2006;164:586–90.
20. Kyle UG, Genton L, Pichard C. Hospital length of stay and nutritional status. Curr Opin Clin Nutr Metab Care 2005;8:397–402.
21. Barisonzo R, Wiedermann W, Unterhuber M, et al. Length of stay as risk factor for inappropriate hospital days: interaction with patient age and co-morbidity. J Eval Clin Pract 2013;19:80–5.
22. Gibbs D, Ehwerhemuepha L, Moreno T, et al. Prolonged hospital length of stay in pediatric trauma: a model for targeted interventions. Pediatr Res 2021;90:464–71.
23. Horn SD, Sharkey PD, Buckle JM, et al. The relationship between severity of illness and hospital length of stay and mortality. Med Care 1991;29:305–17.
24. Agency for Healthcare Research and Quality. Healthcare Cost and Utilization Project (HCUP). HCUP Kids' Inpatient Database (KID). 2006, 2009. URL: www.hcup-us.ahrq.gov/kidoverview.jsp.
25. Ward MM. Development and testing of a systemic lupus-specific risk adjustment index for in-hospital mortality. J Rheumatol 2000;27:1408–13.
26. Shulman AI, Punaro M. Critical care of the pediatric patient with rheumatic disease. Curr Opin Pediatr 2011;23:263–8.
27. Agency for Healthcare Research and Quality. Healthcare Cost and Utilization Project (HCUP). HCUP KID trend supplemental file. 1997. URL: www.hcup-us.ahrq.gov/db/nation/kid/kidtrends.jsp.
28. StataCorp. Stata statistical software: release 16. 2019. College Station (TX): StataCorp LLC.
29. Duncan GJ, Magnuson K, Votruba-Drzal E. Boosting family income to promote child development. Future Child 2014;24:99–120.
30. Heinrich CJ. Parents' employment and children's wellbeing. Future Child 2014;24:121–46.
31. Harris MA, Spiro K, Heywood M, et al. Novel Interventions in Children's Health Care (NICH): innovative treatment for youth with complex medical conditions. Clin Pract Pediatr Psychol 2013;1:137–45.
32. Harris MA, Wagner DV, Wilson AC, et al. Novel interventions in children's healthcare for youth hospitalized for chronic pain. Clin Pract Pediatr Psychol 2015;3:48–58.
33. Harris MA, Wagner DV, Heywood M, et al. Youth repeatedly hospitalized for DKA: proof of concept for Novel Interventions in Children's Healthcare (NICH). Diabetes Care 2014;37:e125–6.
34. Vazzana KM, Daga A, Goilav B, et al. Principles of pediatric lupus nephritis in a prospective contemporary multi-center cohort. Lupus 2021;30:1660–70.
35. Lewis MJ, Jawad AS. The effect of ethnicity and genetic ancestry on the epidemiology, clinical features and outcome of systemic lupus erythematosus. Rheumatology (Oxford) 2017;56 Suppl 1:i67–77.
36. Mersha TB, Abebe T. Self-reported race/ethnicity in the age of genomic research: its potential impact on understanding health disparities. Hum Genomics 2015;9:1. Erratum in: Hum Genomics 2021;15:35.
37. Trent M, Dooley DG, Dougé J, et al. The impact of racism on child and adolescent health. Pediatrics 2019;144:e20191765.
38. Yearby R. Structural racism and health disparities: reconfiguring the social determinants of health framework to include the root cause. J Law Med Ethics 2020;48:518–26.
39. Ringold S, Beukelman T, Nigrovic PA, et al. Race, ethnicity, and disease outcomes in juvenile idiopathic arthritis: a cross-sectional analysis of the Childhood Arthritis and Rheumatology Research Alliance (CARRA) Registry. J Rheumatol 2013;40:936–42.
40. Beukelman T, Kimura Y, Ilowite NT, et al. The new Childhood Arthritis and Rheumatology Research Alliance (CARRA) registry: design, rationale, and characteristics of patients enrolled in the first 12 months. Pediatr Rheumatol Online J 2017;15:30.
41. Bitencourt N, Makris UE, Solow EB, et al. Predictors of adverse outcomes in patients with systemic lupus erythematosus transitioning to adult care. Semin Arthritis Rheum 2021;51:353–9.
42. Kind AJ, Jencks S, Brock J, et al. Neighborhood socioeconomic disadvantage and 30-day rehospitalization: a retrospective cohort study. Ann Intern Med 2014;161:765–74.

Arthritis Care & Research
Vol. 75, No. 10, October 2023, pp 2082–2087
DOI 10.1002/acr.25132

AMERICAN COLLEGE of RHEUMATOLOGY
Empowering Rheumatology Professionals

BRIEF REPORT

HLA–DRB1*15 and Eosinophilia Are Common Among Patients With Systemic Juvenile Idiopathic Arthritis

Alison M. Lerman,[1] Shawn A. Mahmud,[1] Zineb Alfath,[1] Benjamin W. Langworthy,[2] Patricia M. Hobday,[1] Mona M. Riskalla,[1] and Bryce A. Binstadt[1]

Objective. Concern exists that medications used to treat patients with systemic juvenile idiopathic arthritis (JIA), particularly interleukin (IL)-1 and IL-6 blocking agents, might be causing adverse drug reactions and lung disease (systemic JIA-LD). Carriage of HLA–DRB1*15 has been reported as a risk factor for adverse drug reactions among patients with systemic JIA. We performed a retrospective chart review to evaluate these factors at our center.

Methods. We reviewed the records of 86 subjects with systemic JIA followed for at least 6 months between 1996 and 2022. HLA typing was performed in 23 of the subjects. We compared characteristics of patients with or without eosinophilia. Among patients with HLA typing, we compared clinical characteristics of subjects with or without DRB1*15 and with or without systemic JIA-LD.

Results. Among the 23 patients with HLA typing, 74% carried DRB1*15, and 63% of patients without systemic JIA-LD carried DRB1*15. Seven subjects had systemic JIA-LD, all of whom carried DRB1*15. Patients with systemic JIA-LD were younger at the time of diagnosis and more likely to have had macrophage activation syndrome. Exposure to IL-1 and IL-6 blockers was common, occurring in 95% of patients. Eosinophilia occurred in 39% of patients with systemic JIA, often before IL-1 or IL-6 blockade. Eosinophilia was associated with adverse drug reactions and macrophage activation syndrome. There was 1 death, unrelated to active systemic JIA disease.

Conclusion. Carriage of DRB1*15 was more common in this cohort of patients with systemic JIA than in the general population. Eosinophilia and systemic JIA-LD were more common among patients with severe systemic JIA complicated by macrophage activation syndrome.

INTRODUCTION

Over the last 2 decades, some children with systemic juvenile idiopathic arthritis (JIA) have developed a severe form of interstitial lung disease termed systemic JIA–associated lung disease (systemic JIA-LD) (1–3). The emergence of systemic JIA-LD has paralleled increased use of biologic agents targeting interleukin (IL)-1 and IL-6 to treat systemic JIA, leading to concern that these medications might predispose to systemic JIA-LD. In retrospective analyses, we found that the majority of patients with systemic JIA-LD have had exposure to IL-1 and IL-6 inhibitors, but interpreting the relevance of this drug exposure history is confounded by the increasingly widespread use of these agents to treat systemic JIA over the same time period (1–3).

Recently, Saper and colleagues reported an association between carriage of the class II major histocompatibility complex allele DRB1*15 and the development of eosinophilia, systemic JIA-LD, and nonevanescent rashes different from the typical systemic JIA rash after exposure of patients with systemic JIA to IL-1 and/or IL-6 inhibitors (4). The authors classified these reactions as drug reactions with eosinophilia and systemic symptoms (DRESS). Over 80% of patients with systemic JIA and a history of DRESS-like reactions expressed a DRB1*15 allele, whereas only 7% of drug-tolerant patients with systemic JIA did. These observations have led to several working hypotheses to explain the potential mechanisms by which DRB1*15 might lead to apparent drug reactions to IL-1 or IL-6 blocking agents in patients with systemic JIA (4,5).

Dr. Lerman's work was supported by the Arthritis Foundation (Cultivating a New Generation of Rheumatologists Fellowship training award). Dr. Langworthy's work was supported by the NIH/Clinical and Translational Science Institute (grant UL1-TR-002494).

[1]Alison M. Lerman, MD, Shawn A. Mahmud, MD, PhD, Zineb Alfath, MD, Patricia M. Hobday, MD, Mona M. Riskalla, MD, Bryce A. Binstadt, MD, PhD: University of Minnesota and M. Health Fairview Masonic Children's Hospital, Minneapolis; [2]Benjamin W. Langworthy, PhD: University of Minnesota, Minneapolis.

Author disclosures are available online at https://onlinelibrary.wiley.com/doi/10.1002/acr.25132.

Address correspondence via email to Bryce A. Binstadt, MD, PhD, at binstadt@umn.edu.

Submitted for publication December 9, 2022; accepted in revised form April 11, 2023.

SIGNIFICANCE & INNOVATIONS

- The prevalence of HLA–DRB1*15:XX carriage in our patients with systemic juvenile idiopathic arthritis (JIA) was higher than in the general population.
- Eosinophilia was relatively common in our cohort (39%), often occurred prior to interleukin (IL)-1 or IL-6 inhibitor therapy, and did not differ based on the presence or absence of HLA–DRB1*15:XX.
- Many patients who carried HLA–DRB1*15:XX did not have eosinophilia, adverse drug reactions, or lung disease, despite 95% of all patients being exposed to IL-1 or IL-6 inhibitors.
- In our cohort, all patients with systemic JIA–lung disease expressed HLA–DRB1*15:XX and were also significantly younger at the age of diagnosis and more likely to have had a history of macrophage activation syndrome.

Given the degree of clinical concern raised by the prior report (4), we began routinely performing HLA typing on patients with systemic JIA at our center in the autumn of 2021. Of note, some patients with more severe disease, including systemic JIA-LD, had previously undergone HLA typing. We observed in routine clinical practice that many patients with systemic JIA carried DRB1*15:XX. We therefore designed this retrospective study with the primary objective of reporting the carriage rate of DRB1*15:XX among patients with systemic JIA at our center compared to the general population carriage rate. In addition, we performed a chart review of patients with confirmed systemic JIA seen at our center since 1996, regardless of whether HLA typing had been performed, with the objective of describing their drug-exposure history, incidence and type of adverse drug reactions, and incidence and timing of eosinophilia in relation to each other and in relation to DRB1*15:XX carriage status. Furthermore, we sought to determine the incidence of systemic JIA-LD in our cohort and to evaluate its association with previously reported risk factors, including DRB1*15:XX carriage, young age at disease onset, and episodes of macrophage activation syndrome (MAS). Here we describe our findings because we believe they provide critical information for the field, and we acknowledge the limitations inherent in our retrospective study design.

PATIENTS AND METHODS

Study design and population. We conducted a retrospective cohort study. The study design was reviewed by the University of Minnesota institutional review board (IRB) and deemed exempt from IRB oversight. Subjects were considered for inclusion if they received a billing diagnosis code for systemic JIA between 1996 and 2022 (International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, Clinical Modification codes: M08.20, M08.2A, M08.29) before age 17 years and had at least 6 months of follow-up in the University of Minnesota pediatric rheumatology clinics. A total of 124 patients had diagnosis codes for systemic JIA during the time interval of interest. Cases were manually adjudicated to confirm the diagnosis of systemic JIA and to confirm a minimum of 6 months of follow-up at our center. This 6-month minimum was chosen both to ensure that the diagnosis of systemic JIA was confirmed and to allow time to observe potential responses/reactions to medications. Among the 124 patients with systemic JIA, we excluded a total of 38, including 7 with new diagnoses of systemic JIA who have not yet been followed for 6 months, 6 with more remote diagnoses of systemic JIA who were not followed for at least 6 months, 9 who were determined to have an alternative diagnosis, 4 diagnosed at age 17 years or older, 6 never followed by our pediatric rheumatology service, and 6 with medical records too sparse to determine whether the patient had confirmed systemic JIA. After exclusion, 86 patients met criteria for further analysis. Data were abstracted from medical records into a secure spreadsheet. Three patients with systemic JIA-LD in our cohort were previously reported (3,4).

Data extraction. The primary source of information was the rheumatologist's clinical documentation. To determine HLA–DRB1 typing status, charts were searched for "HLA" and "DRB1." Searches for whole exome sequencing or geneticists' notes were also performed. The terminology DRB1*15:XX denotes DRB1*15:01 and DRB1*15:03.

For drug exposure history, we reviewed the treatment timeline in the rheumatologists' notes and the full prescribing history. We performed chart searches for "anakinra," "tocilizumab," "rilonacept", and "canakinumab." The timing of medication initiation and cessation were abstracted directly from the prescribing records in the electronic medical record whenever possible; when a medication initiation or cessation occurred outside of our system, the dates described in the primary rheumatologist's notes were used. The development of MAS and eosinophilia were examined in relationship to these prescribing patterns.

We defined eosinophilia as an absolute eosinophil count ≥700/microliter or eosinophils ≥10% of the total white blood cell count. We also performed a chart search for "eosinophilia" and "eosinophil." Eosinophilia was defined as related to MAS if it occurred within 2 weeks prior to the diagnosis of MAS or if eosinophils remained elevated during an episode of MAS, then decreased with MAS treatment.

To identify possible drug reactions, we reviewed the rheumatologist's notes and the allergy section of the chart. Additionally, we searched for the terms "drug allergy," "medication allergy," "drug reaction," and "medication reaction." We defined a drug reaction as a local injection site reaction, elevated aspartate aminotransferase or alanine aminotransferase (grouped here as liver function tests) without other explanation such as systemic JIA disease activity or MAS, rashes atypical for systemic JIA

(e.g., nonevanescent, pruritic, urticarial), or acute transfusion reactions (respiratory distress or swelling during or near the time of medication administration). One atypical drug reaction included diarrhea, convulsions, and confusion. To identify subjects with systemic JIA-LD, search terms included "lung disease," "fibrosis," and "interstitial," plus review of any prior chest computed tomography studies.

Statistical analysis. We tested whether the carriage rate of DRB1*15:XX was equal to the general population average of 25% (6) using a binomial test. To test the null hypothesis that age was the same between 2 groups of patients, we used a 2-sample *t*-test. For all other variables, we tested for differences in proportions using a chi-square test when all expected counts were >5, and a Fisher's exact test otherwise. The reported *P* values are not adjusted for multiple testing, and therefore should be considered exploratory, rather than confirming any specific hypothesis.

RESULTS

We identified 86 patients with systemic JIA who met our inclusion criteria. Seven of the 86 patients (8%) had systemic JIA-LD. There was 1 death in a patient who had undergone bone marrow transplantation for systemic JIA; this patient did not have systemic JIA-LD and had stable engraftment with no residual systemic JIA activity prior to death.

To date, we have performed HLA typing in 23 patients, including all 7 with systemic JIA-LD. These patients are shown in Table 1. A large fraction of systemic JIA expressed DRB1*15:XX (17 of 23 = 74%). This total included 15 with DRB1*15:01, 1 with DRB1*15:03, and 1 in whom subtyping beyond DRB1*15 was not provided. All 7 patients with systemic JIA-LD carried DRB1*15. Among the patients without systemic JIA-LD, the rate of DRB*15:XX carriage was 10 of 16 (63%). For both patients with and without systemic JIA-LD, the proportion who expressed DRB1*15:XX is larger than the general population average of 25%, with $P < 0.01$ in both cases (6). Similar to other reports, patients with systemic JIA-LD were younger at diagnosis and more likely to have had MAS. Trisomy 21 has been identified as a potential risk factor for systemic JIA-LD (2,3). In our cohort 3 patients had trisomy 21; all 3 expressed DRB1*15:01 but none has developed systemic JIA-LD. Among patients with HLA typing, eosinophilia was somewhat more common in patients with systemic JIA-LD than in those without (71% versus 50%), although this finding did not reach statistical significance.

We next evaluated the association between exposure to IL-1 or IL-6 inhibitor therapy, eosinophilia, and drug reactions. As shown in Table 2, the vast majority of all 86 systemic JIA patients (95%) were treated with IL-1 or IL-6 inhibitors, including 60 of 86 (70%) treated with anakinra, 52 of 86 (60%) with tocilizumab, and 29 of 86 (34%) with canakinumab. Thirty-nine patients (45%) received >1 of these agents. Eosinophilia was common in our cohort (33 of 86, 39%); its timing in relation to drug exposure is described later. Patients with eosinophilia were more likely to have had MAS and a history of a drug reaction.

Apparent adverse reactions to IL-1 or IL-6 blocking agents occurred in 17 of 86 patients; 1 additional patient had an adverse reaction to the interferon gamma (IFNγ) inhibitor emapalumab. HLA typing was performed in 12 of these 18 patients with drug reactions; 9 of 12 patients (75%) expressed DRB1*15:XX, whereas 3 of 12 (25%) did not. Nine patients had adverse reactions to anakinra, including 7 systemic reactions and 2 local injection site reactions; 3 of these reactions occurred in DRB1*15:XX–positive patients (1 liver function test elevation, 2 atypical rash), 2 in DRB1*15:XX–negative patients (1 pruritus and 1 reaction described as confusion, diarrhea, and convulsions), and 4 who did not have HLA typing (1 with atypical rash and liver function test elevation, 1 with petechial rash and eosinophilia, and 2 local injection site reactions). Eight patients had adverse reactions to tocilizumab, including 5 in DRB1*15:XX–positive patients (2 liver function test elevation, 1 urticarial rash, 2 respiratory distress during medication infusion), 1 in a DRB1*15:XX–negative patient (atypical rash), and 2 in patients without HLA typing (1 infusion

Table 1. Comparison of patients with systemic JIA with lung disease (systemic JIA-LD) to those without lung disease (n = 23 patients with HLA–DRB1 typing)*

	Systemic JIA-LD (n = 7)	No lung disease (n = 16)	Effect size (95% CI)	*P* (lung disease vs. no lung disease)
Age at diagnosis, mean (range) years	1.7 (1–3)	7.4 (1–16)	–5.7 (–9.7, –1.7)	0.01
HLA–DRB1*15:XX positive	7 (100)	10 (63)	Inf (0.6, Inf)	0.12
Male	3 (43)	6 (38)	1.2 (0.1, 10.5)	1
IL-1 or IL-6 blockade	7 (100)	15 (94)	Inf (0.0, Inf)	1
MAS	5 (71)	2 (13)	14.5 (1.4, 266.7)	0.01
Eosinophilia	5 (71)	8 (50)	2.4 (0.3, 32.5)	0.41
Trisomy 21	0 (0)	3 (19)	0.0 (0.0, 5.7)	0.53
History of drug reaction	6 (86)	6 (38)	9.0 (0.8, 503.1)	0.07

* Values are the number (%) unless indicated otherwise. For age at diagnosis, the effect size is the mean difference; for all other variables the effect size is the odds ratio. Odds ratio >1 indicates the given variable is more likely in systemic juvenile idiopathic arthritis with lung disease (JIA-LD). 95% CI = 95% confidence interval; IL = interleukin; Inf = infinite; MAS = macrophage activation syndrome.

Table 2. Characteristics of 86 patients with systemic JIA, divided by those with and without eosinophilia*

	All patients with systemic JIA (n = 86)	Eosinophilia (n = 33)	No eosinophilia (n = 53)	Effect size (95% CI)	*P* (eosinophilia vs. no eosinophilia)
Age at diagnosis, mean (range) years	6.7 (1–16)	6.3 (1–16)	7.0 (1–16)	−0.8 (−2.8, 1.3)	0.47
Male	33 (38)	13 (39)	20 (38)	1.1 (0.4, 2.6)	0.88
IL-1 or IL-6 blockade	82 (95)	32 (97)	50 (94)	1.9 (0.2, 103.8)	1
MAS	15 (17)	10 (30)	5 (9)	4.1 (1.3, 14.7)	0.01
Lung disease	7 (8)	5 (15)	2 (4)	4.5 (0.7, 49.8)	0.1
Trisomy 21	3 (3)	2 (6)	1 (2)	3.3 (0.2, 201.4)	0.56
History of drug reaction	18 (21)	11 (33)	7 (13)	3.2 (1.1, 10.0)	0.03

* Values are the number (%) unless indicated otherwise. For age at diagnosis, the effect size is the mean difference; for all other variables the effect size is the odds ratio. Odds ratio >1 indicates the given variable is more likely in patients with eosinophilia. 95% CI = 95% confidence interval; JIA = juvenile idiopathic arthritis; IL = interleukin; MAS = macrophage activation syndrome.

reaction, 1 liver function test elevation). None of these patients met clinical criteria for DRESS. These reactions led to discontinuation of the suspected offending drug in 12 of 18 patients. Medications used after patients had an adverse reaction to anakinra included methotrexate, tocilizumab, canakinumab, and ruxolitinib. Medications used after patients had an adverse reaction to tocilizumab included canakinumab, abatacept, and emapalumab.

We next sought to evaluate the potential association between DRB1*15:XX carriage, eosinophilia, and drug reactions. To avoid skewing our data based on the fact that all patients with systemic JIA-LD had undergone HLA typing, we excluded them from this analysis. Among the 16 patients without systemic JIA-LD who had HLA typing performed, age at diagnosis did not differ between DRB1*15:XX carriers and noncarriers (Table 3). In this group, the rates of eosinophilia were similar among DRB1*15:XX carriers (5 of 10 [50%]) and noncarriers (3 of 6 [50%]). For comparison, among patients without HLA typing, the incidence of eosinophilia was 20 of 63 (32%). We also observed similar rates of drug reactions in DRB1*15:XX carriers (3 of 10 [30%]) and noncarriers (3 of 6 [50%]).

We next sought to describe the timing of eosinophilia in relation to exposure to IL-1 or IL-6 inhibitor therapy. To evaluate this relationship, patients had to have laboratory data available prior to initiation of these agents. Of the total 33 patients with eosinophilia, 4 lacked pretreatment data and 1 was not exposed to IL-1 or IL-6 inhibitors. Among the remaining 28 patients, eosinophilia occurred prior to IL-1 or IL-6 blockade in 12 (43%). Six of those 12 patients had resolution of eosinophilia after initiation of therapy, whereas the other 6 had some degree of ongoing eosinophilia. Table 4 compares the 12 patients who had eosinophilia prior to initiation of therapy to the 16 patients who had eosinophilia only after initiation of therapy. Importantly, we identified no significant differences between these 2 groups. Ten of the 33 total patients with eosinophilia had a history of MAS, and eosinophilia occurred during MAS in 5 of 10 (50%). We also examined the relationship between the development of MAS and the initiation of anti–IL-1 or anti–IL-6 therapy. In the 15 patients who had MAS during their clinical course, 5 had MAS before initiation of treatment but not after, 3 patients had MAS both before and after initiation of treatment, and 7 had MAS only after initiation of treatment.

DISCUSSION

The most striking finding in this cohort of patients with systemic JIA is the high rate of DRB1*15:XX carriage (74%). Even among subjects without lung disease, the rate of DRB1*15:XX carriage was high (63%). This rate is similar to that reported by the Boston Children's Hospital group (7). In our cohort, all 7 patients with systemic JIA-LD expressed DRB1*15:XX. Eosinophilia was also common, occurring in 39% of all patients with systemic JIA and often prior to IL-1 or IL-6 inhibitor therapy.

Table 3. Comparison of systemic JIA patients without lung disease based on DRB1*15:XX status (n = 16)*

	DRB1*15:XX positive (n = 10)	DRB1*15:XX negative (n = 6)	Effect size (95% CI)	*P*
Age at diagnosis, mean (range) years	6.2 (1–16)	9.3 (5–16)	−3.1 (−8.6, 2.3)	0.24
Male	4 (40)	2 (33)	1.31 (0.1, 21.3)	1
IL-1 or IL-6 blockade	9 (90)	6 (100)	0 (0, 64.9)	1
MAS	1 (10)	1(17)	0.6 (0.01, 52.55)	1
Eosinophilia	5 (50)	3 (50)	1 (0.09, 11.69)	1
Trisomy 21	3 (30)	0 (0)	Inf (0.25, Inf)	0.25
History of drug reaction	3 (30)	3 (50)	0.45 (0.03, 5.51)	0.61

* Values are the number (%) unless indicated otherwise. For age at diagnosis, the effect size is the mean difference; for all other variables the effect size is the odds ratio. Odds ratio >1 indicates the given variable is more likely in patients who are DRB1*15:XX positive. 95% CI = 95% confidence interval; IL = interleukin; Inf = infinite; MAS = macrophage activation syndrome; JIA = juvenile idiopathic arthritis.

Table 4. Comparison of systemic JIA patients with complete eosinophilia data prior to initiation of anti–IL-1 or anti–IL-6 blockade versus patients with eosinophilia only after initiation of these treatments (n = 28)*

	Eosinophilia before treatment (n = 12)	Eosinophilia only after treatment (n = 16)	Effect size (95% CI)	*P*
Age at diagnosis, mean (range) years	6.67 (1–16)	6.75 (1–16)	−0.08 (−4.4, 4.2)	0.97
Male	5 (42)	6 (38)	1.18 (0.2, 7.0)	1
MAS	4 (33)	6 (38)	0.84 (0.1, 5.1)	1
Trisomy 21	1 (8)	1 (6)	1.35 (0.0, 114.0)	1
History of drug reaction	4 (33)	6 (38)	0.84 (0.1, 5.1)	1
HLA–DRB1*15:XX positive	4 (10)	5 (6)	Inf (0.2, Inf)	0.49
Lung disease	2 (17)	2 (12)	1.38 (0.1, 22.2)	1

* Values are the number (%) unless indicated otherwise. For age at diagnosis, the effect size is the mean difference; for all other variables the effect size is the odds ratio. Odds ratio >1 indicates the given variable is more likely in those with eosinophilia before treatment. 95% CI = 95% confidence interval; IL = interleukin; Inf = infinite; MAS = macrophage activation syndrome; JIA = juvenile idiopathic arthritis.

One limitation of our retrospective study is the relatively small sample size, particularly the number of patients who had HLA typing performed. Because of this limitation, several comparisons have wide confidence intervals for effect size, indicating uncertainty about the true association, or potential lack thereof. A larger sample is needed to identify differences in characteristics between DRB1*15:XX carriers and noncarriers. We also note that for those without systemic JIA-LD, the patients who were HLA typed were those who were evaluated after autumn 2021. This time frame of routine HLA typing could potentially lead to biased results if the patients who have visited the clinic after this date are systematically different from those who have not, but we do not have a reason to believe this is the case.

A very high fraction (95%) of subjects in our cohort were exposed to IL-1 or IL-6 inhibitors. Apparent drug reactions of any type occurred in 21% of our cohort, with 19% having a systemic reaction and 2% having local injection site reactions only. Of note, 6 of the 7 patients with systemic JIA-LD had a history of a drug reaction during their clinical course; 2 had reactions to anakinra, 3 to tocilizumab, and 1 to emapalumab. Notably, the drug reactions occurred after the diagnosis of systemic JIA-LD in 2 of these 6 patients. One patient with systemic JIA-LD had no history of drug reactions. In the report by Saper and colleagues, all subjects with lung disease scored as having DRESS during treatment with IL-1 and IL-6 inhibitors (4). In contrast, none of our patients with systemic JIA-LD were diagnosed with DRESS.

A critical issue is whether many patients with systemic JIA are developing DRESS in the setting of IL-1 or IL-6 inhibitors or not. The cohort in which the association of DRESS and DRB1*15 was first reported was, by design, enriched for subjects with probable drug reactions to IL-1 or IL-6 blocking agents. The authors applied RegiSCAR scoring criteria for DRESS and considered subjects with scores ≥4 to have "probable DRESS" (4,8). They reported 94 total patients with systemic JIA or adult-onset Still's disease in whom HLA typing was performed. Among these, 64 (68%) were considered to have had probable DRESS, including 45 with lung disease. The other 30 (32%) were classified as drug-tolerant controls. This classification led to the conclusion that HLA–DRB1*15:XX was enriched among subjects with "Still's-DRESS" cases relative to "drug-tolerant" controls as well as to a larger global systemic JIA genetic repository, the International Childhood Arthritis Genetics Consortium (INCHARGE) (4).

A challenge with applying the RegiSCAR scoring for DRESS to patients with systemic JIA, as we have highlighted (5), is the shared features of the 2 conditions, including lymph node enlargement, organ involvement such as elevated liver enzymes or splenomegaly, and exclusion of other potential etiologies such as infections. Adding eosinophilia and/or characteristic DRESS rash elevates the score to at least 4, leading to classification as "probable DRESS," although the majority of subjects reported by Saper et al did have RegiSCAR scores of 6 or higher (4). The high rate of eosinophilia in our cohort (39%) is notable in this regard. Thus, the key issue is whether these patients truly have DRESS reactions to IL-1 or IL-6 blockers, or whether the same symptoms and signs, particularly eosinophilia, are simply common manifestations of systemic JIA, a notion supported by our finding that 43% of patients with eosinophilia had it prior to IL-1 or IL-6 blockade. Eosinophilia could possibly arise due to the biologic activity of IL-1 or IL-6 blockade, as proposed in the cytokine plasticity hypothesis (5).

If one analyzes the Saper cohort through the lens that eosinophilia is common among patients with systemic JIA, a different conclusion can be reached. Specifically, among all 94 systemic JIA and adult-onset Still's disease subjects with HLA typing performed, 52 (55%) carried DRB1*15:XX (4). That carriage frequency is on par with the data we report here (74%) and those from Boston Children's (49%) (7), although again the Saper cohort was intentionally enriched for subjects considered to have probable drug reactions. Those carriage frequencies are all higher than those reported in the INCHARGE cohort (approximately 25%) as well as those reported in a recent abstract with 65 Dutch patients (26%), and also higher than those reported for the general US population (DRB1*15 allele frequency in the US White population is 15.8%) (4,6,9,10). These findings could be due to differences in demographic or ancestral characteristics of the populations, although this possibility seems unlikely since all were

predominantly White. Of note, the 3 studies reporting higher rates of DRB1*15:XX carriage, including this one, were retrospective and applied chronicity metrics for study inclusion (4,7). Similarly, the subgroup of patients who had HLA typing performed in these studies included a higher fraction of subjects with systemic JIA-LD compared to the entire systemic JIA cohorts, so the groups are enriched for patients with more severe disease. In contrast, the Dutch study was prospective, and thus may have been more likely to enroll subjects with shorter disease duration (so-called monophasic systemic JIA). The Dutch cohort also included only 1 subject with systemic JIA-LD (10). Considering these factors, we hypothesize that DRB1*15:XX might not represent a risk factor for adverse reactions to IL-1 or IL-6 inhibitors, but rather a risk factor for more chronic or severe systemic JIA. The other MHC class II allele previously identified as a risk factor for systemic JIA is DRB1*11 (9). Considering these data alongside our own, both DRB1*11 and DRB1*15:XX seem to be enriched among patients with systemic JIA relative to the healthy control population, and DRB1*15:XX might also increase the risk of more severe disease.

How might DRB1*15:XX promote chronic or severe systemic JIA? MHC class II molecules present antigens to CD4+ T cells. We envision that patients with systemic JIA and DRB1*15:XX may harbor populations of CD4+ T cells that recognize particular, nondrug DRB1*15:XX–presented endogenous or exogenous antigenic peptides, for instance from the microbiome or common infectious agents. As proposed in the cytokine plasticity hypothesis (5), introduction of IL-1 or IL-6 blocking agents may alter the cytokine production profile of those T cells in ways that promote eosinophilia (e.g., Th2 skewing) and/or MAS and systemic JIA-LD (e.g., Th1 skewing with overproduction of IFNγ and IL-18). A current challenge to the field is to define these hypothetical DRB1*15–presented antigens and to identify and better characterize the responding T cell populations. DRB1*15:XX could of course promote more severe systemic JIA through other mechanisms, but to date none has been hypothesized.

If DRB1*15:XX is indeed a risk factor for more chronic or severe systemic JIA, including systemic JIA-LD, how would that fact inform clinical practice? We suggest that HLA typing may be useful at the time of diagnosis. Based on our data and clinical experience, we feel that patients who carry DRB1*15:XX can still be treated with IL-1 or IL-6 blockers, but concomitant use of corticosteroids or traditional nonbiologic DMARDS such as methotrexate should be considered. Careful monitoring for the development of systemic JIA-LD, particularly in younger patients and those with MAS, would be reasonable.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Lerman, Mahmud, Binstadt.

Acquisition of data. Lerman, Mahmud, Alfath, Hobday, Riskalla, Binstadt.

Analysis and interpretation of data. Lerman, Mahmud, Alfath, Langworthy, Binstadt.

REFERENCES

1. Kimura Y, Weiss JE, Haroldson KL, et al. Pulmonary hypertension and other potentially fatal pulmonary complications in systemic juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2013;65:745–52.
2. Saper VE, Chen G, Deutsch GH, et al. Emergent high fatality lung disease in systemic juvenile arthritis. Ann Rheum Dis 2019;78:1722–31.
3. Schulert GS, Yasin S, Carey B, et al. Systemic juvenile idiopathic arthritis-associated lung disease: characterization and risk factors. Arthritis Rheumatol 2019;71:1943–54.
4. Saper VE, Ombrello MJ, Tremoulet AH, et al. Severe delayed hypersensitivity reactions to IL-1 and IL-6 inhibitors link to common HLA-DRB1*15 alleles. Ann Rheum Dis 2022;81:406–15.
5. Binstadt BA, Nigrovic PA. The conundrum of lung disease and drug hypersensitivity-like reactions in systemic juvenile idiopathic arthritis. Arthritis Rheumatol 2022;74:1122–31.
6. Gonzalez-Galarza FF, McCabe A, Santos E, et al. Allele frequency net database (AFND) 2020 update: gold-standard data classification, open access genotype data and new query tools. Nucleic Acids Res 2020;48:D783–88.
7. Wobma H, Arvila SR, Taylor ML, et al. Incidence and risk factors for eosinophilia and lung disease in biologic-exposed children with systemic juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2023;75:2063–72.
8. Kardaun SH, Sekula P, Valeyrie-Allanore L, et al. Drug reaction with eosinophilia and systemic symptoms (DRESS): an original multisystem adverse drug reaction. Results from the prospective RegiSCAR study. Br J Dermatol 2013;169:1071–80.
9. Ombrello MJ, Arthur VL, Remmers EF, et al. Genetic architecture distinguishes systemic juvenile idiopathic arthritis from other forms of juvenile idiopathic arthritis: clinical and therapeutic implications. Ann Rheum Dis 2017;76:906–13.
10. Erkens R, Sinha R, Pickering A, et al. First line treatment using recombinant IL-1 receptor antagonist in new onset systemic juvenile idiopathic arthritis is an effective treatment strategy, irrespective of HLA DRB1 background [abstract]. Arthritis Rheumatol 2022;74 Suppl 9. URL: https://acrabstracts.org/abstract/first-line-treatment-using-recombinant-il-1receptor-antagonist-in-new-onset-systemic-juvenile-idiopathic-arthritis-is-an-effective-treatment-strategy-irrespective-of-hla-drb1-background/.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 75, No. 10, October 2023, pp 2088–2095
DOI 10.1002/acr.25107

Role of Platelet-Bound C4d (PC4d) in Predicting Risk of Future Thrombotic Events in Systemic Lupus Erythematosus

Yevgeniya Gartshteyn,[1] John Conklin,[2] Michelle A. Petri,[3] Vasileios C. Kyttaris,[4] Daniel W. Goldman,[3] Anja Kammesheidt,[2] Anca D. Askanase,[1] and Roberta Vezza Alexander[2]

Objective. Platelet-bound complement activation product C4d (PC4d) levels correlate with history of thrombosis in patients with systemic lupus erythematosus (SLE). The present study evaluated whether PC4d levels could assess risk of future thrombosis events.

Methods. PC4d level was measured by flow cytometry. Thromboses were confirmed by electronic medical record data review.

Results. The study included 418 patients. Nineteen events (13 arterial and 6 venous) occurred in 15 subjects in the 3 years post-PC4d level measurement. PC4d levels above the optimum cutoff of 13 mean fluorescence intensity (MFI) predicted future arterial thrombosis with a hazard ratio of 4.34 (95% confidence interval [95% CI] 1.03–18.3) ($P = 0.046$) and a diagnostic odds ratio (OR) of 4.30 (95% CI 1.19–15.54). Negative predictive value of PC4d level of ≤13 MFI for arterial thrombosis was 99% (95% CI 97–100%). Although a PC4d level of >13 MFI did not reach statistical significance for prediction of total thrombosis (arterial and venous) (diagnostics OR 2.50 [95% CI 0.88–7.06]; $P = 0.08$), it was associated with all thrombosis (n = 70 historic and future arterial and venous events in the 5 years pre- to 3 years post-PC4d level measurement) with an OR of 2.45 (95% CI 1.37–4.32; $P = 0.0016$). In addition, the negative predictive value of PC4d level of ≤13 MFI for all future thrombosis events was 97% (95% CI 95–99%).

Conclusions. A PC4d level of >13 MFI predicted future arterial thrombosis and was associated with all thrombosis. Patients with SLE presenting with a PC4d level of ≤13 MFI had high probability of not experiencing arterial or any thrombosis in the 3 years afterwards. Taken together, these findings indicate that PC4d levels may help predict the risk of future thrombosis events in SLE.

INTRODUCTION

The incidence of both venous and arterial thrombosis is increased in patients with systemic lupus erythematosus (SLE) compared to the general population. Overall, the risk of cardiovascular disease and events such as stroke and myocardial infarction is ∼2 times higher than in the general population (1,2). A recent meta-analysis showed that cardiovascular events occurred within a median of 8 years in 25.4% of patients with SLE, with 4.1% of patients experiencing a myocardial infarction and 7.3% a stroke (3). It is estimated that 9% of the patients with SLE will develop a venous thrombosis within 20 years of SLE diagnosis (2).

Among the autoantibodies, antiphospholipid antibodies (aPL), in particular lupus anticoagulant (LAC), are strongly associated with thrombosis. In patients with SLE, aPL are present in 30–40% and are well-established predictive biomarkers of thrombosis. In fact, aPL positive SLE patients have higher prevalence of both arterial and venous thrombosis compared to aPL negative patients (∼40% versus 10–20%) (4). LAC is the aPL most strongly and conclusively associated with arterial and venous thrombosis (5,6) and is the best predictor of thrombosis (all, venous, and

Supported by Exagen, Inc.
[1]Yevgeniya Gartshteyn, MD, Anca D. Askanase, MD, MPH: Columbia University, New York, New York; [2]John Conklin, MS (current address: Jexa Analytics, Oceanside, California), Anja Kammesheidt, PhD (current address: Anja-Gen Consulting, Laguna Beach, California), Roberta Vezza Alexander, PharmD, PhD (current address: Biological Dynamics, San Diego, California): Exagen Inc., Vista, California; [3]Michelle A. Petri, MD, MPH, Daniel W. Goldman, PhD: Johns Hopkins University, Baltimore, Maryland; [4]Vasileios C. Kyttaris, MD: Beth Israel Deaconess Medical Center, Boston, Massachusetts.
Dr. Gartshteyn and Mr. Conklin contributed equally to this work.

Data availability. All data generated and analyzed in this study are available from Exagen Inc. upon reasonable request. Requests for access to data generated in this study should be made to the corresponding author.
Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.25107&file=acr25107-sup-0001-Disclosureform.pdf.
Address correspondence via email to Yevgeniya Gartshteyn, MD, at yg2372@cumc.columbia.edu.
Submitted for publication October 17, 2022; accepted in revised form February 16, 2023.

SIGNIFICANCE & INNOVATIONS

- This is the first study to investigate whether the cell-bound complement activation product, C4d bound to platelets (PC4d), can assess risk of future thrombosis events in patients with systemic lupus erythematosus (SLE).
- PC4d levels above the optimum cutoff of 13 mean fluorescence intensity (MFI), as determined by receiver operating characteristic curve analysis, predicted future arterial thrombosis.
- On the other hand, patients with a PC4d level of ≤13 MFI had a high probability of not experiencing arterial or any thrombosis event in the 3 years after PC4d level measurement.
- As PC4d level may help evaluate the risk of future thrombosis events in patients with SLE, it may inform patient monitoring, control of other thrombosis risk factors, and lead to institution of appropriate pharmacologic preventive treatment.

arterial) in patients with SLE (2,7). LAC positivity is, however, insufficient to explain the entire thrombotic risk in SLE.

We and others have shown that cell-bound complement activation products and, in particular, platelet-bound complement activation product C4d (PC4d), is associated with a history of thrombosis in SLE (8–11). The present study was conducted to evaluate the predictive value of PC4d by analyzing 3 patient cohorts enrolled at Johns Hopkins University (8), Columbia University (10), and Beth Israel Deaconess Medical Center.

PATIENTS AND METHODS

Patient cohorts. Patients from the lupus cohorts of Johns Hopkins University (JH), Columbia University (COL), and Beth Israel Deaconess Medical Center (BI) were included in this study between April and September 2017 (JH patient cohort), August 2018 and January 2020 (COL patient cohort), and October 2018 and March 2022 (BI patient cohort). The study protocols and consent forms were approved by the Institutional Review Boards of Columbia University (AAAN0550), Johns Hopkins University (study number IRB00118914), and Beth Israel Deaconess Medical Center (2006-P000298). All patients provided written informed consent.

To expand on the data reported by Petri et al (8) and Gartshteyn et al (10), we combined the patient cohorts in those 2 studies (enrolled at JH and COL, respectively) with the patients enrolled at BI. Evaluating the prospective data across these 3 cohorts allowed us to study not only the association of PC4d with history of thrombosis, but also the ability of PC4d level to predict future events.

All patients fulfilled the 1997 American College of Rheumatology (12) and/or the Systemic Lupus International Collaborating Clinics (13) classification criteria for SLE. Disease activity was evaluated using the Safety of Estrogens in Lupus Erythematosus National Assessment version of the SLE Disease Activity Index (SLEDAI) (14) or the SLEDAI 2000 (15). Body mass index was calculated as weight in kilograms divided by the height in meters squared.

Following PC4d measurement, we continued to collect data on thrombotic events in the 3 years after the baseline visit to evaluate whether PC4d level could predict future thrombotic events. Arterial and venous events were confirmed by medical record review.

Biomarkers. Apart from lupus anticoagulant, all biomarkers were measured at Exagen in blood samples collected at the same time at the study visit. Lupus anticoagulant was collected at the same time but measured at JH for the patients in that cohort. PC4d level was measured by flow cytometry following Exagen's standard operating procedures, as described (8,10,16). Briefly, upon receipt of samples collected at the visit at Exagen's clinical laboratory, red blood cells from ethylenediamine-tetraacetic acid (EDTA)–anticoagulated blood were lysed and platelets were stained using a mouse monoclonal antibody against human C4d (Quidel) or a mouse IgG1 isotype control (MOPC-21). After incubation, samples were stained with a goat anti-mouse antibody conjugated to fluorescein isothiocyanate. A mouse anti-human monoclonal antibody against human CD42b conjugated to phycoerythrin (PE) was used to identify platelets. Flow cytometry analysis was performed using a Gallios flow cytometer (Beckman Coulter). Light scatter (forward and side) gating parameters were used to isolate the platelet population, followed by secondary gating based on positive CD42b PE staining. Quantification of the non-specific (isotype control) and specific (C4d) fluorescence was determined for the CD42b PE gated platelets (5,000 events).

Values are reported as net mean fluorescence intensity (MFI), which was determined by subtraction of isotype control MFI from the specific C4d MFI on gated platelets. For 1 subject without any thrombosis, the PC4d value was not available at the baseline visit. Serum complement proteins C3 and C4 were measured by standard immunoturbidimetry assay (The Binding Site) and were considered low if below the manufacturer's cutoff levels (81.1 mg/dl and 12.9 mg/dl, respectively).

All aPLs were measured in serum or plasma from venous blood collected in EDTA when serum was not available. All assays were carried out in Exagen's clinical laboratory upon receipt of samples collected at the clinical sites. Manufacturers' instructions were followed, and the cutoffs used by the manufacturers were applied to calculate positivity. In all the COL and JH cohort samples and in the majority of the BI cohort samples, anti-cardiolipin (aCL) and anti-beta2 glycoprotein 1 (aB2GP1) were measured by chemiluminescence (QUANTA Flash; Werfen) and all the isotypes were considered positive if >20 chemiluminescent units

(CU). Anti-phosphatidylserine/prothrombin complex antibodies (aPS/PT) IgG and IgM were measured by enzyme-linked immunosorbent assay (ELISA) (QUANTA Lite; Werfen) and were considered positive if >30 units. A subset of samples from the BI cohort were tested for aB2GP1 and aCL by ELISA fluorescence enzyme immunoassay (Phadia; ThermoFisher Scientific) instead of chemiluminescence. aCL IgG and IgM measured by ELiA were considered positive if >40 units/ml; aCL IgA if >20 units/ml; all the isotypes of aB2GP1 if >10 units/ml. We demonstrated previously that results obtained with the Quanta Flash and the ELiA platforms are well correlated (16). Thus, no distinctions between platforms were made when calculating aPL positivity. Patients were considered aCL, aB2GP1, and aPS/PT positive if positive for at least 1 of the corresponding isotypes and were considered triple positive if positive for at least 1 aCL isotype, 1 aB2GP1 isotype, and 1 aPS/PT isotype.

Statistical analysis. Data were analyzed by Fisher's exact, analysis of variance, Mann-Whitney U test, logistic regression, and hazard ratio as reported herein. Confusion matrix analysis was performed to calculate the performance characteristics of PC4d, including sensitivity, specificity, diagnostics odds ratio (OR), and negative and positive predictive value. Logistic regression was performed to estimate multivariate effects. Kaplan-Meier with log-rank test and Cox proportional hazards model was performed for time to arterial events. Analysis was conducted in R (R Core Team, version 4.1) with salient packages pROC (17) for receiver operating characteristic (ROC) curve analysis and epiR (version 2.0, 2022; https://CRAN.R-project.org/package=epiR) for confusion matrix derived statistics.

For a given subject, analysis of thrombotic events was based on the event closest in time to a given PC4d level measurement. If a PC4d level was measured more than once for a given patient, only the PC4d level value preceding and closest to a thrombosis was included in the analysis of events that occurred after PC4d. This approach allows a one-to-one relation between subject, event, and PC4d measurement for analysis.

RESULTS

A total of 418 SLE patients were enrolled in the study, including 149 at JH, 150 at COL, and 119 at BI. Main demographic and clinical characteristics are reported in Table 1. The majority of patients were female, and average age ranged between 38.7 and 48.7 years. Disease activity, measured by SLEDAI (in all cohorts) and by physician global assessment (in 2 of the 3 cohorts), was consistent with mild–moderate disease. Approximately 81% of patients were prescribed hydroxychloroquine, while glucocorticoids were used by approximately one-third of the patients.

Nineteen events (13 arterial and 6 venous) occurred in 15 subjects in the 3 years after PC4d level measurement, including 8 cerebrovascular accidents, 2 gastrointestinal infarctions, 2 myocardial infarctions, 3 deep vein thromboses, 1 pulmonary embolism, 2 venous and 1 arterial thrombosis not specified. PC4d levels closest to the event were higher, although not statistically significant, in the 15 subjects with as compared to the 403 subjects without future thrombosis events (median 11.35 MFI [interquartile range (IQR) 4.6–19.7] versus median 5.00 MFI [IQR 2.66–13.34]; 2,385, $n_1 = 15$, $n_2 = 403$; $P = 0.173$ by Mann-Whitney U test). BC4d levels and EC4d levels closest to the event were also higher in the subjects with thrombosis (median 43.30 MFI [IQR 25.8–64.3] and median 13.64 MFI [IQR 6.02–31.90], respectively) as compared to the subjects without future thrombosis events (median 33.87 MFI [IQR 19.5–70.49] and median 10.07 MFI [IQR 6.22–19.16], respectively). However, these differences between subjects with and without thrombosis were not statistically significant for both BC4d levels (2,421.5, $n_1 = 15$, $n_2 = 403$; $P = 0.570$ by Mann-Whitney U test) and EC4d levels (2,703, $n_1 = 15$, $n_2 = 403$; $P = 0.487$ by Mann-Whitney U test).

Because platelet dysfunction is classically associated with arterial vascular disease, and more arterial than venous events occurred in the 3 years after PC4d, we evaluated whether PC4d level might predict future arterial thrombosis events. Two subjects had 2 arterial events each after PC4d level measurement, and 1 subject had 2 PC4d level measurements and 1 arterial event after each PC4d level determination. Thus, of the 13 total arterial events, the 11 arterial events closest to PC4d level measurements (in 10 subjects) were included in the analysis to predict future arterial thrombosis. An ROC curve analysis established that the optimum cutoff was 13 MFI (Figure 1). A PC4d level of >13 MFI was statistically significant ($P = 0.026$) at predicting future arterial thrombosis with sensitivity of 60% (95% confidence interval [95% CI] 26–88%) and specificity of 74% (95% CI 70–78%). The diagnostics OR was 4.30 (95% CI 1.19–15.54), indicating that the odds of correctly predicting a thrombosis event if PC4d level was >13 MFI was 4.30 times greater than a false prediction if PC4d level was ≤13 MFI. Kaplan-Meier analysis with Cox proportional hazard model showed a statistically significant difference for time to arterial events in patients with a PC4d level of >13 MFI versus a PC4d level of ≤13 MFI (hazard ratio 4.34 [95% CI 1.03–18.3]; $P = 0.046$). The negative predictive value of a PC4d level ≤13 MFI for arterial thrombosis was 99% (95% CI 97–100%), indicating that the estimated probability of not having an arterial thrombosis within 3 years after PC4d level measurement was 99% if PC4d level was ≤13 MFI. Not surprisingly, given the low prevalence of events, the positive predictive value of a PC4d level of >13 MFI for arterial thrombosis was 5% (95% CI 2–11%), indicating that the estimated probability of having an arterial thrombosis within 3 years after PC4d level measurement was 5% if PC4d level was >13 MFI.

Similarly, a PC4d level of ≤13 MFI had negative predictive value of 97% (95% CI 95–99%) and a PC4d level of >13 had a positive predictive value of 6% (95% CI 3–13%) for all future thrombosis risk when the 11 arterial and 6 venous thrombosis

Table 1. Demographic and clinical variables at baseline in the entire (total) population and in the 3 cohorts (JH, COL, and BI cohorts)*

	Total	JH	COL	BI
No.	418	149 (35.6)	150 (35.8)	119 (28.5)
Female sex	374 (89.5	128 (85.9	137 (91.3	109 (91.6
Age, mean (95% CI) years	42.7 (41.3–44.0)†	48.7 (46.3–51.0)	39.8 (37.8–41.8)	38.7 (36.4–41.0)
Age <65 years	383 (91.4)	127 (85.2)	139 (92.7)	117 (97.5)
Race and ethnicity				
Black	124 (29.6)	51 (34.2)	39 (26.0)	34 (28.3)
Asian	38 (9.1)	6 (4.0)	14 (9.3)	18 (15.0)
White	174 (41.6)	83 (55.7)	40 (26.7)	51 (42.9)
Hispanic	65 (15.5)	0 (0.0)	49 (32.7)	16 (13.3)
Other	12 (2.9)	9 (6.0)	3 (2.0)	0 (0.0)
Not available	5 (1.2)	0 (0.0)	5 (3.3)	0 (0.0)
BMI, mean (95% CI)	27.8 (27.1–28.5)	27.6 (26.5–28.7)	27.9 (26.8–28.9)	27.9 (26.0–29.8)
SLEDAI, mean (95% CI)	4.0 (3.6–4.4)†	2.5 (2.1–2.9)	5.8 (5.0–6.5)	3.5 (2.7–4.2)
PGA, mean (95% CI)	0.7 (0.6–0.8)	0.7 (0.5–0.8)	N/A	0.7 (0.6–0.9)
aCL positive‡	84 (20.1)	39 (26.2)	30 (20.0)	15 (12.6)
Low C3	75 (17.9)	17 (11.4)	27 (18.0)	31 (26.1)
Low C4	75 (17.9)	24 (16.1)	24 (16.0)	27 (22.7)
aB2GP1 positive‡	86 (20.6)	36 (24.2)	34 (22.7)	16 (13.4)
aPS/PT positive§	151 (36.1)	72 (48.3)	44 (29.3)	35 (29.4)
Any aPL positive	187 (44.7)	80 (53.7)	64 (42.7)	43 (36.1)
aPL triple positive¶	46 (11.0)	24 (16.1)	13 (8.7)	9 (7.6)
LAC positive#	N/A	88 (59.1)	N/A	N/A
BC4d >60 MFI	117 (28.0)	38 (25.5)	44 (29.3)	35 (29.4)
EC4d >14 MFI	150 (35.9)	49 (32.9)	60 (40.0)	41 (34.5)
PC4d >10 MFI	132 (31.6)	52 (34.9)	43 (28.7)	37 (31.1)
PC4d >13 MFI	111 (26.6)	46 (30.9)	36 (24.0)	29 (24.4)
PC4d >20 MFI	77 (18.4)	33 (22.1)	22 (14.7)	22 (18.5)
HCQ	338 (80.7)	131 (87.9)	122 (81.3)	85 (71.4)
Glucocorticoids	135 (32.3)	48 (32.2)	47 (31.3)	40 (33.6)
Azathioprine	38 (9.1)	12 (8.1)	14 (9.3)	12 (10.1)
Mycophenolate	110 (26.3)	35 (23.5)	40 (26.7)	35 (29.4)

* Values are the number (%) unless indicated otherwise. Positivity rate for PC4d was based on 3 cutoffs, 10, 13, and 20 net MFI. 95% CI = 95% confidence interval; aB2GP1 = anti-β2 glycoprotein 1 antibodies; aCL = anti-cardiolipin antibodies; aPL = antiphospholipid antibodies; aPS/PT = anti-phosphatidylserine/prothrombin complex antibodies; BI = Beth Israel Deaconess Medical Center; BMI = body mass index; COL = Columbia University; HCQ = hydroxychloroquine; JH = Johns Hopkins University; LAC = historic lupus anticoagulant; MFI = mean fluorescence intensity; N/A = not applicable; PC4d = platelet-bound C4d; PGA = physician global assessment; SLEDAI = SLE disease activity index.
† $P < 0.001$ by analysis of variance.
‡ Considered aCL or aB2GP1 positive if at least 1 of the isotypes (IgG, IgM, or IgA) was positive.
§ Considered aPS/PT positive if IgG and/or IgM positive.
¶ Considered triple positive if positive for at least 1 aCL isotype, at least 1 anti-β2-GP1 isotype, and at least 1 aPS/PT isotype.
# LAC was available for 145 patients of the JH cohort (97% of the JH cohort and 34.7% of the total patients) while no LAC data were available for patients in the COL and BI cohorts.

events that happened after PC4d measurement (and closest to PC4d level measurement) were analyzed. However, likely because of the small number of events, a PC4d level of >13 MFI did not reach statistical significance for prediction of all future thrombotic events (arterial and venous) (diagnostics OR 2.50 [95% CI 0.88–7.06]; $P = 0.08$). Similar results were obtained when only patients younger than 65 years of age were included in the analysis (data not shown).

We evaluated the association between presence of antiphospholipid antibodies (aCL, aB2GP1, and aPS/PT) and future thrombosis risk. Of the 15 subjects who experienced a thrombosis event after PC4d level measurement, only 1 was triple aPL positive. Among the remaining 14 patients who were not triple aPL positive, 6 (43%) had a PC4d level of >13 MFI (Table 2), indicating that the probability of having a thrombosis event if PC4d level was >13 is independent of triple aPL positivity. In the population without thrombosis, 114 of the 297 subjects with a PC4d level of ≤13 MFI (38%) were positive for at least 1 aPL measured in this study, and 24 of 297 (8.1%) were triple aPL positive (Table 2), indicating that the probability of not having a thrombosis event if PC4d level was ≤13 MFI holds true for the population that was aPL positive and at higher risk. Because LAC was measured at the baseline visit only in the JH cohort and historic LAC data were available only for the subjects in the JH cohort, the contribution of LAC to thrombosis could not be evaluated in this study.

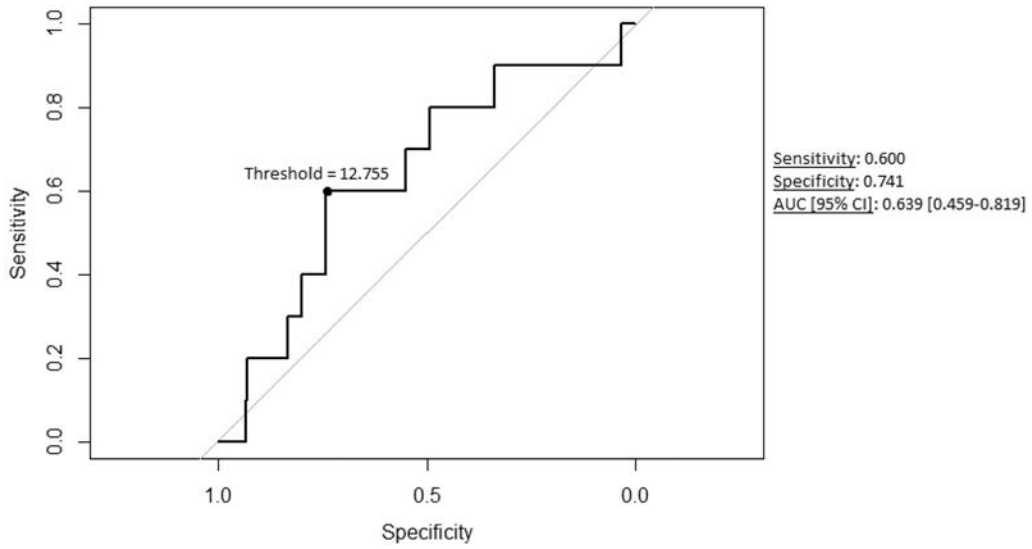

Figure 1. Receiver operating characteristic (ROC) curve of arterial thrombotic events in the 3 years after platelet-bound complement activation product C4d (PC4d) measurement. The arterial thrombotic events that occurred after PC4d measurement were used for ROC curve analysis and a cutoff of 13 mean fluorescence intensity (MFI) was established. The actual PC4d cutoff value (12.755) with the specificity and sensitivity of PC4d at that cutoff (0.741 and 0.600, respectively) and the area under the ROC curve (AUC) (0.639) with the 95% confidence interval (95% CI; 0.459–0.819) are shown. The actual PC4d level cutoff was rounded to 13 MFI for simplicity.

We also analyzed the events that happened from 5 years before to 3 years after PC4d level measurement to expand on the previous analysis in the studies by Petri et al (8) and Gartshteyn et al (10). A total of 70 individuals (22 in the JH cohort, 43 in the COL cohort, and 5 in the BI cohort) experienced at least 1 event for a total of 134 events recorded in the medical records in this 8-year time period. For each subject, we analyzed the event closest to PC4d level measurement. These 70 events included 23 deep vein thromboses, 2 pulmonary embolisms, 1 portal vein thrombosis, 12 venous thromboses not specified, 12 cerebrovascular accidents, 6 myocardial infarctions, 4 gastrointestinal infarctions, and 10 arterial thromboses not specified. The median PC4d level in the 70 individuals with thrombotic history was 9.3 MFI (mean 40.18 MFI) compared to 4.7 MFI (mean 20.43 MFI) in the 348 individuals without history of thrombosis. The difference between PC4d levels in subjects with no events versus subjects with events (arterial or venous), subjects with venous events, and patients with arterial events was statistically significant (P = 0.0058, P = 0.0356, and P = 0.0066 by Mann-Whitney U test, respectively) (Figure 2). In addition, 30 of the 70 patients with thrombosis (42.9%) had a PC4d level of >13 MFI, while 265 of the 348 patients without thrombosis (76.1%) had a PC4d level of ≤13 MFI (OR 2.45 [95% CI 1.37–4.32]) (Fisher's test P = 0.0016). These data, combining the arterial and venous events that occurred in the 5 years before PC4d level measurement to 3 years after, confirm that PC4d is associated with history of thrombosis in SLE, as previously shown when the analysis was conducted on the historic thrombosis in the JH and COL cohorts (8,10).

DISCUSSION

SLE is a complex autoimmune disease linked to classical complement pathway activation, consumption of complement proteins C3 and C4, and production of C4d split fragments covalently bound to hematopoietic cells, including erythrocytes,

Table 2. Triple aPL positivity in patients with or without thrombosis stratified by PC4d ≤13 or >13 MFI*

	Not triple aPL positive	Triple aPL positive	Total
Subjects with thrombosis after PC4d measurement			
PC4d ≤13 MFI	8	0	8
PC4d >13 MFI	6	1	7
Subjects without thrombosis			
PC4d ≤13 MFI	273	24	297
PC4d >13 MFI	83	21	104

* aPL = antiphospholipid antibodies; MFI = mean fluorescence intensity; PC4d = platelet-bound C4d. Patients were considered triple aPL positive if positive for at least 1 anti-cardiolipin antibodies isotype, at least 1 anti-β2 glycoprotein 1 antibodies isotype, and at least 1 anti-phosphatidylserine/prothrombin complex antibodies isotype.

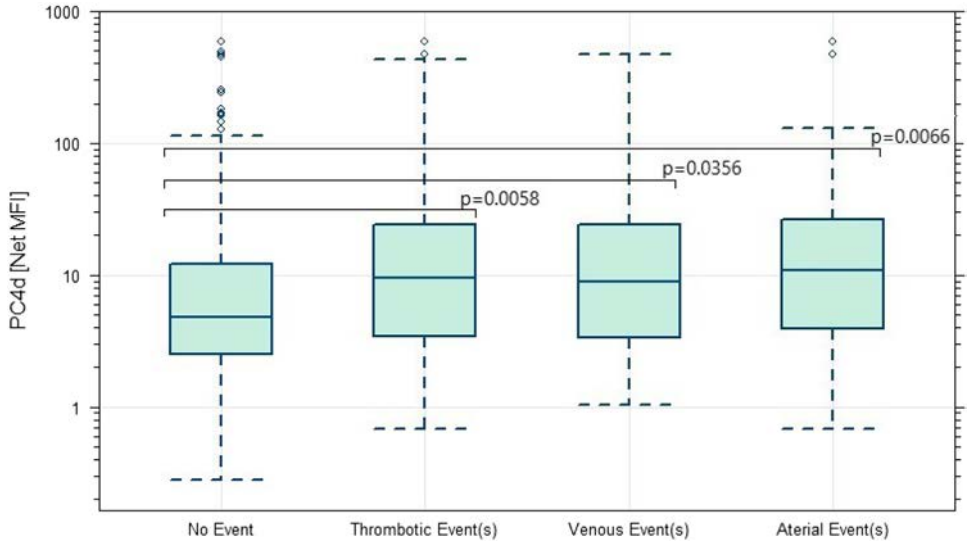

Figure 2. Platelet-bound complement activation product C4d (PC4d) values in patients with and without thrombotic events (all, arterial, and venous) before or after PC4d measurement. Box plot of PC4d levels (expressed as net mean fluorescence intensity) in the 348 subjects who did not experience (no event) and in the 70 subjects who experienced any thrombosis (thrombotic event[s]), venous thrombosis (venous event[s]), or arterial thrombosis (arterial event[s]) before or after PC4d measurement. For subjects without events, the PC4d value plotted is the baseline value; for subjects with events, the PC4d value plotted is the value closest to the event. Lines inside the boxes represent the median, each box represents the interquartile range (25^{th} and 75^{th} percentile), whiskers represent 1.5 times the interquartile range, and circles indicate outliers.

B lymphocytes, and platelets. We and others (8–11) have shown that PC4d is associated with a history of thrombosis in SLE. In particular, PC4d, low serum complement proteins, LAC, and anti-PS/PT IgG were associated with history of thrombosis in the JH cohort (8). In the COL cohort, we showed that PC4d was associated with a history of thrombotic events and that platelets with a PC4d level of ≥10 MFI were hyperactive (10).

To increase the number of thrombotic events for analysis, we combined the 2 cohorts from JH and COL with a third cohort from BI and evaluated the association of PC4d level with history of thrombosis and, more importantly, the ability of PC4d level to predict occurrence of future thrombosis events. This study is the first to show that PC4d level can predict future thrombosis events in SLE by collecting clinical data up to 3 years after PC4d level measurement. Although epidemiologic studies to measure incidence of thrombosis ideally have a longer follow-up (1,3,18,19), we chose a relatively short period of time post-PC4d because the utility of a biomarker to predict thrombosis and guide patient management is especially relevant for the immediate future. Although 418 patients in this study were enrolled at 3 different lupus centers under different protocols, patient characteristics, including sex, age, proportion of patients younger than 65 years of age, disease activity, use of medications, and positivity rate for aPL, were similar.

PC4d above the optimum cutoff of greater than 13 MFI determined by ROC curve analysis was statistically significant at predicting future arterial thrombosis ($P = 0.026$) with sensitivity of 60%, specificity of 74%, and diagnostics OR of 4.30. In addition, there was a statistically significant difference for time to an arterial event in patients with a PC4d level of >13 MFI as determined by log-rank test (hazard ratio 4.34).

Data did not reach statistical significance, neither in the entire population nor in the subgroup younger than 65 years of age, when arterial and venous thrombosis were analyzed together. Lack of statistical significance at predicting arterial and venous thrombosis may be due to the small number of prospective events observed in these cohorts, as PC4d has been shown to be associated not only with arterial but also with venous thrombotic events in SLE (8,9). EC4d and BC4d, on the other hand, were not associated with history of thrombotic events and did not predict either arterial or venous events in our study.

Although ability to evaluate risk of future thrombosis is important, equally important is being reassured that thrombosis is unlikely. A PC4d level of ≤13 MFI had excellent negative predictive value for future arterial thrombosis and any thrombosis (99% and 97%, respectively), indicating very high probability of not having a thrombosis within 3 years if PC4d level was ≤13.

We previously showed that aPL and PC4d level are weakly correlated, suggesting an additive value (16). We now demonstrate that the increased risk of thrombosis conferred by a PC4d level of >13 MFI, as well as the protective effects of a PC4d level of ≤13 MFI, were independent of aPL positivity. In fact, only 1 of 15 patients who had a thrombosis and presented with a PC4d level of >13 MFI was triple aPL positive at baseline, and 6 of the 14 patients who were not triple positive had a PC4d level of >13 MFI. On the other hand, 8.1% of patients with a PC4d level of ≤13 MFI did not

experience a thrombosis even if they were triple aPL positive, indicating that the probability of not having a thrombosis if PC4d level was ≤13 MFI holds true for the aPL positive population and presumed at higher risk. We also analyzed the thrombotic events pre- and post-PC4d and confirmed previous data (8,10) that PC4d level is associated with history of thrombosis in SLE ($P = 0.0016$).

This study has several limitations that warrant acknowledgment. Although we observed 19 venous and arterial events in the 3 years post-PC4d, corresponding to an incidence of 15 events per 1,000 patient-years, the sample size was small. This may explain the modest effect of PC4d level in predicting occurrence of venous and arterial thrombosis combined. In addition, we were unable to evaluate the contribution of traditional risk factors of thrombosis, such as hypertension, diabetes mellitus, hypercholesterolemia, or smoking. The contribution of LAC to thrombosis could not be evaluated, given that LAC was performed only in 1 of the 3 study cohorts. However, the importance of LAC has been demonstrated previously (2,5–7).

PC4d, as a biomarker for increased cardiovascular disease risk, may be useful to identify at-risk patients with SLE who would benefit from further screening with cardiac stress test or coronary calcium score computerized tomography evaluations. However, additional prospective studies are needed to validate our findings as well as to evaluate PC4d level as an actionable biomarker for risk-reducing interventions, such as the use of aspirin or lipid-lowering therapeutics.

In conclusion, PC4d level can predict arterial thrombosis in the 3 years post measurement and can also help to rule out risk of thrombosis. As PC4d level may help evaluate the risk of future thrombosis events in SLE, it may inform patient monitoring, control of other thrombosis risk factors, and lead to institution of appropriate pharmacologic preventive treatment.

ACKNOWLEDGMENTS

The authors thank the patients who participated in the study, Suzanne Krishfield for contributing to patient enrollment and data collection, Wei Tang for contributing to data collection, and Touba Warsi for contributing to data analysis. The authors also thank Exagen's laboratory personnel for technical assistance.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Gartshteyn, Conklin, Petri, Kyttaris, Kammesheidt, Askanase, Alexander.

Acquisition of data. Gartshteyn, Conklin, Petri, Kyttaris, Goldman, Askanase, Alexander.

Analysis and interpretation of data. Gartshteyn, Conklin, Petri, Kyttaris, Kammesheidt, Askanase, Alexander.

ROLE OF THE STUDY SPONSOR

Exagen facilitated the study design, provided writing assistance for the manuscript, and reviewed and approved the manuscript prior to submission. The authors independently collected the data, interpreted the results, and had the final decision to submit the manuscript for publication. Publication of this article was not contingent upon approval by Exagen.

ADDITIONAL DISCLOSURES

Authors Conklin, Kammesheidt, and Vezza Alexander are former employees and own stock or stock options in Exagen.

REFERENCES

1. Lu X, Wang YH, Zhang J, et al. Patients with systemic lupus erythematosus face a high risk of cardiovascular disease: a systematic review and meta-analysis. Int Immunopharmacol 2021;94:107466.
2. Somers E, Magder LS, Petri M. Antiphospholipid antibodies and incidence of venous thrombosis in a cohort of patients with systemic lupus erythematosus. J Rheumatol 2002;29:2531–6.
3. Ballocca F, D'Ascenzo F, Moretti C, et al. Predictors of cardiovascular events in patients with systemic lupus erythematosus (SLE): a systematic review and meta-analysis. Eur J Prev Cardiol 2015;22:1435–41.
4. Unlu O, Zuily S, Erkan D. The clinical significance of antiphospholipid antibodies in systemic lupus erythematosus. Eur J Rheumatol 2016; 3:75–84.
5. Galli M, Luciani D, Bertolini G, et al. Lupus anticoagulants are stronger risk factors for thrombosis than anticardiolipin antibodies in the antiphospholipid syndrome: a systematic review of the literature. Blood 2003;101:1827–32.
6. Galli M, Luciani D, Bertolini G, et al. Anti-β2-glycoprotein I, antiprothrombin antibodies, and the risk of thrombosis in the antiphospholipid syndrome. Blood 2003;102:2717–23.
7. Demir S, Li J, Magder LS, et al. Antiphospholipid patterns predict risk of thrombosis in systemic lupus erythematosus. Rheumatology (Oxford) 2021;60:3770–7.
8. Petri MA, Conklin J, O'Malley T, et al. Platelet-bound C4d, low C3 and lupus anticoagulant associate with thrombosis in SLE. Lupus Sci Med 2019;6:e000318.
9. Svenungsson E, Gustafsson JT, Grosso G, et al. Complement deposition, C4d, on platelets is associated with vascular events in systemic lupus erythematosus. Rheumatology (Oxford) 2020;59:3264–74.
10. Gartshteyn Y, Mor A, Shimbo D, et al. Platelet bound complement split product (PC4d) is a marker of platelet activation and arterial vascular events in systemic lupus erythematosus. Clin Immunol 2021; 228:108755.
11. Kao AH, McBurney CA, Sattar A, et al. Relation of platelet C4d with all-cause mortality and ischemic stroke in patients with systemic lupus erythematosus. Transl Stroke Res 2014;5:510–8.
12. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 1997;40:1725.
13. Petri M, Orbai AM, Alarcon GS, et al. Derivation and validation of the Systemic Lupus International Collaborating Clinics classification criteria for systemic lupus erythematosus. Arthritis Rheum 2012;64: 2677–86.
14. Petri M, Kim MY, Kalunian KC, et al. Combined oral contraceptives in women with systemic lupus erythematosus. N Engl J Med 2005;353: 2550–8.
15. Gladman D, Ibanez D, Urowitz M. Systemic Lupus Erythematosus Disease Activity Index 2000. J Rheumatol 2002;29:288–91.

16. Sciascia S, Bloch R, Malley TO, et al. Antiphospholipid antibodies are persistently positive at high titers. Additive value of platelet-bound C4d. Front Immunol 2022;13:949919.
17. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 2011;12:77.
18. Urowitz MB, Gladman DD, Farewell V, et al. Accrual of atherosclerotic vascular events in a multicenter inception systemic lupus erythematosus cohort. Arthritis Rheumatol 2020;72:1734–40.
19. Urowitz MB, Su J, Gladman DD. Atherosclerotic vascular events in systemic lupus erythematosus: an evolving story. J Rheumatol 2020; 47:66–71.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 75, No. 10, October 2023, pp 2096–2106
DOI 10.1002/acr.25093

Sex and Racial Differences in Systemic Lupus Erythematosus Among US Adults in the All of Us Research Program

Christopher Rice,[1] Deepak Nag Ayyala,[2] Hong Shi,[2] Adria Madera-Acosta,[3] Stephen Bell,[3] Anam Qureshi,[3] Laura D. Carbone,[3] Steven S. Coughlin,[2] and Rachel E. Elam[3]

Objective. Men with systemic lupus erythematosus (SLE) are an understudied population. The present study characterized differences between men and women with SLE.

Methods. We examined cross-sectionally participants with SLE in the All of Us Research Program, a US cohort with a participant survey at enrollment (May 2018 to June 2022) and linked electronic health record (EHR) data. We described and compared characteristics of men and women with SLE encompassing disease manifestations and prescribed medications from EHR data and socioeconomic factors, including health literacy and health care access and utilization, from surveys. We reported racial variations stratified by sex.

Results. Of 1,462 participants with SLE, 126 (9%) were male. Men reported lower educational attainment and less fatigue than women. Myocardial infarction was significantly more common in men. Men had significantly less confidence in completing medical forms than women and exhibited a trend toward requiring more help in reading health-related materials. Barriers to health care access and utilization were common in both men and women (40% versus 47%, respectively, reporting some reason for delay in care; $P = 0.35$). Women of race other than Black or African American or White more often reported delaying care due to cultural differences between patient and provider.

Conclusion. Our study demonstrated major clinical and health literacy differences in men and women with SLE. Socioeconomic factors were significant barriers to health care in both sexes. Our study suggests men have disproportionately poorer health literacy, which may exacerbate preexisting disparities. Further large prospective studies, focusing on recruiting men, are needed to better characterize racial differences in men with SLE.

INTRODUCTION

Systemic lupus erythematosus (SLE) is ~9 times more common in women than men in the US (1). Sex-specific differences in sex hormones, toll-like receptor expression, and microRNA profiles may play a role in the sex-dependent susceptibility to SLE (2). SLE is still a substantial burden in men, with an estimated prevalence of 14.6 per 100,000 person-years (1) and perhaps a more aggressive clinical course (3). In men and women, SLE disproportionately affects certain racial and ethnic underserved populations (1). Black or African American and Hispanic or Latino persons with SLE have higher disease activity, more disease-related complications, and excess mortality compared with non-Hispanic or Latino White persons (4).

Several studies report sex differences in SLE disease manifestations, but findings are inconsistent across cohorts, and sample sizes of men are often small (5–10). Even less understood is the role of race in the clinical phenotype of SLE in men. Rare reports have identified racial differences in disease manifestations in men with SLE (5,11–15), but differences are incompletely characterized. The limited data on how socioeconomic factors differ between men and women

The All of Us Research Program is supported by the NIH, Office of the Director (Regional Medical Centers grants 1OT2-OD-026549, 1OT2-OD-026554, 1OT2-OD-026557, 1OT2-OD-026556, 1OT2-OD-026550, 1OT2-OD-026552, 1OT2-OD-026553, 1OT2-OD-026548, 1OT2-OD-026551, and 1OT2-OD-026555, and IAA grant AOD 16037), Federally Qualified Health Centers (grant HHSN 263201600085-U), Data and Research Center (grant 5U2C-OD-023196), Biobank (grant 1U24-OD-023121), The Participant Center (grant U24-OD-023176), Participant Technology Systems Center (grant 1U24-OD-023163), Communications and Engagement (grants 3OT2-OD-023205 and 3OT2-OD-023206), and Community Partners (grants 1OT2-OD-025277, 3OT2-OD-025315, 1OT2-OD-025337, and 1OT2-OD-025276).

[1]Christopher Rice, MD: Cone Health, Greensboro, North Carolina; [2]Deepak Nag Ayyala, PhD, Hong Shi, MD, Steven S. Coughlin, PhD, MPH: Augusta University, Augusta, Georgia; [3]Adria Madera-Acosta, MD, Stephen Bell, MD, Anam Qureshi, MD, Laura D. Carbone, MD, MS, Rachel E. Elam, MD, ScM: Augusta University and Charlie Norwood Veterans Affairs Medical Center, Augusta, Georgia.

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.25093&file=acr25093-sup-0001-Disclosureform.pdf.

Address correspondence via email to Rachel E. Elam, MD, ScM, at relam@augusta.edu.

Submitted for publication July 5, 2022; accepted in revised form January 24, 2023.

SIGNIFICANCE & INNOVATIONS

- Male participants with systemic lupus erythematosus (SLE) in this US national cohort exhibited a trend toward lower educational attainment and significantly less confidence in completing medical forms than female participants, suggesting disproportionately poorer health literacy in men with SLE.
- Socioeconomic barriers to health care access and utilization are common and numerous in both men and women with SLE in this US national cohort.
- Barriers to health care access and utilization may differ by race in women with SLE, with women of race other than Black or African American or White more frequently reporting delayed care due to cultural differences between patient and provider.

with SLE is conflicting, but education level and income may vary by sex (3,5).

Health care cost and access disparities have been reported to disproportionately impact persons with SLE (16). Men with SLE have reported more perceived difficulty in accessing health care than women (3), and this access disparity is corroborated by literature demonstrating less outpatient clinic visits (including rheumatology subspecialty) in men (17,18). Sex differences in health care utilization may be altered by race as Black or African American women with SLE were found to be significantly less likely to be referred to a rheumatologist compared to their White male counterparts (18). However, further data on sex differences in health care access and utilization by race is needed.

The purpose of this study was to characterize differences in SLE clinical manifestations, prescribed medications, and socioeconomic determinants of health, including health literacy and barriers to health care access and utilization, by sex and race among men with SLE. To this end, we utilized data from the All of Us Research Program, a US national, deidentified data repository consisting of both patient survey and linked electronic health record (EHR) data.

PATIENTS AND METHODS

Study cohort. The National Institutes of Health (NIH) All of Us Research Program (hereafter referred to as the All of Us) is an ongoing longitudinal study aimed at recruiting 1 million volunteers representative of the US population to contribute data to the All of Us data repository with the goal of accelerating biomedical research and improving health. The All of Us study procedures have been previously described (19). In brief, adults ages 18 years and older who reside within the US or a US territory are eligible to participate. The All of Us program initiated enrollment in May 2018 and participants enroll and consent to participate either via the All of Us website (https://joinallofus.org) or a smart-phone application. Volunteer participants are invited to complete several health surveys, composed of validated instruments or questions, when appropriate (20). Participants may opt to provide authorization to share EHR data, in which case survey data is linked to billing codes, medication history, laboratory results, and encounter records from EHR data from any of 60 health care provider organizations in the All of Us program's network and also (in a subset) from other providers using Fast Healthcare Interoperability Resources–based connections. Each participant is also eligible to undergo an initial evaluation for physical measurements. The All of Us protocol was approved by the Institutional Review Board of the All of Us Research Program, which follows the regulations and guidance of the NIH Office for Human Research Protections.

Participants with SLE were identified from the All of Us database, version 6. This version was released in June 2022 and accessed in November 2022. We restricted our cohort to participants with linked EHR data, as the accuracy of self-reported SLE diagnosis has been demonstrated to be poor (21). We included a participant in the All of Us database as having SLE if they had ≥3 SLE diagnosis codes on separate occurrences and had ever been prescribed an antimalarial medication (including hydroxychloroquine, chloroquine, or quinacrine). SLE diagnosis codes accepted included International Classification of Diseases, ninth revision (ICD-9) code 710.0; ICD-10 codes M32.1, M32.8, or M32.9; or Standardized Nomenclature of Medicine (SNOMED) code 55464009. We excluded participants with a medical billing diagnosis code for dermatomyositis (ICD-9 code 710.3, ICD-10 codes M33.0, M33.1) or systemic sclerosis (ICD-9 code 710.1, ICD-10 code M34). This algorithm has been validated to have 88–91% positive predictive value in correctly identifying participants with SLE from EHR databases using ICD-9 billing codes (22).

Participant characteristics. Participants were categorized as male or female based on self-reported biological sex assigned at birth. Other sociodemographic data was self-reported in The Basics enrollment survey and included age at consent for study participation, race (Black or African American, Other race, or White), ethnicity (Hispanic or Latino versus any other response), highest level of educational attainment, health insurance provider, employment status, and annual household income category. Other race is a composite of responses "another single population," "Asian," "more than one population," "none indicated," or "none of these."

The enrollment Lifestyle questionnaire captured self-reported cigarette smoking status, current alcohol use, and current and ever marijuana use. Current cigarette smoking was defined by self-report of smoking (either "some days" or "every day") of cigarettes. Former cigarette smokers reported having ever smoked at least 100 cigarettes in their lifetime, but now smoking cigarettes "not at all." Current heavy alcohol use was defined as present if the participant self-reported drinking a drink containing alcohol 2 times per week or more and self-reported ≥3 alcoholic drinks

on a typical day when they drink. Current marijuana use described any self-reported marijuana use (including cannabis, pot, grass, hash, weed, etc.) in the past 3 months.

An enrollment survey on overall health ascertained general health metrics of average pain level, general health perception, and fatigue level. Average pain level was defined by self-reported average pain over the past 7 days on a scale of 0 (no pain) to 10 (worst pain imaginable). We categorized these scores as mild (0–3), moderate (4–7), and severe (8–10). Participants were asked, "In general, would you say your health is: poor, fair, good, very good, or excellent?" Participants rated their fatigue over the past 7 days as none, mild, moderate, severe, or very severe.

Physical measurements were obtained per standardized protocol by a trained program staff member following enrollment with patient consent or from linked EHR data (19). Physical metrics included height and weight, from which body mass index was calculated.

SLE disease manifestations and prescribed medications. Organ-specific SLE disease involvement was identified for each participant from EHR data using ICD-9, ICD-10, SNOMED, and/or Current Procedural Terminology, 4th edition codes (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25093). Cardiovascular disease (including myocardial infarction, coronary artery disease with or without angina, limb claudication, congestive heart failure, peripheral arterial disease, and stroke), lupus nephritis, end-stage renal disease, antiphospholipid syndrome, lupus pericarditis, lupus lung disease, and Raynaud's phenomenon were evaluated because sex-specific differences have previously been reported for these features (6,23–28).

Medications commonly prescribed in the treatment of SLE were ascertained from prescription drug history in linked EHR data, including mycophenolate mofetil or mycophenolic acid, azathioprine, cyclophosphamide (oral or intravenous), methotrexate, tacrolimus, rituximab, and belimumab (29). Participants were considered as having had a prescribed medication if it was currently or ever previously prescribed to that participant. There were no participants in All of Us, version 6, prescribed anifrolumab or voclosporin.

Health literacy. The Overall Health enrollment survey includes the following 3 questions modified from the Brief Health Literacy Screen (20): 1) how confident are you filling out medical forms by yourself: extremely, quite a bit, somewhat, a little bit, not at all, prefer not to answer? 2) how often do you have someone help you read health-related materials: always, often, sometimes, occasionally, never, prefer not to answer? and 3) how often do you have problems learning about your medical condition because of difficulty understanding written information: always, often, sometimes, occasionally, never, prefer not to answer? Factor analysis showed these questions to be the 3 most strongly explanatory factors of health literacy among the survey items, with coefficients of –0.586 (question 1), 0.755 (question 2), and 0.748 (question 3) (20). We defined lack of confidence in completing medical forms as present if a participant responded somewhat, a little bit, or not at all to question 1. We defined requiring help reading health-related materials as present if the participant responded always, often, or sometimes to question 2. We defined having difficulty understanding written health information as present if the participant responded always, often, or sometimes to question 3.

Health care access and utilization. Domains of health care access and utilization were assessed on the Health Care Access & Utilization survey administered after enrollment. On this survey, participants were queried if they had delayed getting care in the past 12 months for any of the following reasons: 1) could not afford the copay, 2) insurance deductible was too high or could not afford the deductible, 3) had to pay out of pocket for some or all of the procedure, 4) could not get time off of work, 5) could not get child care, 6) could not get elderly care, 7) did not have transportation, 8) live in a rural area where distance to the health care provider is too far, or 9) were nervous to see a health care provider. We collapsed these delays into meaningful categories for analysis: delay due to affordability (reasons 1–3); delay due to time constraint (reasons 4–6); delay due to transportation (reasons 7–8); and delay due to nervousness to see a health care provider (reason 9).

Any medication challenge due to a cost barrier was defined as presence of any of the following actions reported by the participants to save money in the past 12 months: skipped medication doses, took less medicine, delayed filling a prescription, bought prescription medications from another country, requested a lower cost medication from their doctor, or used alternative therapies. Participants were also asked "how often have you either delayed or not gone to see doctors or health care providers because they were different from you in any of these ways (race, ethnicity, religion, beliefs, or native language)? Always, most of the time, some of the time, none of the time, don't know, did not answer." We defined ever having delayed care because their health care provider was different from them if a participant responded always, most of the time, or some of the time on this item.

Statistical analyses. Descriptive statistics were calculated as number (%) by sex and by race stratified by sex. According to the All of Us Research Program data and statistics dissemination policy, cell values and aggregate statistics that correspond to 1 to 20 participants are obscured to protect participant privacy (reported as ≤20 in tables). Inferential statistics compared sociodemographic characteristics, disease manifestations, prescribed medications, health literacy items, and barriers to health care access and utilization by chi-square test with Yates' continuity correction. All analyses were performed using the All of Us Researcher Workbench (version 6) and R environment for

statistical computing (30). An alpha level of 0.05 was prespecified. A Bonferroni -adjusted alpha level was also considered and reported, adjusting for the 25 simultaneous chi-square tests (adjusted alpha level of 0.002). *P* values that, in combination with presented summary statistics, would allow the reader to deduce a participant count of 1 to 20 were reported as not applicable.

RESULTS

We identified 1,462 participants with SLE, including 126 men (9% of the total cohort) (Figure 1). Baseline characteristics of the study cohort by sex are summarized in Table 1.

Sex differences in sociodemographic and lifestyle characteristics. The age distribution of men and women in the SLE cohort was similar, with most participants of either sex being ages <65 years at consent for participation in the study. Among men, White race was the most commonly represented (41%), followed by Other race (36%), and then Black or African American race (24%). Twenty-eight percent of men identified as Hispanic or Latino. Among women, the ethnic and racial distribution was similar: 31% Black or African American, 28% Other race, 41% White, and 24% Hispanic or Latino.

Most men (73%) and women (73%) were overweight or obese. Men showed lower educational attainment ($P = 0.02$). The majority of participants of both sexes were not employed (63–65%) and exclusively on government-issued insurance (60–62%). Annual household income <$35,000 was common in both men (45%) and women (54%). Men were more likely to be current or former cigarette smokers ($P = 0.02$) and were less likely to report moderate-to-severe fatigue ($P = 0.01$).

Sex differences in SLE disease manifestations and prescribed medications. Organ-specific SLE disease manifestations and prescribed medications by sex are summarized in Table 2. Men were more likely than women to have any cardiovascular event, and in particular, myocardial infarction and coronary artery disease with or without angina. The difference in cross-sectional incidence of myocardial infarction in men (18%) compared to women (8%) remained statistically significant with Bonferroni correction of alpha significance level ($P < 0.002$). There was a trend toward more men having lupus nephritis (25% versus 18%) and antiphospholipid syndrome (19% versus 13%) compared to women, but these differences did not reach statistical significance. Medication prescription rates of mycophenolate mofetil or mycophenolic acid and methotrexate did not differ by sex. The remainder of comparisons by sex for prescribed medications were limited due to the small sample size in men.

Racial differences in SLE disease manifestations and prescribed medications. Organ-specific SLE disease manifestations and prescribed medications by race, stratified by sex, are summarized in Supplementary Table 2 (available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25093). White women reported more Raynaud's phenomenon (11%) compared to Black or African American (4%) or Other race (5%). Otherwise, findings were similar across racial groups studied among women.

Sex differences in health literacy. Sex differences in the 3 domains of health literacy are shown in Table 3. A significantly larger proportion of men reported a lack of confidence completing medical forms than women (23% in men versus 12% in women; $P = 0.0017$). This difference remained statistically significant with Bonferroni correction of alpha significance level ($P < 0.002$). Men also more frequently reported requiring help reading health-related materials (30% of men versus 19% of women; $P = 0.009$). The proportion of men and women having difficulty understanding written health information was similar.

Racial differences in health literacy. Racial differences, stratified by sex, in the 3 domains of health literacy are shown in Table 3. In women, participants of Black or African American and Other race more frequently reported a lack of confidence in completing medical forms (16–18% versus 6%), requiring help reading health-related materials (21–28% versus 12%), and having difficulty understanding written health information (21–24% versus 10%) than participants of White race.

Sex differences in health care access and utilization. Only a minority of participants completed at least 1 item on the Health Care Access & Utilization survey (53 men and 634 women). Sex differences in the evaluated domains of health care access and utilization are shown in Table 4. Barriers to health care access leading to delays in care were common in both men and women, with 40% of men and 47% of women reporting at least some reason for delay in care in the past 12 months, and these rates did not differ by sex ($P = 0.35$). A sizeable minority of women reported delays in care due to affordability (23%), time constraints (18%), transportation (17%), and nervousness to see a health care provider (17%).

Racial differences in health care access and utilization. Racial differences, stratified by sex, in the evaluated domains of health care access and utilization are shown in Table 5. In women, delays in care in the past 12 months for any reason were reported more often by women of Other race (57%) than Black or African American (45%) or White (44%) race, however, delays were common across all races. Women of Other race reported more delays due to time constraints (24% versus 15–16%) and transportation (23% versus 16%) than women of

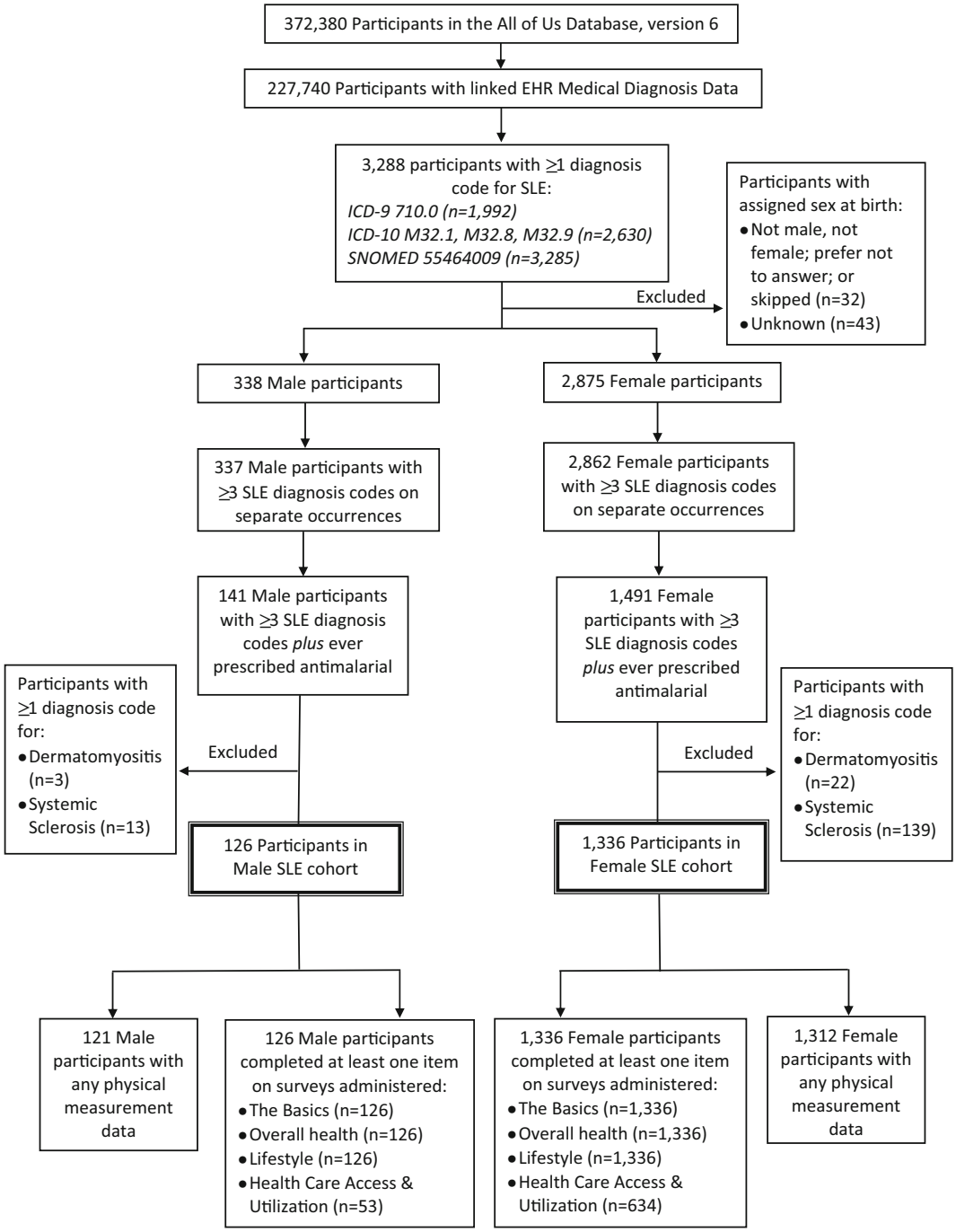

Figure 1. Study flow chart. EHR = electronic health record; ICD-9 = International Classification of Diseases, ninth revision; SLE = systemic lupus erythematosus; SNOMED = Standardized Nomenclature of Medicine. * ICD-9 code 710.3, ICD-10 codes M33.0 and M33.1 = dermatomyositis; ICD-9 code 710.1, ICD-10 code M34 = systemic sclerosis. Antimalarial includes hydroxychloroquine, chloroquine, or quinacrine.

Black or African American or White race. Women of Other race also more often reported ever delaying care because their health care provider was different from them in race, ethnicity, sex, religion, beliefs, or native language (25% versus 14–16%) compared to women of Black race or White race. In contrast, women of White race reported more medication challenges due to cost barriers (43%) compared with Black or African American (34%) or women of Other race (34%).

DISCUSSION

A total of 9% of participants with SLE in our cross-sectional study were male, consistent with the US SLE male:female prevalence ratio (1). Our study has one the of largest cohort of men among studies systematically examining sex differences in SLE (5–10). We observed differences between men and women in a number of clinically relevant features,

Table 1. Baseline characteristics of study cohort by sex*

Population characteristics	Men (n = 126)	Women (n = 1,336)	*P*
Age (at consent for study participation), years			
18–44	41 (33)	512 (38)	0.08
45–64	57 (45)	625 (47)	
≥65	28 (22)	199 (15)	
Race			0.13
Black or African American	29 (24)	403 (31)	
Other†	44 (36)	371 (28)	
White	50 (41)	538 (41)	
Hispanic or Latino ethnicity	35 (28)	315 (24)	0.34
BMI category (kg/m^2)			1.00
Underweight or normal (<25)	33 (27)	354 (27)	
Overweight or obese (≥25)	89 (73)	947 (73)	
Highest educational attainment‡			0.02§
At most high school completion	49 (39)	373 (28)	
At least 1 year of college	77 (61)	941 (72)	
Health insurance status¶			0.73
Exclusively government insurance	56 (62)	608 (60)	
Exclusively non-government insurance	34 (38)	410 (40)	
Employment status			0.76
Employed for wages or self-employed	41 (35)	461 (37)	
Not currently employed for wages	76 (65)	788 (63)	
Annual household income category			0.09
<$35,000	44 (45)	566 (54)	
>$35,000	54 (55)	477 (46)	
Cigarette smoking status#			0.02§
Current smoker	24 (20)	153 (12)	
Former smoker	29 (24)	253 (20)	
Never smoker	70 (57)	874 (68)	
Current heavy alcohol use**	≤20 (≤16)	39 (29)	N/A
Ever marijuana use	64 (51)	560 (42)	0.07
Current marijuana use††	≤20 (≤16)	126 (9)	N/A
Pain level‡‡			0.06
Mild	48 (41)	382 (30)	
Moderate	46 (39)	596 (48)	
Severe	23 (20)	276 (22)	
General health perception§§			0.96
Good to excellent	53 (43)	555 (42)	
Fair to poor	71 (57)	765 (58)	
Fatigue level¶¶			0.01§
Mild or less	55 (45)	426 (32)	
Moderate to very severe	68 (55)	895 (68)	

* Values are the number (%) unless indicated otherwise. According to the All of Us Research Program data and statistics dissemination policy, cell values and aggregate statistics that correspond to 1 to 20 participants are obscured to protect participant privacy (in these cases, *P* values are designated as not applicable [N/A]). BMI = body mass index.
† Participants who self-reported as another single population, Asian, more than 1 population, none indicated, or none of these.
‡ At most high school completion included less than a high school degree or equivalent and highest grade completed as 12 or General Educational Development equivalent; at least 1 year of college included persons who completed 1–3 years of college, college graduates, or those having another advanced degree.
§ Significant at alpha level 0.05 (no *P* values remained statistically significant at Bonferroni -adjusted alpha level of 0.002 [25 simultaneous tests]).
¶ Government insurance included Indian, Medicaid, Medicare, military, and Veterans Affairs health insurance categories; non-government insurance included employer or union, purchased, or other health plan health insurance categories.
# Current smoker defined by self-report of smoking (either some days or every day) of cigarettes; former cigarette smoker defined by self-report of having ever smoked at least 100 cigarettes in their lifetime but now smoking cigarettes not at all.
** Defined by self-report of drinking a drink containing alcohol ≥2 times per week and self-report of ≥3 alcoholic drinks on a typical day when drinking (an alcoholic drink included a bottle of beer, a glass of wine or a wine cooler, a shot of liquor, or a mixed drink that included liquor).
†† Defined by self-report of marijuana use in the past 3 months.
‡‡ Defined by self-reported average pain over the past 7 days on a scale of 0 to 10, with 0 being no pain and 10 being worst pain imaginable; mild corresponded to a score of 0–3; moderate a score of 4–7; and severe a score of 8–10.
§§ Defined by self-reported response to survey query "In general, would you say your health is:" rated as fair to poor (fair, poor) versus good to excellent (good, very good, or excellent).
¶¶ Defined by self-reported fatigue over the past 7 days rated as mild or less (mild, none) versus moderate to very severe (moderate, severe, or very severe).

Table 2. Systemic lupus erythematosus disease manifestations and prescribed medications by sex*

	Men (n = 126)	Women (n = 1,336)	*P*
Cardiovascular disease manifestations			
Any cardiovascular event†	58 (46)	442 (33)	0.005‡
Myocardial infarction	22 (18)	102 (8)	<0.001§
Coronary artery disease with or without angina	41 (33)	292 (22)	0.009‡
Limb claudication	≤20 (≤16)	≤20 (≤1.5)	N/A
Congestive heart failure	25 (20)	188 (14)	0.10
Peripheral arterial disease	≤20 (≤16)	47 (4)	N/A
Stroke	≤20 (≤16)	77 (6)	N/A
Other organ-specific manifestations			
Lupus nephritis	31 (25)	235 (18)	0.07
End-stage renal disease	≤20 (≤16)	70 (5)	N/A
Antiphospholipid syndrome	24 (19)	172 (13)	0.07
Lupus pericarditis	≤20 (≤16)	56 (4)	N/A
Lupus lung disease	≤20 (≤16)	47 (4)	N/A
Raynaud's phenomenon	≤20 (≤16)	241 (18)	N/A
SLE medication ever prescribed¶			
Mycophenolate mofetil or mycophenolic acid	40 (32)	328 (25)	0.09
Azathioprine	≤20 (≤16)	253 (19)	N/A
Cyclophosphamide (oral or intravenous)	≤20 (≤16)	45 (3)	N/A
Methotrexate	29 (23)	324 (24)	0.84
Tacrolimus	≤20 (≤16)	157 (12)	N/A
Rituximab	≤20 (≤16)	84 (6)	N/A
Belimumab	≤20 (≤16)	108 (8)	N/A

* Values are the number (%) unless indicated otherwise. According to the All of Us Research Program data and statistics dissemination policy, cell values and aggregate statistics that correspond to 1–20 participants are obscured to protect participant privacy (in these cases, *P* values are designated as not applicable [N/A]). SLE = systemic lupus erythematosus.
† Also included undergoing percutaneous transluminal coronary angioplasty or undergoing coronary artery bypass grafting.
‡ Significant at alpha level 0.05.
§ Significant at alpha level 0.05; significant at Bonferroni -adjusted alpha level of 0.002 (25 simultaneous tests).
¶ Defined as ever being prescribed a specific medication or class if there was a past or current prescription documented in the electronic health record for that medication.

including sociodemographic characteristics, disease manifestations, and health literacy. Our findings contribute to the currently small body of literature reporting differences in socioeconomic factors and health care disparities between men and women in SLE (3,17,18).

Men with SLE in this US national cohort had lower educational attainment than women. Lower education in men with SLE has been reported in other cohorts (5) and may reflect national patterns of higher college enrollment rates in women, especially among Hispanic or Latina and Black or African American women (31). Lower educational attainment may contribute to the reported risk of more severe SLE in men (3). More years of education were protective from SLE-associated organ damage in men and women (3) and SLE-associated death in White men with SLE, but not Black or African American men (32). In our study, most men and women were unemployed, congruent with reported low employment rates in SLE (33). Men may be at particular risk for health-related work cessation due to higher disease activity and associated damage (3,5,25,34) and lower levels of educational attainment, both of which have been associated with work status (33). In another large SLE cohort, disability was 70% more common in men (5). Despite lower educational attainment, there was no sex difference in annual household income in our study, consistent with 1 prior report (5). In contrast, a study by Andrade et al found that women with SLE are more likely to experience poverty (3).

In our study, men were more likely to be current or former cigarette smokers than women. More tobacco smoking in men may contribute to increased SLE-related damage accrual in men compared to women (35).

Fatigue of at least moderate severity was reported in the majority of both men and women with SLE in our study, consistent with known high incidence of self-reported fatigue in SLE (36). Fatigue in SLE is multidimensional and has been associated with SLE disease activity and cumulative disease damage (37). Therefore, a trend toward less fatigue in men despite literature suggestive of a more aggressive SLE clinical course in men (3) is surprising and merits further investigation. It is possible that men perceive and report fatigue differently than women (38), as no difference in general health perception between men and women was observed.

In our study, men had more cross-sectional cardiovascular events, including coronary artery disease with or without angina, and statistically significantly more myocardial infarctions than women with SLE. These findings are in accord with previous studies (3,5,23,28,39) and may reflect the increased

Table 3. Health literacy by sex and by race, stratified by sex*

	Men	Women	*P*
Participants of all races	n = 125	n = 1,326	
Lack of confidence in completing medical forms†	28 (23)	162 (12)	0.0017‡
Requiring help reading health-related materials§	37 (30)	254 (19)	0.009¶
Having difficulty understanding written health information#	28 (23)	231 (18)	0.22
Black or African American	n = 29	n = 401	
Lack of confidence in completing medical forms†	≤20 (≤69)	64 (16)	N/A
Requiring help reading health-related materials§	≤20 (≤69)	86 (21)	N/A
Having difficulty understanding written health information#	≤20 (≤69)	85 (21)	N/A
Other race**	n = 44	n = 366	
Lack of confidence in completing medical forms†	≤20 (≤45)	66 (18)	N/A
Requiring help reading health-related materials§	≤20 (≤45)	102 (28)	N/A
Having difficulty understanding written health information#	≤20 (≤45)	88 (24)	N/A
White	n = 49	n = 535	
Lack of confidence in completing medical forms†	≤20 (≤41)	30 (6)	N/A
Requiring help reading health-related materials§	≤20 (≤41)	62 (12)	N/A
Having difficulty understanding written health information#	≤20 (≤41)	51 (10)	N/A

* Values are the number (%) unless indicated otherwise. According to the All of Us Research Program data and statistics dissemination policy, cell values and aggregate statistics that correspond to 1–20 participants are obscured to protect participant privacy (in these cases, *P* values are designated as not applicable [N/A]).
† Present if participant answered the survey question "How confident are you filling out medical forms by yourself?" with somewhat, a little bit, or not at all.
‡ Significant at alpha level 0.05; significant at Bonferroni -adjusted alpha level of 0.002 (25 simultaneous tests).
§ Present if participant answered the survey question "How often do you have someone help you read health-related materials?" with always, often, or sometimes.
¶ Significant at alpha level 0.05.
# Present if participant answered the survey question "How often do you have problems learning about your medical condition because of difficulty understanding written information?" with always, often or sometimes.
** Participants who self-reported as another single population, Asian, more than 1 population, none indicated, or none of these.

cardiovascular risk in men compared to women in the general population (40). Existing literature suggests that men are at increased risk for lupus nephritis (14) and antiphospholipid syndrome (3,23,39), and our findings corroborate these studies' results with a trend toward more men with lupus nephritis and antiphospholipid syndrome compared to women in our cohort.

Table 4. Domains of health care access and utilization by sex*

	Men (n = 53)	Women (n = 634)	*P*
Any reason for delay in care (past 12 months)†	21 (40)	300 (47)	0.35
Delay due to affordability‡	≤20 (≤38)	147 (23)	N/A
Delay due to time constraint§	≤20 (≤38)	112 (18)	N/A
Delay due to transportation¶	≤20 (≤38)	108 (17)	N/A
Delay due to nervousness to see health care provider	≤20 (≤38)	110 (17)	N/A
Any medication challenge due to cost barrier (past 12 months)#	≤20 (≤38)	249 (39)	N/A
Ever delayed care because their health care provider was different from them in race, ethnicity, sex, religion, beliefs, or native language**	≤20 (≤38)	106 (17)	N/A

* Values are the number (%) unless indicated otherwise. According to the All of Us Research Program data and statistics dissemination policy, cell values and aggregate statistics that correspond to 1 to 20 participants are obscured to protect participant privacy (in these cases, *P* values are designated as not applicable [N/A]).
† Present if participant reported any delay due to affordability, time constraint, and transportation, or a delay due to nervousness to see health care provider in the past 12 months.
‡ Present if participant reported they had delayed getting care in the past 12 months because they could not afford the copay, the insurance deductible was too high or could not afford the deductible, or they had to pay out of pocket for some or all of the procedure.
§ Present if participant reported they had delayed getting care in the past 12 months because they could not get time off work, could not get child care, or could not get elderly care.
¶ Present if participant reported they had delayed getting care in the past 12 months because they did not have transportation or they live in a rural area where distance to the health care provider is too far.
# Present if participant reported during the past 12 months they did any of the following to save money: skipped medication doses, took less medicine, delayed filling a prescription, bought prescription medications from another country, requested a lower cost medication from their doctor, or used alternative therapies.
** Present if participant self-reported they always, most of the time, or some of the time had delayed or not gone to see a health care provider because the provider was different from the participant in race, ethnicity, sex, religion, beliefs or native language.

Table 5. Domains of health care access and utilization by race, stratified by sex*

	Black or African American	Other race†	White
Male participants, no.	≤20	≤20	29
Any reason for delay in care (past 12 months)‡	≤20 (≤100)	≤20 (≤100)	≤20 (≤69)
Any medication challenge due to cost barrier (past 12 months)§	0 (0)	≤20 (≤100)	≤20 (≤69)
Ever delayed care because their health care provider was different from them in race, ethnicity, sex, religion, beliefs or native language¶	≤20 (≤100)	≤20 (≤100)	≤20 (≤69)
Female participants, no.	142	128	353
Any reason for delay in care (past 12 months)‡	64 (45)	73 (57)	157 (44)
Delay due to affordability#	34 (24)	28 (22)	85 (24)
Delay due to time constraint**	23 (16)	31 (24)	53 (15)
Delay due to transportation††	23 (16)	29 (23)	55 (16)
Delay due to nervousness to see health care provider	21 (15)	25 (20)	60 (17)
Any medication challenge due to cost barrier (past 12 months)§	48 (34)	43 (34)	151 (43)
Ever delayed care because their health care provider was different from them in race, ethnicity, sex, religion, beliefs, or native language¶	23 (16)	32 (25)	48 (14)

* Values are the number (%) unless indicated otherwise; according to the All of Us Research Program data and statistics dissemination policy, cell values and aggregate statistics that correspond to 1 to 20 participants are obscured to protect participant privacy (in these cases, *P* values are designated as not applicable [N/A]).
† Participants who self-reported as another single population, Asian, more than 1 population, none indicated, or none of these.
‡ Present if participant reported any of a delay due to affordability, time constraint, transportation, or due to nervousness to see health care provider in the past 12 months.
§ Present if participant reported during the past 12 months they did any of the following to save money: skipped medication doses, took less medicine, delayed filling a prescription, bought prescription medications from another country, requested a lower cost medication from their doctor, or used alternative therapies.
¶ Present if participant self-reported they always, most of the time, or some of the time had delayed or not gone to see a health care provider because the provider was different from the participant in race, ethnicity, sex, religion, beliefs, or native language.
# Present if participant reported they had delayed getting care in the past 12 months because they could not afford the copay, the insurance deductible was too high or could not afford the deductible, or they had to pay out of pocket for some or all of the procedure.
** Present if participant reported they had delayed getting care in the past 12 months because they could not get time off work, could not get child care, or could not get elderly care.
†† Present if participant reported they had delayed getting care in the past 12 months because they did not have transportation or live in a rural area where distance to the health care provider was too far.

Prevalence estimates of impaired health literacy in SLE vary widely in the literature, from 8.5–48% (41). Our study suggests men with SLE have less health literacy than women with SLE. A study by Hearth-Holmes et al demonstrated that Black or African American race and education level were associated with lower health literacy, but sex was not significantly associated (42). Men in our cohort also exhibited a trend toward lower educational attainment. The interplay of sex, education level, and health literacy in SLE merits further study. Lower health literacy, together with biological factors, may contribute to reported worse SLE prognosis in men, as low health literacy in SLE is associated with poorer patient-reported outcomes (43). In contrast, health literacy did not have a significant association with hydroxychloroquine adherence in a predominately Hispanic SLE cohort (44).

In women in our study, participants of Black or African American or Other race more frequently reported items consistent with lower health literacy than participants of White race. Similar to Hearth-Holmes et al (42), Maheswaranathan and colleagues also reported Black or African American race was associated with lower health literacy (45), but to our knowledge none have reported similar trends for races other than Black or African American compared to White race. Given the lack of specificity within our Other race category, further studies are needed to better elucidate the relationship of health literacy in SLE by race in demographic characteristics other than Black or African American or White persons.

Socioeconomic barriers to health care access and utilization are common and numerous in both men and women in our SLE cohort. Close to half of both men and women had some delay in their health care in the past 12 months. Men with SLE have been reported to have both more perceived difficulty in accessing health care (3) as well as objectively less clinical encounters for care, including less outpatient clinic (17) and, specifically, rheumatology subspecialty (18) visits compared to women. The potential impact of lower outpatient health care utilization in men with SLE remains unclear (17,32).

Reasons for delays in care in women with SLE were often multiple and could be attributed to a variety of causes, both directly related to affordability, but also only indirectly related to cost (time constraints or transportation issues). Furthermore, nervousness to see a health care provider was equally commonly reported as a barrier to care in women with SLE as other categories of reasons. Medication challenges due to cost barriers were also common in women in our cohort. Prescription medication challenges have been associated with more emergency department visits in persons with SLE (18).

Barriers to health care access and utilization may differ by race in women with SLE. Our study raises the possibility that women of race other than Black or African American or White may be disproportionately affected by delays in care. A study by Yazdany et al demonstrated that racial and ethnic minorities in a predominately female SLE cohort are less likely to receive recommended health care for SLE (46). In their study, all racial groups including Black or African American were compared to White race. Health care fragmentation has been shown to disproportionately affect Black or African American patients in another predominately female SLE cohort (47).

In our study, women of races other than Black or African American or White more commonly reported delayed care due to perceived differences between themselves and their provider with respect to race, ethnicity, sex, religion, beliefs, or native language. To our knowledge, this finding has not previously been reported. A study by Vina et al showed that, in a predominately female SLE cohort, African American individuals with SLE were less willing to receive cyclophosphamide if their SLE worsened compared to White patients, and this difference was mediated by less trust in the physician. However, this study did not include races other than African American or White (48). Further study is needed to refine our understanding of barriers to health care access and utilization in persons with SLE of all races.

Our study has limitations. Low response rates and low sample sizes in men limited the strength of our conclusions that can be drawn for this study on racial differences in SLE in men, and this remains an area in need of further research. Participation in the All of Us Research Program is voluntary and participants may not be representative of the US SLE population. Incomplete ascertainment of EHR data from consenting participants who receive care outside of the All of Us' network health care provider organizations may bias our results. Given its cross-sectional design, our study cannot determine whether clinical manifestations (i.e., myocardial infarction) occurred following SLE diagnosis and the relative attribution of these complications to SLE itself. Data identifying provider type (i.e., rheumatologist, nephrologist, etc.) in medication prescribing patterns and to whom delays in care refer was not included in this study. We cannot provide any information on causal relationships in features explored. The number of participants who identified within our Other race category was large, and likely heterogeneous in racial backgrounds. Finally, in our data collection of barriers to access to care in SLE, barriers were nonweighted, so priority of those variables could not be established.

Our study demonstrated major clinical and health literacy differences between men and women with SLE. Socioeconomic factors were significant barriers to health care access and utilization in both sexes. However, men reported poorer health literacy in our study compared to women with SLE, which may exacerbate preexisting socioeconomic disparities and barriers and lead to worse outcomes. Further large, prospective studies of SLE with an aim at recruiting men are needed to better understand racial differences in men with SLE in all domains affecting health care.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Rice, Carbone, Elam.

Acquisition of data. Rice, Ayyala, Coughlin, Elam.

Analysis and interpretation of data. Rice, Ayyala, Shi, Madera-Acosta, Bell, Qureshi, Carbone, Coughlin, Elam.

ACKNOWLEDGMENT

The authors thank the participants of the All of Us Research Program.

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;73:991–6.
2. Nusbaum JS, Mirza I, Shum J, et al. Sex differences in systemic lupus erythematosus: epidemiology, clinical considerations, and disease pathogenesis. Mayo Clin Proc 2020;95:384–94.
3. Andrade RM, Alarcón GS, Fernández M, et al. Accelerated damage accrual among men with systemic lupus erythematosus: XLIV. Results from a multiethnic US cohort. Arthritis Rheum 2007;56:622–30.
4. Lim SS, Drenkard C, McCune WJ, et al. Population-based lupus registries: advancing our epidemiologic understanding. Arthritis Rheum 2009;61:1462–6.
5. Tan TC, Fang H, Magder LS, et al. Differences between male and female systemic lupus erythematosus in a multiethnic population. J Rheumatol 2012;39:759–69.
6. Alonso MD, Martínez-Vázquez F, Riancho-Zarrabeitia L, et al. Sex differences in patients with systemic lupus erythematosus from Northwest Spain. Rheumatol Int 2014;34:11–24.
7. Ramírez Sepúlveda JI, Bolin K, Mofors J, et al. Sex differences in clinical presentation of systemic lupus erythematosus. Biol Sex Differ 2019;10:60.
8. Schwartzman-Morris J, Putterman C. Gender differences in the pathogenesis and outcome of lupus and of lupus nephritis. Clin Dev Immunol 2012;2012:604892.
9. Rastin M, Mahmoudi M, Sahebari M, et al. Clinical & immunological characteristics in systemic lupus erythematosus patients. Indian J Med Res 2017;146:224–9.
10. Zhang S, Su J, Li X, et al. Chinese SLE Treatment and Research group (CSTAR) registry: V. gender impact on Chinese patients with systemic lupus erythematosus. Lupus 2015;24:1267–75.
11. Arbuckle MR, James JA, Dennis GJ, et al. Rapid clinical progression to diagnosis among African-American men with systemic lupus erythematosus. Lupus 2003;12:99–106.
12. Scalzi LV, Hollenbeak CS, Wang L. Racial disparities in age at time of cardiovascular events and cardiovascular-related death in patients with systemic lupus erythematosus. Arthritis Rheum 2010;62:2767–75.

13. Fisher ES, Whaley FS, Krushat WM, et al. The accuracy of Medicare's hospital claims data: progress has been made, but problems remain. Am J Public Health 1992;82:243–8.
14. Seligman VA, Lum RF, Olson JL, et al. Demographic differences in the development of lupus nephritis: a retrospective analysis. Am J Med 2002;112:726–9.
15. Golder V, Connelly K, Staples M, et al. Association of Asian ethnicity with disease activity in SLE: an observational study from the Monash Lupus Clinic. Lupus 2013;22:1425–30.
16. Minhas D, Marder W, Harlow S, et al. Access and cost-related nonadherence to prescription medications among lupus patients and controls: the Michigan Lupus Epidemiology and Surveillance Program. Arthritis Care Res (Hoboken) 2021;73:1561–7.
17. Feldman CH, Broder A, Guan H, et al. Sex differences in health care utilization, end-stage renal disease, and mortality among Medicaid beneficiaries with incident lupus nephritis. Arthritis Rheumatol 2018; 70:417–26.
18. Ow MY, Ho PC, Thumboo J, et al. Factors associated with health services utilization in patients with systemic lupus erythematosus: a systematic review. Clin Exp Rheumatol 2010;28:892–904.
19. Denny JC, Rutter JL, Goldstein DB, et al. The "All of Us" Research Program. N Engl J Med 2019;381:668–76.
20. Cronin RM, Jerome RN, Mapes B, et al. Development of the initial surveys for the All of Us Research Program. Epidemiology 2019;30: 597–608.
21. Walitt BT, Constantinescu F, Katz JD, et al. Validation of self-report of rheumatoid arthritis and systemic lupus erythematosus: the Women's Health Initiative. J Rheumatol 2008;35:811–8.
22. Barnado A, Casey C, Carroll RJ, et al. Developing electronic health record algorithms that accurately identify patients with systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2017;69:687–93.
23. Stefanidou S, Benos A, Galanopoulou V, et al. Clinical expression and morbidity of systemic lupus erythematosus during a post-diagnostic 5-year follow-up: a male:female comparison. Lupus 2011;20:1090–4.
24. Jacobsen S, Petersen J, Ullman S, et al. A multicentre study of 513 Danish patients with systemic lupus erythematosus. I. Disease manifestations and analyses of clinical subsets. Clin Rheumatol 1998;17:468–77.
25. Voulgari PV, Katsimbri P, Alamanos Y, et al. Gender and age differences in systemic lupus erythematosus. A study of 489 Greek patients with a review of the literature. Lupus 2002;11:722–9.
26. Soto ME, Vallejo M, Guillén F, et al. Gender impact in systemic lupus erythematosus. Clin Exp Rheumatol 2004;22:713–21.
27. Mok CC, Lau CS, Chan TM, et al. Clinical characteristics and outcome of southern Chinese males with systemic lupus erythematosus. Lupus 1999;8:188–96.
28. Molina JF, Drenkard C, Molina J, et al. Systemic lupus erythematosus in males. A study of 107 Latin American patients. Medicine (Baltimore) 1996;75:124–30.
29. Basta F, Fasola F, Triantafyllias K, et al. Systemic lupus erythematosus (SLE) therapy: the old and the new. Rheumatol Ther 2020;7:433–46.
30. R Development Core Team. R: the R project for statistical computing. URL: https://R-project.org.
31. Lopez MH, Gonzalez-Barrera A. Women's college enrollment gains leave men behind. Pew Research Center. 2014. URL: https://www.pewresearch.org/fact-tank/2014/03/06/womens-college-enrollment-gains-leave-men-behind/.
32. Ward MM. Access to care and the incidence of endstage renal disease due to systemic lupus erythematosus. J Rheumatol 2010;37: 1158–63.
33. Scofield L, Reinlib L, Alarcón GS, et al. Employment and disability issues in systemic lupus erythematosus: a review. Arthritis Rheum 2008;59:1475–9.
34. Medina G, Vera-Lastra O, Barile L, et al. Clinical spectrum of males with primary antiphospholipid syndrome and systemic lupus erythematosus: a comparative study of 73 patients. Lupus 2004;13:11–6.
35. McKown T, Schletzbaum M, Unnithan R, et al. The effect of smoking on cumulative damage in systemic lupus erythematosus: an incident cohort study. Lupus 2021;30:620–9.
36. Nyman E, Vaughan T, Desta B, et al. Characteristics and symptom severity of patients reporting systemic lupus erythematosus in the PatientsLikeMe online health community: a retrospective observational study. Rheumatol Ther 2020;7:201–13.
37. Da Costa D, Dritsa M, Bernatsky S, et al. Dimensions of fatigue in systemic lupus erythematosus: relationship to disease status and behavioral and psychosocial factors. J Rheumatol 2006;33:1282–8.
38. Bensing JM, Hulsman RL, Schreurs KM. Gender differences in fatigue: biopsychosocial factors relating to fatigue in men and women. Med Care 1999;37:1078–83.
39. Garcia MA, Marcos JC, Marcos AI, et al. Male systemic lupus erythematosus in a Latin-American inception cohort of 1214 patients. Lupus 2005;14:938–46.
40. Walli-Attaei M, Joseph P, Rosengren A, et al. Variations between women and men in risk factors, treatments, cardiovascular disease incidence, and death in 27 high-income, middle-income, and low-income countries (PURE): a prospective cohort study. Lancet 2020; 396:97–109.
41. Maheswaranathan M, Cantrell S, Eudy AM, et al. Investigating health literacy in systemic lupus erythematosus: a descriptive review. Curr Allergy Asthma Rep 2020;20:79.
42. Hearth-Holmes M, Murphy PW, Davis TC, et al. Literacy in patients with a chronic disease: systemic lupus erythematosus and the reading level of patient education materials. J Rheumatol 1997;24: 2335–9.
43. Katz P, Dall'Era M, Trupin L, et al. Impact of limited health literacy on patient-reported outcomes in systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2021;73:110–9.
44. Geraldino-Pardilla L, Perel-Winkler A, Miceli J, et al. Association between hydroxychloroquine levels and disease activity in a predominantly Hispanic systemic lupus erythematosus cohort. Lupus 2019; 28:862–7.
45. Maheswaranathan M, Eudy AM, Bailey SC, et al. Low health numeracy is associated with higher disease activity in systemic lupus erythematosus. Lupus 2021;30:489–94.
46. Yazdany J, Trupin L, Tonner C, et al. Quality of care in systemic lupus erythematosus: application of quality measures to understand gaps in care. J Gen Intern Med 2012;27:1326–33.
47. Walunas TL, Jackson KL, Chung AH, et al. Disease outcomes and care fragmentation among patients with systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2017;69:1369–76.
48. Vina ER, Masi CM, Green SL, et al. A study of racial/ethnic differences in treatment preferences among lupus patients. Rheumatology (Oxford) 2012;51:1697–706.

Arthritis Care & Research
Vol. 75, No. 10, October 2023, pp 2107–2116
DOI 10.1002/acr.25103

American College of Rheumatology
Empowering Rheumatology Professionals

Influence of Sex on Early Axial Spondyloarthritis: A Six-Year Longitudinal Analysis From a Large National Cohort

Krystel Aouad,[1] Anne Tournadre,[2] Florian Lucasson,[3] Daniel Wendling,[4] Anna Molto,[5] Bruno Fautrel,[6] and Laure Gossec[6]

Objective. The objective was to determine sex differences in disease outcomes in recent axial spondyloarthritis (SpA) over time.

Methods. We analyzed the first 6 years of follow-up of the prospective French multicenter DESIR cohort. Patients analyzed had <3 years of disease, were naive to disease-modifying antirheumatic drugs, and fulfilled the Assessment of SpondyloArthritis international Society classification criteria for axial SpA. Disease activity (Ankylosing Spondylitis Disease Activity Score [ASDAS] using the C-reactive protein [CRP] level), patient global assessment (PtGA), CRP level, and radiographic sacroiliitis were compared between men and women (self-reported sex) by linear and logistic mixed-effects models. Models were created for trajectories of disease activity over 6 years in men and women, using k-means.

Results. Of 494 patients analyzed (mean ± SD age 31.9 ± 7.5 years, symptoms duration 20.7 ± 11.7 months), 50.4% were men. Over 6 years of follow-up, both men and women showed clear improvements in ASDAS-CRP, PtGA, and CRP level. Women had higher ASDAS-CRP and PtGA over time compared to men (both $P < 0.0001$) with overall similar CRP levels ($P = 0.089$), whereas structural damage increased more in men ($P < 0.0001$). One-third of both men (33%) and women (34%) belonged to persistent high/very high disease activity trajectories, but ASDAS-CRP was globally higher in women in these trajectories.

Conclusion. In early axial SpA, clinical outcomes (disease activity and symptoms) were worse in women than men over 6 years of follow-up, whereas CRP was similar and structural damage was more frequent in men. Although similarly distributed, disease activity scores were higher in women in high/very high disease activity trajectories. Sex appears to be an important contextual factor in axial SpA.

INTRODUCTION

In many diseases, sex-related differences have become a major focus of interest (1). In rheumatic and musculoskeletal diseases, women are increasingly recognized as differing from men regarding disease expression, clinical manifestations, disease progression, and treatment response (2,3). Sex-related differences may be multifactorial and complex, driven by genetic, hormonal, sociocultural, and/or psychological factors, thus leading to differences in disease pathophysiology, articular pattern, quality of life, and perception of pain between men and women. Recognizing and understanding these disease-specific differences may result in adaptations related to sexes, leading to a tailored management of the disease and better outcomes (1).

Historically, ankylosing spondylitis was considered a predominantly male disease; however, currently, the male to female

ClinicalTrials.gov identifier: NCT01648907.

The DESIR cohort is conducted as a programme hospitalier de recherche clinique with Assistance Publique–Hôpitaux de Paris as the sponsor. This study is conducted under the umbrella of the French Society of Rheumatology, which is also financially supporting this cohort. An unrestricted grant from Pfizer was allocated for the first 10 years. The present analysis was funded through an investigator-initiated grant from UCB.

[1]Krystel Aouad, MD, MSc: INSERM UMRS 1136-6, Institut Pierre Louis d'Epidémiologie et de Santé Publique, Sorbonne Université, Paris, France, and Saint George Hospital University Medical Center, Saint George University of Beirut, Beirut, Lebanon; [2]Anne Tournadre, MD, PhD: Université Clermont Auvergne, CHU Clermont-Ferrand, INRAE, UNH UMR 1019, Clermont-Ferrand, France; [3]Florian Lucasson, MD, MSc: INSERM UMRS 1136-6, Institut Pierre Louis d'Epidémiologie et de Santé Publique, Sorbonne Université, Paris, France; [4]Daniel Wendling, MD, PhD: CHRU de Besançon, University Teaching Hospital, and EA 4266 EPILAB Université Bourgogne Franche-Comté, Besançon, France; [5]Anna Molto, MD, PhD: Cochin Hospital AP-HP and INSERM U1153, Sorbonne Paris Cité Research Center, Paris University, Paris, France; [6]Bruno Fautrel, MD, PhD, Laure Gossec, MD, PhD: INSERM UMRS 1136-6, Institut Pierre Louis d'Epidémiologie et de Santé Publique and Pitié-Salpêtrière Hospital AP-HP, Sorbonne Université, Paris, France.

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.25103&file=acr25103-sup-0001-Disclosureform.pdf.

Address correspondence via email to Laure Gossec, MD, PhD, at: laure.gossec@aphp.fr.

Submitted for publication February 24, 2022; accepted in revised form February 9, 2023.

SIGNIFICANCE & INNOVATIONS

- Less improvement in disease activity and patient-reported outcomes was observed in women over time.
- Structural damage was more common in men than women over 6 years of follow-up.
- Sex is an important contextual parameter to consider in the management of axial spondyloarthritis.

ratio is 2–3:1 in radiographic axial spondyloarthritis (SpA), with an equal sex distribution in its nonradiographic form (4). Recent data suggest differences in disease phenotypes between men and women with axial SpA, in terms of disease activity and disease progression (5–12). Women seem to have a higher disease burden and worse functional impairment but less structural damage (13–15). However, data on sex differences in axial SpA available today are mostly issued from cross-sectional, retrospective studies or long-standing disease cohorts.

In chronic inflammatory arthritis, studying the early period of the disease is key since many changes occur in the early stages (16). Furthermore, cohort studies allow prospective longitudinal analyses over time (17). Therefore, identifying the consequences of sex on axial SpA in an inception cohort is of interest.

We hypothesized that sex may influence disease activity progression and that different factors at baseline might be related to disease progression in men and women. We have previously explored disease activity trajectories in axial SpA, and homogeneous groups of disease activity were identified (18). Such methods could be applied to explore the influence of sex in axial SpA.

The objective of this study was to explore differences in outcomes, including disease activity, patient-reported outcomes, inflammatory markers, and radiographic damage between men and women with axial SpA, and to identify sex-specific factors associated with disease activity trajectories, in an inception cohort with a longitudinal follow-up.

MATERIALS AND METHODS

Study design and patients. The DESIR cohort is a French longitudinal, prospective, multicenter observational cohort including patients age >18 and <50 years with early inflammatory back pain fulfilling either the Calin or Berlin criteria and with symptom duration >3 months and <3 years suggestive of SpA according to the rheumatologist's assessment (e.g., a score of ≥5 on a 0–10 numerical rating scale, where 0 = not suggestive and 10 = very suggestive) (19). Of 708 patients included in the cohort, we analyzed only patients fulfilling the Assessment of SpondyloArthritis international Society (ASAS) classification criteria of axial SpA (20) at entry or at any time point during follow-up visits (see Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25103). All patients were naive to biologic disease-modifying antirheumatic drugs (DMARDs). Baseline analyses according to sex have been reported elsewhere (5); here, the first 6 years of follow-up were analyzed (with data collected at 0, 6, 12, 18, 24, 36, 48, 60, and 72 months). The DESIR cohort was approved by the French ethical committee and the study was conducted according to Good Clinical Practice guidelines. All participants provided written informed consent (21). Ethics committee approval was already obtained through the DESIR cohort agreements. The data underlying this article were provided by the DESIR cohort under permission and cannot be shared publicly. Data will be shared on reasonable request to the corresponding author with permission of the DESIR cohort scientific committee.

Clinical outcomes. Sex as reported by the patient at baseline was the main variable of interest. The outcome criteria collected over time and analyzed according to sex were: 1) a validated composite disease activity score: the Ankylosing Spondylitis Disease Activity Score (ASDAS) using the C-reactive protein (CRP) level (22); 2) a patient-reported outcome: the patient global assessment (PtGA) using the Bath Ankylosing Spondylitis Patient Global Score over the past week; 3) an acute-phase reactant: CRP level; and 4) radiographic sacroiliitis according to the modified New York criteria. The first 3 outcomes were evaluated at 0, 6, 12, 18, 24, 36, 48, 60, and 72 months, whereas radiographs were obtained and scored at baseline, 2 years, and 5 years. Demographic and clinical data at baseline included age, relevant family history, duration, and vertebral level of inflammatory back pain, HLA–B27 status, extraarticular manifestations (uveitis, inflammatory bowel disease, psoriasis), and other manifestations (dactylitis, enthesitis, arthritis).

Secondary analyses were performed for other disease activity scores: Bath Ankylosing Spondylitis Disease Activity Index (BASDAI; range 0–100), Ritchie articular index (53 joints weighted from 0 to 3), 28-swollen joint score (range 0–28), and Maastricht Ankylosing Spondylitis Enthesitis Score (MASES; range 0–13) (22). Physical function and quality of life were assessed using the Health Assessment Questionnaire for Ankylosing Spondylitis (HAQ-AS) and the Bath Ankylosing Spondylitis Functional Index (BASFI), the Ankylosing Spondylitis Quality of Life (ASQoL) questionnaire, and the Short Form-36 health survey questionnaire (SF-36) (22).

Treatment intake, defined as conventional synthetic DMARDs (csDMARDs; yes/no), and separately, as tumor necrosis factor inhibitor (TNFi) use (yes/no) at each timepoint, was analyzed by comparing the sexes using a mixed-effects model. Radiographs of sacroiliac joints and the spine were assessed by local reading at baseline, 2 years, and 5 years. Spine radiographs were scored according to the modified Stoke Ankylosing Spondylitis Spine Score (mSASSS).

Statistical analysis. Baseline characteristics of men and women were compared using *t*-tests or Mann-Whitney U tests for continuous variables and chi-square tests or Fisher's exact test for categorical variables. To assess the impact of sex over time, taking into account longitudinal repetitive measures, mixed-effects linear and logistic regression models were performed (using lme4 R package) for continuous and binary dependent variables, respectively. In the mixed models, sex and time were determined as fixed effects and subjects as random effects (more details in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25103). To illustrate the evolution over time, the mean of the last 2 values of each outcome was calculated. Missing data were handled using mixed models; imputation of missing data was not needed and not performed.

Models were created of trajectories in men and in women separately through homogeneous clusters of disease activity based on ASDAS-CRP, using the R package kml (23). The optimal number of clusters (from 2 to 6) was chosen using the Calinski and Harabatz criterion. Only patients for whom at least 3 ASDAS values were available during the 6 years of follow-up were included in this analysis, as was done in previous DESIR analyses (see Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25103). Predisposing baseline factors (demographic factors, disease phenotype, disease severity, quality of life, radiographic sacroiliitis) associated with the different trajectories were analyzed by univariate analysis. Covariates that were significantly associated with at least 1 trajectory ($P < 0.1$) in univariate analysis were included in the multivariable multinomial logit regression, after excluding colinear variables. For the analyses, the persistent low disease activity trajectory was used as the reference trajectory because it included the largest number of patients (more details in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25103). Statistical analyses were performed using the Software R Studio, version 1.4.1106.

RESULTS

Of the 708 patients included in the DESIR cohort at baseline, 494 (69.7%) who fulfilled the ASAS classification criteria for axial SpA at any time point were analyzed (see Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25103). The mean ± SD age was 31.9 ± 7.5 years, disease duration was 20.7 ± 11.7 months, and 249 patients (50.4%) were men and 245 (49.6%) were women.

Comparison of baseline characteristics between men and women. At baseline (Table 1), the mean age was similar between men and women, but the duration of inflammatory back pain was longer in women. At baseline, the level of disease activity was high and similar in men and women (mean ± SD ASDAS-CRP was 2.6 ± 1.0 in men versus 2.6 ± 0.9 in women, $P = 0.75$), whereas subjective measures such as PtGA and other patient-reported outcomes (SF-36, HAQ-AS, ASQoL, BASFI), but also objective signs such as tender and swollen joint counts and MASES, were worse in women at baseline. However, CRP levels were higher and sacroiliitis was more common in men.

Impact of sex on disease outcomes. *Main outcomes.* Although disease activity was comparable at baseline and decreased in both groups, women had significantly higher

Table 1. Comparison of baseline characteristics between men and women with axial spondyloarthritis*

	All (n = 494)	Men (n = 249)	Women (n = 245)	*P*
Age, years	31.9 ± 7.5	31.2 ± 7.6	32.6 ± 7.3	0.03†
Duration of inflammatory back pain, years	1.7 ± 1.0	1.7 ± 0.9	1.8 ± 1.0	<0.0001†
Past history, no. (%)				
Dactylitis	70 (14.2)	32 (12.9)	38 (15.5)	0.40
Enthesitis	265 (53.6)	131 (52.6)	134 (54.7)	0.64
Uveitis	47 (9.5)	20 (8.0)	27 (11.0)	0.26
Peripheral arthritis	138 (28.0)	73 (29.3)	65 (26.5)	0.51
Inflammatory bowel disease	26 (5.3)	10 (4.0)	16 (6.5)	0.21
Psoriasis	78 (16.6)	38 (16.0)	40 (17.1)	0.50
HLA–B27 positivity, no. (%)	378 (76.5)	201 (80.7)	177 (72.2)	0.02†
Tender joint score (0–53)	2.6 ± 4.8	1.8 ± 4.0	3.5 ± 5.4	<0.0001†
Swollen joint score (0–28)	0.1 ± 0.7	0.1 ± 0.9	0.2 ± 0.3	<0.0001†
MASES (0–13)	2.4 ± 3.0	1.4 ± 2.2	3.4 ± 3.3	<0.0001†
MRI sacroiliitis, no. (%)‡	220 (44.5)	132 (53.0)	87 (35.5)	0.0002†
ASAS criteria clinical arm verified, no. (%)	189 (40.1)	78 (32.9)	111 (47.4)	0.0013†
ASAS criteria imaging arm verified, no. (%)	245 (52.0)	147 (62.1)	98 (41.9)	<0.0001†

* Values are the mean ± SD unless indicated otherwise. ASAS = Assessment of SpondyloArthritis international Society; MASES = Maastricht Ankylosing Spondylitis Enthesitis Score; MRI = magnetic resonance imaging.
† Statistically significant.
‡ MRI sacroiliitis according to the local investigator.

ASDAS-CRP over 6 years of follow-up ($P < 0.0001$ in mixed models) (Figure 1 and Table 2). To illustrate the mean ASDAS-CRP was 1.9 ± 1.0 in men versus 2.3 ± 0.9 in women at the last 2 time points (Table 2 and see Supplementary Tables 2 and 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25103, for more details and data availability for each timepoint). Women had significantly higher PtGA that persisted over 6 years of follow-up ($P < 0.0001$ in mixed models) (Figure 1 and Table 2).

Although significantly higher CRP levels were seen in men at baseline, overall CRP levels were similar over 6 years of follow-up ($P = 0.090$ in mixed models) At baseline and over time, more men had radiographic sacroiliitis ($P < 0.0001$ in mixed models) (Figure 1 and Table 2). Radiographic sacroiliitis was seen in 72 men (28.9%) and 32 women (13.1%) at 5 years (Table 2). At 6 years, the main outcomes (ASDAS-CRP, PtGA, and CRP level) showed significantly higher values in women compared to men in both radiographic and nonradiographic axial SpA (see Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25103).

Secondary outcomes. As seen in Table 2, women had significantly worse patient-reported outcomes (BASDAI, BASFI, ASQoL, HAQ-AS, SF-36) than men over time, higher MASES, and higher tender/swollen joint counts but less sacroiliitis on magnetic resonance imaging (MRI) and less structural spinal damage (mSASSS) over time (all $P < 0.05$ in mixed models). Drug intake at least once over follow-up (csDMARDs and TNFi) was similar between sexes over time (Table 2).

Trajectories. When analyzing trajectories separately in men and women (total 352 patients), we identified 6 different trajectories in both in men and women (Figure 2 and Supplementary Figure 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25103). A nomenclature for trajectories of disease activity in DESIR has been previously described and was used here (18,24). The distribution of patients in different trajectories was globally similar in men and women, e.g., 68 men (38%) and 73 women (42%) had an inactive or low disease activity evolution, whereas 59 men (33%) and 60 women (34%) had a high/very high disease evolution. ASDAS-CRP was globally higher in women in high/very high disease trajectories as shown in Supplementary Figures 2A and 2B, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25103. On the

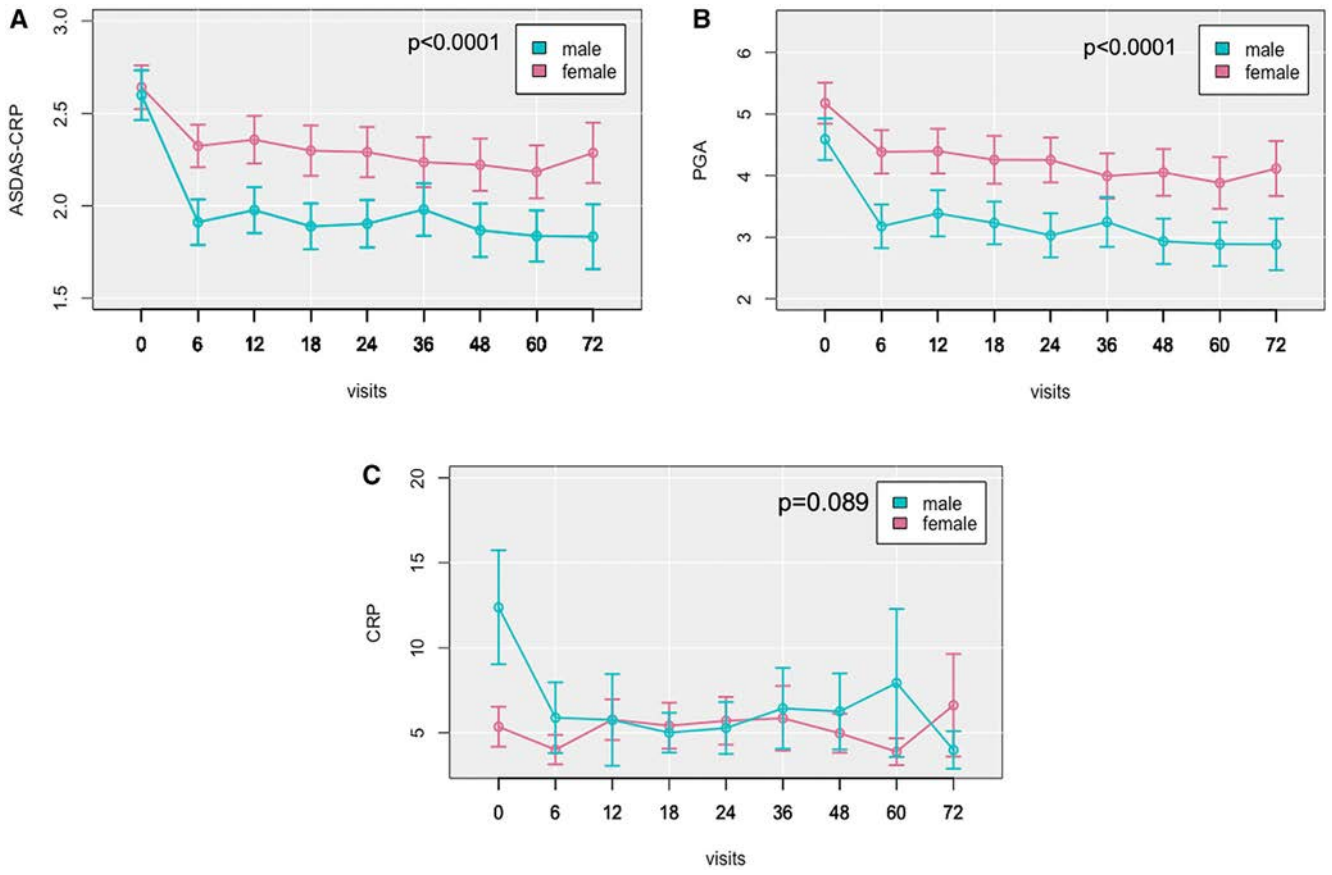

Figure 1. Evolution of disease activity over 6 years in men and women: **A**, Ankylosing Spondylitis Disease Activity Score using the C-reactive protein level (ASDAS-CRP); **B**, Patient global assessment (PGA); **C**, CRP level. The *P* values shown in the figures are the results of linear mixed-effects models. x-axis = visits (months).

Table 2. Effect of sex on different clinical outcomes in axial spondyloarthritis over 6 years of follow-up using mixed-effects models*

	Baseline value		Final value†		
	Men	Women	Men	Women	*P* for effect estimate of sex by mixed-effects models over time (n = 494)
ASDAS-CRP	2.6 ± 1.0	2.6 ± 1.0	1.9 ± 1.0	2.25 ± 0.9	<0.0001‡
PtGA (0–10)	4.6 ± 2.6	5.2 ± 2.6	2.9 ± 2.5	4.1 ± 2.8	<0.0001‡
CRP, mg/liter	10.7 ± 16.5	6.2 ± 8.3	6.0 ± 11.6	4.9 ± 6.5	0.09
mNY radiographic sacroiliitis, follow-up (local reading), no. (%)§	64 (25.7)	42 (17.1)	72 (28.9)	32 (13.1)	<0.0001‡
Exploratory outcomes					
BASDAI (0–100)	40.5 ± 20.2	46.3 ± 18.5	25.7 ± 19.3	37.5 ± 21.2	<0.0001‡
BASFI (0–100)	26.9 ± 21.3	32.4 ± 23.3	15.9 ± 18.3	26.0 ± 22.3	<0.0001‡
ASQoL (0–18)	8.0 ± 5.0	10.1 ± 4.8	4.5 ± 4.7	8.0 ± 5.5	<0.0001‡
SF-36					
PCS (0–100)	41.3 ± 8.9	39.0 ± 9.1	47.0 ± 8.7	41.0 ± 9.6	<0.0001‡
MCS (0–100)	42.3 ± 11.2	39.6 ± 11.3	47.1 ± 10.1	43.0 ± 11.6	0.0002‡
HAQ-AS (0–3)	0.5 ± 0.5	0.8 ± 0.5	0.3 ± 0.4	0.63 ± 0.5	<0.0001‡
Tender joint count (0–53)	2.5 ± 5.2	4.5 ± 7.5	1.1 ± 3.6	3.8 ± 6.2	<0.0001‡
Swollen joint count (0–28)	0.1 ± 0.3	0.2 ± 0.9	0.02 ± 0.2	0.1 ± 0.7	0.04‡
MASES (0–13)	2.1 ± 3.4	5.4 ± 6.1	1.1 ± 2.9	4.1 ± 5.4	<0.0001‡
mSASSS (0–72)	0.5 ± 1.7	0.2 ± 0.6	1.5 ± 5.1	0.6 ± 2.0	0.02‡
csDMARDs intake, no. (%)	35 (14.5)	33 (14.1)	30 (12.7)	28 (12.0)	0.74
TNFi intake at least once during follow-up, no. (%)	0	0	72.5 (30.5)	70 (29.9)	0.85
MRI sacroiliitis during follow-up (local reading), no. (%)§	132 (53.0)	87 (35.5)	27 (10.8)	19 (7.8)	<0.0001‡

* Values are the mean ± SD unless indicated otherwise. ASDAS-CRP = Ankylosing Spondylitis Disease Activity Score using the C-reactive protein level; ASQoL = Ankylosing Spondylitis Quality of Life; BASDAI = Bath Ankylosing Spondylitis Disease Activity Index; BASFI = Bath Ankylosing Spondylitis Functional Index; csDMARDs = conventional synthetic disease-modifying antirheumatic drugs; HAQ-AS = Health Assessment Questionnaire for Ankylosing Spondylitis; MASES = Maastricht Ankylosing Spondylitis Enthesitis Score; MCS = mental component summary; MRI = magnetic resonance imaging; mNY = modified New York criteria; mSASSS = modified Stoke Ankylosing Spondylitis Spine Score; PCS = physical component summary; PtGA = patient global assessment; SF-36 = Short Form-36 health survey questionnaire; TNFi = tumor necrosis factor inhibitors.
† The final value represents the mean of the last 2 values at months 60 and 72 for all clinical outcomes except for radiographic sacroiliitis, MRI sacroiliitis, and mSASSS (only values at month 60).
‡ Significant results ($P < 0.05$). The *P* value reflects differences between sexes over time, using all available timepoints, through mixed models. For illustrative purposes, the baseline and final values of each variable are presented, though the *P* values represent all timepoints.
§ According to the local investigator.

other hand, 22 men (12%) and 15 women (9%) had almost major improvement on ASDAS-CRP.

Baseline characteristics associated with each trajectory in men and women. The multivariable analysis showed that higher HAQ-AS scores at baseline were associated with increased odds of being in the persistent high/very high disease activity group compared to persistent low disease activity in both men and women (Table 3). On the other hand, higher HAQ-AS scores at baseline in men (odds ratio 0.09 [95% confidence interval 0.01–0.79]) were significantly associated with decreased odds of being in the persistent inactive disease activity compared to persistent low disease activity.

DISCUSSION

This large cohort of patients with early axial SpA demonstrates the differences in disease phenotypes between men and women. For both sexes, disease activity, function, and pain improved over time. However, although disease activity was similar between sexes at baseline, there was less improvement in disease activity, functioning, and quality of life in women than in men over 6 years of follow-up, leading to higher disease activity and symptoms in women at 6 years. The higher level of disease activity over time was confirmed by a second statistical technique: even though one-third of both men and women had a persistent high/very high disease activity evolution, disease activity scores were globally higher in women in these trajectories. Finally, structural damage and MRI inflammation were more prevalent in men than women, although they had similar acute-phase reactants and treatment intake.

This study has some strengths and limitations. This longitudinal analysis of the impact of sex on disease outcomes in axial SpA brings different and valuable information compared to published data, which are mainly cross-sectional (5–8,10,12,13,15,25,26). A strength of our study is the validated methodologies used to analyze longitudinal data: mixed-effects models and trajectory analyses. The use of k-means algorithm presents an original and innovative methodology to study longitudinal data; trajectory modeling was recently used in axial SpA in previous studies on

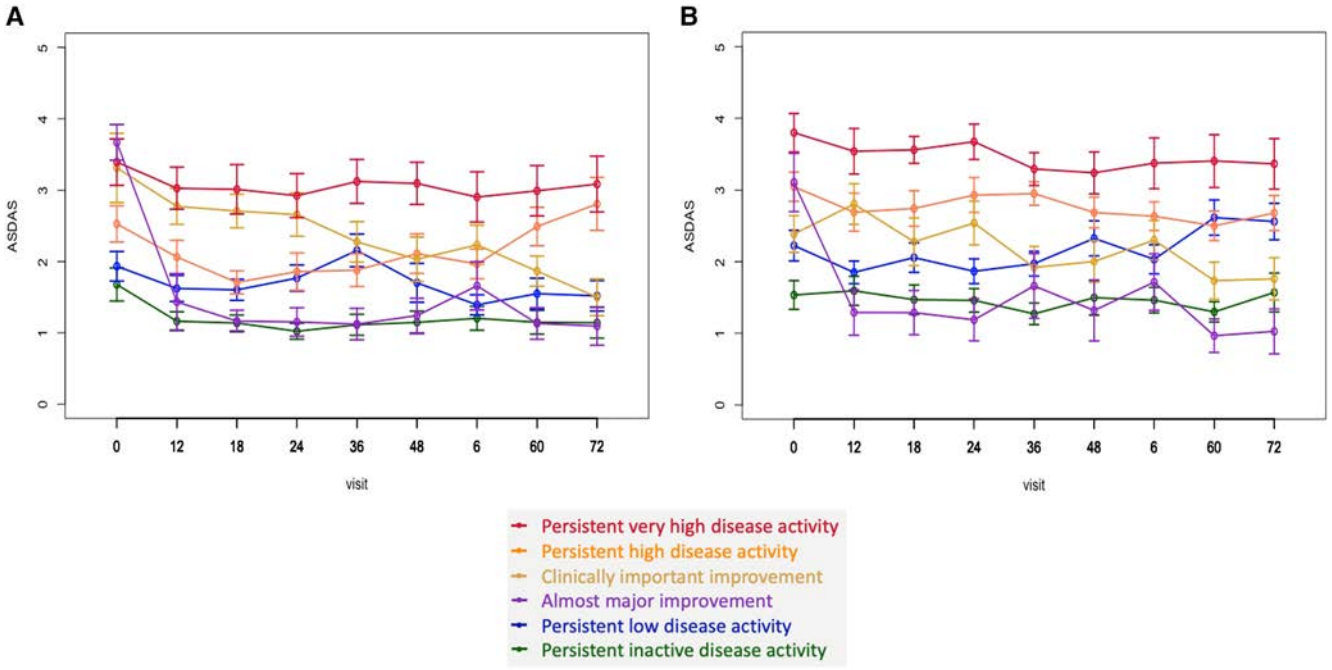

Figure 2. Trajectories of disease activity in men and women with recent axial spondyloarthritis. **A**, All trajectories of disease activity in men. **B**, All trajectories of disease activity in women. ASDAS = Ankylosing Spondylitis Disease Activity Score. Error bars indicate for each trajectory, the sample mean ±1 SD is represented at each time point. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25103/abstract.

the same cohort after 3 years and 5 years of follow-up (18,24). The DESIR cohort included a large sample of patients; however, many patients could not be analyzed because they did not satisfy the ASAS criteria for axial SpA. The sample of patients allowed us to define distinctive trajectories of disease activity stable over time. However, given the small sample size in each specific trajectory we chose not to overinterpret baseline characteristics of the patients in each trajectory. One may argue that patient-reported outcomes included in the ASDAS may have contributed more than CRP level to trajectory clustering. However, both patient-reported outcomes and acute-phase reactants are recommended for disease activity monitoring in axial SpA (27). Although treatment intake over time was analyzed (and was similar between men and women over time), the types and dosages of TNFi and treatment retention were not analyzed. As in many longitudinal studies, data were missing over time; however, the DESIR cohort is one of the largest cohorts of early axial SpA, and using mixed-effects models allowed us to overcome missing data and analyze all the data of patients included in the cohort.

Our study showed that disease activity scores were higher in women over time as reflected by both ASDAS and BASDAI. This finding is in line with previous cross-sectional and observational studies reporting that female sex was associated with higher disease activity, mainly observed on the BASDAI scores (5–8,10,13,25,26), and a few recent studies also reported higher ASDAS in women (12,28). In this study, women had more peripheral arthritis and enthesitis than men. Past studies have reported similar findings (6,8,15,26). As for patient-reported outcomes measures, PtGA, BASFI, ASQoL, HAQ-AS, and SF-36 scores were worse in women, determining poor functional status and well-being. This finding is consistent with previous observations showing overall higher disease burden and functional impairment in female sex with axial SpA (6,10,15,26,29). On the other hand, there was no difference in CRP levels between sexes over time, and therefore, the difference in ASDAS was mainly driven by the patient-reported outcomes (axial and peripheral pain, duration of morning stiffness, PtGA). This difference in ASDAS highlights the fact that disease activity scores may be somewhat subjective and driven by central sensitization, in contrast to MRI inflammation, which may be a more subjective tool.

Altogether, in the present study, both disease activity scores and patient-reported outcomes were worse in women. Notably, similar findings were seen in women with chronic rheumatic and musculoskeletal diseases such as rheumatoid arthritis and psoriatic arthritis (2,30), thus emphasizing increased pain sensitivity widely acknowledged by the female subjects (6,31). Women had a higher percentage of nonradiographic axial SpA, with higher disease activity scores compared to men (32). These findings may be partly explained by the higher prevalence of concomitant fibromyalgia among women with axial SpA (33). The ASDAS-CRP used in our analysis and trajectory modeling is a rather objective score that seems to be less influenced by secondary fibromyalgia in patients with axial SpA (34). Nevertheless, studies have shown that coexistent fibromyalgia characterized by chronic musculoskeletal widespread pain and tender points may impact disease activity scores and treatment response in axial SpA (32), making

Table 3. The association of baseline characteristics with trajectories of disease activity in men and women with early axial spondyloarthritis*

	Trajectory 1: persistently inactive (n = 69)		Trajectory 2: persistently low (n = 72)†		Trajectory 3: almost major improvement (n = 37)		Trajectory 4: clinically important improvement (n = 55)		Trajectory 5: persistently high (n = 68)		Trajectory 6: persistently very high (n = 51)	
Variables in multinomial logit regression	Men (n = 35, 19.8%)	Women (n = 34, 19.4%)	Men (n = 33, 18.6%)	Women (n = 39, 22.3%)	Men (n = 22, 12.4%)	Women (n = 15, 8.57%)	Men (n = 28, 15.8%)	Women (n = 27, 15.4%)	Men (n = 31, 17.5%)	Women (n = 37, 21.1%)	Men (n = 28, 15.8%)	Women (n = 23, 13.1%)
HLA–B27 positivity	1.61 (0.25–10.45)	NI	–	–	2.72 (0.35–21.42)	NI	0.89 (0.19–4.05)	NI	2.91 (0.45,19.07)	NI	0.42 (0.01–1.84)	NI
Baseline												
TJC score (0–53)	1.22 (0.69–2.16)	0.89 (0.68–1.15)	–	–	1.71 (1.09–2.69)‡	1.16 (0.97–1.39)	1.36 (0.88–2.10)	1.01 (0.86–1.18)	1.51 (0.96–2.35)	1.02 (0.89–1.17)	1.52 (0.98–2.37)	1.02 (0.88–1.18)
SJC score (0–28)	2.59 (0.19–35.90)	NI	–	–	2.88 (0.29–28.69)	NI	1.73 (0.19–16.12)	NI	0.19 (0.01–3.94)	NI	$1.49.10^{-7}$ (NS)	NI
MASES (0–13)	0.64 (0.39–1.05)	0.95 (0.08–1.37)	–	–	0.89 (0.64–1.21)	0.70 (0.50–0.97)‡	1.03 (0.79–1.33)	0.96 (0.78–1.18)	0.80 (0.58–1.11)	1.05 (0.88–1.26)	0.87 (0.63–1.20)	1.12 (0.91–1.37)
HAQ-AS (0–3)	0.09 (0.01–0.79)‡	0.34 (0.07–1.54)	–	–	3.22 (0.55–19.04)	1.34 (0.28–6.40)	4.92 (0.96–25.34)	2.07 (0.58–7.32)	7.72 (1.49–39.97)‡	5.98 (1.78–20.04)‡	17.24 (3.13–9.51)‡	15.00 (3.55–62.40)‡
Radiographic sacroiliitis, mNY (local reading)	0.82 (0.22–3.07)	NI	–	–	6.74 (1.72–26.45)‡	NI	1.63 (0.46–5.80)	NI	1.66 (0.49–5.66)	NI	2.53 (0.70–9.12)	NI

* Values are the odds ratio (95% confidence interval). Covariates significantly associated with any trajectory ($P < 0.1$) in univariate analysis were included in the model. Variables showing collinearity were excluded from the analysis. HAQ-AS = Health Assessment Questionnaire for Ankylosing Spondylitis; MASES = Maastricht Ankylosing Spondylitis Enthesitis Score; mNY = modified New York criteria; NI = not included in the model; NS = not significant; SJC = swollen joint count; TJC = tender joint count.
† Trajectory 2 was used as the reference trajectory.
‡ Significant results ($P < 0.05$).

the diagnosis particularly difficult in the presence of axial SpA with entheseal symptoms (33). Thus, there is a challenge to differentiate entheseal tenderness due to central sensitization from enthesitis resulting in real inflammation of the entheses in axial SpA patients. Using more objective imaging tools in the evaluation of peripheral symptoms in axial SpA as well as pain pathways might be of interest.

Because of the association between disease activity and structural progression (35), one would expect women with higher disease activity to have more structural damage. However, this expectation was not confirmed by our study and other past studies showing significantly more structural progression in men (7,36–38). Whether men tend to minimize their symptoms leading to worse clinical outcomes remains unclear. The question of whether lower cutoffs are needed for disease activity scores in men is raised. We also believe that difficulties to link clinical outcomes to radiographic progression exist in axial SpA, in particular when radiographic progression is overall low, as is the case in the DESIR cohort (39).

Treatment intake (TNFi and csDMARD) in men and women was similar over time, even though women had higher disease activity scores and disease burden. Several hypotheses could explain this finding. We think it could be due to the higher treatment failure and lower time taking drugs observed in female patients (7), or on the contrary, due to the better recognition by rheumatologists of the overlap with fibromyalgia, resulting in less treatment initiation or switching in women. This possibility needs to be further clarified in future studies.

Approximately one-third of men and women remained in high/very high disease activity trajectories, thus implying an insufficient control of the disease despite the wide use of biologics. This finding was also reported by Molto et al in the analysis of the whole cohort population at 3 years of follow-up (18). Women had globally higher ASDAS in these 2 trajectories, which are consistent with the results from mixed models. Notably, coexistent fibromyalgia in women may have influenced the persistence of increased disease activity scores, which include subjective elements such as PtGA and pain (40). Furthermore, high HAQ-AS scores at baseline were strongly associated with both high and very high disease activity evolution in both sexes. This finding is probably linked to the existing high correlation between HAQ-AS and other patient-reported outcomes, including those in the ASDAS. The limitations of the ASDAS are again highlighted, and appear to be driven by patient-reported outcomes (35,41). Factors associated with disease evolution may help physicians predict the evolution of symptoms in axial SpA while taking into account sex in the equation.

These sex-related differences may be explained by sociocultural, genetic, hormonal, and immunologic factors that may result in distinct pathogenesis and phenotypes of the same disease in both sexes and do not seem to be linked to differences in treatment intake. The perception of pain is a subjective experience and differs between sexes. Andromativity and hegemonic masculinity can influence the expression and coping with pain based on gender norms (31). From a sociocultural perspective, women tend to verbalize and report more pain, whereas chronic pain in men may be seen as a sign of weakness and a threat to masculine identity (42). Higher masculinity was shown to be associated with higher pain tolerance and pain thresholds, whereas femininity was associated with greater sensitivity to pain (31).

As for genetic factors, the most relevant one, HLA–B27 positivity, appears to be more prevalent in men, explaining somewhat the higher structural damage (43). Sex hormones (estrogen and testosterone levels) may also play a role in the dissimilar immune response and pain response between sexes (44). Cytokine release may also differ: higher levels of TNF and interleukin-17 were detected in men with axial SpA, leading to osteoblastic differentiation, osteoproliferation, and stimulation of inflammatory pathways, inducing more bone damage (45,46). Further studies are needed to better identify pathogenic mechanisms and factors causing differential clinical patterns and disease progression between sexes.

In conclusion, our analysis confirms that women seem to have a different disease phenotype than men in axial SpA. Over 6 years of follow-up, women had higher disease activity and worse functioning and well-being but less structural damage. Notably, treatment intake was generally similar over time between sexes, although this observational study, which did not control for treatment received, may have overlooked more subtle differences in treatments between sexes. Recognizing sex as an important contextual factor in axial SpA is essential. Sex differences should be widely recognized and taken into account in the management of the disease. More studies are needed to determine the factors related to sex differences in the view of predicting and improving outcomes.

ACKNOWLEDGMENTS

The authors would like to thank the scientific committee of the DESIR cohort for accepting our application of the current research initiative, the investigators for collecting the data permitting such analyses, and the patients for agreeing to participate.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Aouad, Gossec.

Acquisition of data. Aouad, Gossec.

Analysis and interpretation of data. Aouad, Tournadre, Lucasson, Wendling, Molto, Fautrel, Gossec.

ROLE OF THE STUDY SPONSOR

UCB 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 UCB.

REFERENCES

1. Mauvais-Jarvis F, Bairey Merz N, Barnes PJ, et al. Sex and gender: modifiers of health, disease, and medicine. Lancet 2020;396:565–82.
2. Sokka T, Toloza S, Cutolo M, et al. Women, men, and rheumatoid arthritis: analyses of disease activity, disease characteristics, and treatments in the QUEST-RA study. Arthritis Res Ther 2009;11:R7.
3. Højgaard P, Ballegaard C, Cordtz R, et al. Gender differences in biologic treatment outcomes: a study of 1750 patients with psoriatic arthritis using Danish Health Care Registers. Rheumatology (Oxford) 2018;57:1651–60.
4. Sieper J, Poddubnyy D. Axial spondyloarthritis. Lancet 2017;390: 73–84.
5. Tournadre A, Pereira B, Lhoste A, et al. Differences between women and men with recent-onset axial spondyloarthritis: results from a prospective multicenter French cohort. Arthritis Care Res (Hoboken) 2013;65:1482–9.
6. Mease PJ, McLean RR, Dube B, et al. Comparison of men and women with axial spondyloarthritis in the US-based Corrona Psoriatic Arthritis/Spondyloarthritis Registry. J Rheumatol 2021;48:1528–36.
7. Rusman T, van Bentum RE, van der Horst-Bruinsma IE. Sex and gender differences in axial spondyloarthritis: myths and truths. Rheumatology (Oxford) 2020;59 Suppl 4:iv38–46.
8. Neuenschwander R, Hebeisen M, Micheroli R, et al. Differences between men and women with nonradiographic axial spondyloarthritis: clinical characteristics and treatment effectiveness in a real-life prospective cohort. Arthritis Res Ther 2020;22:233.
9. Van der Horst-Bruinsma IE, Zack DJ, Szumski A, et al. Female patients with ankylosing spondylitis: analysis of the impact of gender across treatment studies. Ann Rheum Dis 2013;72:1221–4.
10. De Carvalho HM, Bortoluzzo AB, Gonçalves CR, et al. Gender characterization in a large series of Brazilian patients with spondyloarthritis. Clin Rheumatol 2012;31:687–95.
11. Chen HH, Chen YM, Lai KL, et al. Gender difference in ASAS HI among patients with ankylosing spondylitis. PLoS One 2020;15: e0235678.
12. Capelusnik D, Zhao SS, Boonen A, et al. Individual and country-level socio-economic factors and health outcomes in spondyloarthritis: analysis of the ASAS-perSpA study. Rheumatology (Oxford) 2022; 61:2043–53.
13. De Jong H, Paramarta JE, de Winter J, et al. Differences between females and males in axial spondyloarthritis: data from a real-life cross-sectional cohort. Scand J Rheumatol 2020;49:28–32.
14. Van Tubergen A. The changing clinical picture and epidemiology of spondyloarthritis. Nat Rev Rheumatol 2015;11:110–8.
15. Landi M, Maldonado-Ficco H, Perez-Alamino R, et al. Gender differences among patients with primary ankylosing spondylitis and spondylitis associated with psoriasis and inflammatory bowel disease in an iberoamerican spondyloarthritis cohort. Medicine (Baltimore) 2016;95:e5652.
16. Molto A, López-Medina C, Van den Bosch FE, et al. Efficacy of a tight-control and treat-to-target strategy in axial spondyloarthritis: results of the open-label, pragmatic, cluster-randomised TICOSPA trial. Ann Rheum Dis 2021;80:1436–44.
17. Hirano F, van der Heijde D, van Gaalen FA, et al. Determinants of the patient global assessment of well-being in early axial spondyloarthritis: 5-year longitudinal data from the DESIR cohort. Rheumatology (Oxford) 2021;60:316–21.
18. Molto A, Tezenas du Montcel S, Wendling D, et al. Disease activity trajectories in early axial spondyloarthritis: results from the DESIR cohort. Ann Rheum Dis 2017;76:1036–41.
19. Dougados M, d'Agostino MA, Benessiano J, et al. The DESIR cohort: a 10-year follow-up of early inflammatory back pain in France: study design and baseline characteristics of the 708 recruited patients. Joint Bone Spine 2011;78:598–603.
20. Rudwaleit M, van der Heijde D, Landewé R, et al. The development of Assessment of SpondyloArthritis International Society classification criteria for axial spondyloarthritis (part II): validation and final selection. Ann Rheum Dis 2009;68:777–83.
21. La cohorte DESIR. URL: http://www.lacohortedesir.fr/.
22. Zochling J. Measures of symptoms and disease status in ankylosing spondylitis: Ankylosing Spondylitis Disease Activity Score (ASDAS), Ankylosing Spondylitis Quality of Life Scale (ASQoL), Bath Ankylosing Spondylitis Disease Activity Index (BASDAI), Bath Ankylosing Spondylitis Functional Index (BASFI), Bath Ankylosing Spondylitis Global Score (BAS-G), Bath Ankylosing Spondylitis Metrology Index (BASMI), Dougados Functional Index (DFI), and Health Assessment Questionnaire for the Spondylarthropathies (HAQ-S). Arthritis Care Res (Hoboken) 2011;63 Suppl 11:S47–58.
23. Genolini C, Falissard B. KmL: a package to cluster longitudinal data. Comput Methods Programs Biomed 2011;104:e112–21.
24. Roussou E, Sultana S. Spondyloarthritis in women: differences in disease onset, clinical presentation, and Bath Ankylosing Spondylitis Disease Activity and Functional indices (BASDAI and BASFI) between men and women with spondyloarthritides. Clin Rheumatol 2011;30: 121–7.
25. Benattar L, Molto A, Gossec L, et al. Different disease activity trajectories in early axial spondyloarthritis lead to significantly different long-term outcomes: a cluster-based analysis of the DESIR Cohort [abstract]. Arthritis Rheumatol 2020;72 Suppl 10.
26. Ortega Castro R, Font Ugalde P, Casto Villegas MC, et al. Different clinical expression of patients with ankylosing spondylitis according to gender in relation to time since onset of disease: data from REGISPONSER. Reumatol Clin 2013;9:221–5.
27. Smolen JS, Schöls M, Braun J, et al. Treating axial spondyloarthritis and peripheral spondyloarthritis, especially psoriatic arthritis, to target: 2017 update of recommendations by an international task force. Ann Rheum Dis 2018;77:3–17.
28. Putrik P, Ramiro S, Moltó A, et al. Individual-level and country-level socioeconomic determinants of disease outcomes in SpA: multinational, cross-sectional study (ASAS-COMOSPA). Ann Rheum Dis 2019;78:486–93.
29. Rusman T, van Vollenhoven RF, van der Horst-Bruinsma IE. Gender differences in axial spondyloarthritis: women are not so lucky. Curr Rheumatol Rep 2018;20:35.
30. Duruöz MT, Gezer HH, Nas K, et al. Gender-related differences in disease activity and clinical features in patients with peripheral psoriatic arthritis: a multi-center study. Joint Bone Spine 2021;88:105177.
31. Alabas OA, Tashani OA, Tabasam G, et al. Gender role affects experimental pain responses: a systematic review with meta-analysis. Eur J Pain 2012;16:1211–23.
32. Moltó A, Etcheto A, Gossec L, et al. Evaluation of the impact of concomitant fibromyalgia on TNF alpha blockers' effectiveness in axial spondyloarthritis: results of a prospective, multicentre study. Ann Rheum Dis 2018;77:533–40.
33. Bello N, Etcheto A, Béal C, et al. Evaluation of the impact of fibromyalgia in disease activity and treatment effect in spondyloarthritis. Arthritis Res Ther 2016;18:42.

34. Baraliakos X, Tsiami S, Dukatz P, et al. Performance of standardized scores for disease assessment and pain in patients with spondyloarthritis and fibromyalgia. Ann Rheum Dis 2021;80 Suppl 1:344.
35. Aouad K, Ziade N, Baraliakos X. Structural progression in axial spondyloarthritis. Joint Bone Spine 2020;87:131–6.
36. Wright GC, Kaine J, Deodhar A. Understanding differences between men and women with axial spondyloarthritis. Semin Arthritis Rheum 2020;50:687–94.
37. Webers C, Essers I, Ramiro S, et al. Gender-attributable differences in outcome of ankylosing spondylitis: long-term results from the Outcome in Ankylosing Spondylitis International Study. Rheumatology (Oxford) 2016;55:419–28.
38. Lee W, Reveille JD, Davis JC Jr, et al. Are there gender differences in severity of ankylosing spondylitis? Results from the PSOAS cohort. Ann Rheum Dis 2007;66:633–8.
39. Dougados M, Sepriano A, Molto A, et al. Sacroiliac radiographic progression in recent onset axial spondyloarthritis: the 5-year data of the DESIR cohort. Ann Rheum Dis 2017;76:1823–8.
40. Alunno A, Carubbi F, Stones S, et al. The impact of fibromyalgia in spondyloarthritis: from classification criteria to outcome measures. Front Med (Lausanne) 2018;5:290.
41. Aouad K, Wendling D, Breban M, et al. Exploring remission concept in axial spondyloarthritis through the perception of rheumatologists using vignettes and priority ratings. Rheumatology (Oxford) 2022;30:2603–8.
42. Samulowitz A, Gremyr I, Eriksson E, et al. "Brave men" and "emotional women": a theory-guided literature review on gender bias in health care and gendered norms towards patients with chronic pain. Pain Res Manag 2018;2018:6358624.
43. Chung HY, Machado P, van der Heijde D, et al. HLA-B27 positive patients differ from HLA-B27 negative patients in clinical presentation and imaging: results from the DESIR cohort of patients with recent onset axial spondyloarthritis. Ann Rheum Dis 2011;70:1930–6.
44. Fish EN. The X-files in immunity: sex-based differences predispose immune responses. Nat Rev Immunol 2008;8:737–44.
45. Gracey E, Yao Y, Green B, et al. Sexual dimorphism in the Th17 signature of ankylosing spondylitis. Arthritis Rheumatol 2016;68:679–89.
46. Huang WN, Tso TK, Kuo YC, et al. Distinct impacts of syndesmophyte formation on male and female patients with ankylosing spondylitis. Int J Rheum Dis 2012;15:163–8.

Arthritis Care & Research
Vol. 75, No. 10, October 2023, pp 2117–2126
DOI 10.1002/acr.25135

American College of Rheumatology
Empowering Rheumatology Professionals

Factors Associated With Adherence to a Supervised Exercise Intervention for Osteoarthritis: Data From the Swedish Osteoarthritis Registry

Simone Battista,[1] Ali Kiadaliri,[2] Thérése Jönsson,[2] Kristin Gustafsson,[3] Martin Englund,[2] Marco Testa,[4] and Andrea Dell'Isola[2]

Objective. To explore how lifestyle and demographic, socioeconomic, and disease-related factors are associated with supervised exercise adherence in an osteoarthritis (OA) management program and the ability of these factors to explain exercise adherence.

Methods. A cohort register-based study on participants from the Swedish Osteoarthritis Registry who attended the exercise part of a nationwide Swedish OA management program. We ran a multinomial logistic regression to determine the association of exercise adherence with the abovementioned factors. We calculated their ability to explain exercise adherence with the McFadden R^2.

Results. Our sample comprises 19,750 participants (73% female, mean ± SD age 67 ± 8.9 years). Among them, 5,862 (30%) reached a low level of adherence, 3,947 (20%) a medium level, and 9,941 (50%) a high level. After a listwise deletion, the analysis was run on 16,685 participants (85%), with low levels of adherence as the reference category. Some factors were positively associated with high levels of adherence, such as older age (relative risk ratio [RRR] 1.01 [95% confidence interval (95% CI) 1.01–1.02] per year), and the arthritis-specific self-efficacy (RRR 1.04 [95% CI 1.02–1.07] per 10-point increase). Others were negatively associated with high levels of adherence, such as female sex (RRR 0.82 [95% CI 0.75–0.89]), having a medium (RRR 0.89 [95% CI 0.81–0.98] or a high level of education (RRR 0.84 [95% CI 0.76–0.94]). Nevertheless, the investigating factors could explain 1% of the variability in exercise adherence (R^2 = 0.012).

Conclusion. Despite the associations reported above, the poorly explained variability suggests that strategies based on lifestyle and demographic, socioeconomic, and disease-related factors are unlikely to improve exercise adherence significantly.

INTRODUCTION

In osteoarthritis (OA), exercise is considered a first-line intervention by international clinical practice guidelines (1,2) due to its ability to improve symptoms and levels of functionality (3,4). Exercise positively affects body weight, lipid metabolism, glycemic control, and systemic inflammation, preventing and treating OA-related chronic diseases (5). Despite these benefits, adherence to exercise in OA is suboptimal (6,7).

Adherence is described by the World Health Organization (WHO) as "the extent to which a person's behavior, taking medication, following a diet, and/or executing lifestyle changes, corresponds with agreed recommendations from a health care provider" (8). Poor adherence to exercise can severely compromise its long-term effectiveness, limiting its benefits (9). Considering the rising prevalence (10) and economic burden of OA (11), identifying factors associated with exercise adherence is fundamental to creating specific interventions to improve it.

Several elements have been hypothesized to be associated with exercise adherence, including lifestyle and demographic, socioeconomic, and disease-related factors (12–17). However, evidence on this topic arises mainly from other chronic conditions than OA, qualitative studies whose aims are not to generalize knowledge, as well as studies with small samples (12–19).

[1]Simone Battista, MSc: Lund University, Lund, Sweden, and University of Genoa, Campus of Savona, Savona, Italy; [2]Ali Kiadaliri, PhD, Thérése Jönsson, PhD, Martin Englund, PhD, Andrea Dell'Isola, PhD: Lund University, Lund, Sweden; [3]Kristin Gustafsson, MSc: Ryhov County Hospital Jönköping, Jönköping, and Linköping University, Linköping, Sweden; [4]Marco Testa, PhD: University of Genoa, Campus of Savona, Savona, Italy.

Author disclosures are available online at https://onlinelibrary.wiley.com/doi/10.1002/acr.25135.

Address correspondence via email to Marco Testa, PhD, marco.testa@unige.it.

Submitted for publication November 17, 2022; accepted in revised form April 13, 2023.

SIGNIFICANCE & INNOVATIONS

- Though exercise is a first-line intervention in osteoarthritis (OA), levels of exercise adherence among people with OA are suboptimal. Several elements have been hypothesized to be associated with exercise adherence, including lifestyle and demographic, socioeconomic, and disease-related factors in conditions other than OA.
- Analyzing real-world data from a first-line intervention provided nationwide in Swedish primary care, we found that high levels of adherence were positively associated with increased age, frequent pain, walking difficulties, and higher levels of self-efficacy. Conversely, high levels of adherence were negatively associated with female sex, higher body mass index, and high socioeconomic positions. However, these factors could explain 1% of the exercise variability.
- In OA, strategies based on lifestyle and demographic, socioeconomic, and disease-related factors are unlikely to improve exercise adherence significantly. Therefore, to improve adherence significantly, we need to consider other elements.

Moreover, the WHO has stated that the combination of different factors, rather than a single one, determines adherence (8). In contrast, the abovementioned studies focused primarily on single factors and their measures of mean association with adherence (e.g., odds ratio). Relying just on measures of association corresponds to an abstraction that does not take into account the variability of individual-level effects (20).

Therefore, we aimed to investigate the associations between lifestyle and demographic, socioeconomic, and disease-related factors with adherence to supervised exercise as a part of an OA management program delivered nationwide in Swedish primary care. Furthermore, we aimed to investigate these factors' ability to explain exercise adherence variability.

MATERIALS AND METHODS

Study design and setting. This study is a cohort register-based study on individual-level data retrieved from the Swedish Osteoarthritis Registry (SOAR; for data on the OA management program) and the Longitudinal Integration Database for Health Insurance and Labour Market Studies (LISA) administered by Statistics Sweden (for data on socioeconomic positions). These data sets were merged using personal identity numbers unique to all citizens in Sweden.

SOAR includes data from approximately 195,000 people with OA who attended an OA management program provided nationwide by the Swedish health care system (21,22). This program has already been thoroughly described elsewhere (23,24). Briefly, it is composed of 2 parts: education and exercise. The education part is mandatory, while the exercise part is optional. The education part is based on 3 sessions that revolve around the pathophysiology of the disease and its self-care management. The first 2 sessions are mandatory and held by a physiotherapist. The third is optional and held by a person with OA, trained as an OA communicator. The exercise (optional) part starts with an individual encounter with a physiotherapist to tailor the exercise program to the participants' needs and characteristics. At this point, participants can decide whether to exercise at home or with a physiotherapist. Those who decide to exercise with a physiotherapist are offered the opportunity to attend 12 sessions over 6 to 8 weeks (2 sessions/week) following OA Swedish clinical practice guidelines (25). LISA provides socioeconomic data such as cohabitation, institutionally based education level, employment, income, and residential area (26). The research was conducted in respect of the Declaration of Helsinki and reported following the Strengthening the Reporting of Observational studies in Epidemiology guidelines. Ethical approval was obtained from the Swedish Ethics Committee (Dnr: 2019-02570).

Population. The study cohort comprises all the participants in the SOAR with a first registration (baseline) between 2012 and 2015. We included only those who started the exercise group sessions supervised by the physiotherapists after the initial encounter with them. We selected participants with knee or hip OA who were recorded in the SOAR only once.

Variables. The level of adherence to the supervised exercise part, reported in the SOAR, is the dependent variable of this study. This is a predetermined categorical variable recorded by the physiotherapists and stratified on the number of sessions participants attended (low levels of adherence: 1–6 training sessions; medium levels of adherence: 7–9 training sessions; or high levels of adherence: 10–12 sessions). In this study, high levels of adherence represent >80% of the adherence with the recommended interventions (12 sessions) (25), which is typically considered a satisfactory level of adherence (27). The collected independent variables are reported hereafter and divided as demographic and lifestyle characteristics, socioeconomic characteristics, and disease-related characteristics.

Demographic and lifestyle characteristics. Participants' demographic and lifestyle characteristics were reported by the participants themselves at the baseline and recorded in the SOAR. These characteristics were assigned sex at birth (binary variable: male/female), age (continuous variable), body mass index (BMI; continuous variable computed from self-reported height and weight), weekly physical activity (continuous variable: hours) that was assessed with the question "How active are you during a regular, typical week?" (21), and health-related quality of life (HRQoL; continuous variable: EuroQol 5-domain instrument visual analog scale [EQ-5D VAS]). In the EQ-5D VAS,

the respondents reported their perceived HRQoL on a VAS scale that scored from 0 (the worst possible) to 100 (the best possible). The EQ-5D VAS is part of the EQ-5D scale (28).

Socioeconomic characteristics. Each socioeconomic position indicator from the year before the enrollment to the SOAR was considered for the analysis. In particular, the following socioeconomic position factors were retrieved and categorized: living alone (binary variable: living alone/living with someone), institutionally based education level (categorical variable: low [primary school: 0–9 years], medium [secondary school up to postsecondary education <3 years: 10–14 years], or high [postsecondary education: ≥15 years]), employment (binary variable: employed/retired-unemployed), residential area (categorical variable: rural/suburban/urban) and the net income.

Residential area was classified based on the Swedish Association of Local Authorities and Regions classification of Swedish municipalities. Specifically, rural areas are smaller towns/urban areas and rural municipalities, suburban areas are medium-sized towns (≥40,000 inhabitants) and municipalities near medium-sized towns, and urban areas are large cities (≥200,000 inhabitants) and municipalities near large cities (29). The individual yearly net income was categorized into quartiles based on the sample income distribution: lowest income quartile (<146,500 Swedish krona [SEK]), second income quartile (146,501–198,100 SEK), third income quartile (198,101–278,800 SEK), and highest income quartile (>278,800 SEK) (29).

Disease-related characteristics. The physiotherapists recorded the index joint (categorical variable: hip or knee) (21), namely, the joint with OA. They assessed this variable based on the participant's medical history, symptoms, and clinical assessment. In the case of multiple joints with OA, the most symptomatic joint was considered the index joint for the treatment. The participants self-recorded the numbers of painful joints (continuous variable); their desire for surgery (binary variable: yes/no) that was assessed by asking them: "Are your knee/hip symptoms so severe that you wish to undergo surgery?" (21); their pain intensity (ordinal variable: 0–10 on a numeric rating scale [NRS] [30]) in their index joint; their pain frequency (binary variable: infrequent pain [less than every week], frequent pain [almost every day]) that was assessed with the question: "How often do you have pain in your knee/hip" (21); their fear of movement (binary variable: yes/no) that was assessed with the question "Are you afraid your joints will be injured by physical training/activity?"; the Charnley score (categorical variable: A = unilateral hip or knee OA, B = bilateral hip or knee OA, C = multiple joint OA or some other condition) that categorizes people with OA into 3 classes based on the diseases that affect walking ability (31); and arthritis-specific self-efficacy (continuous variable: 10–100, pain and symptoms on the Arthritis Self-Efficacy Scale [ASES], using the Swedish version of the scale) (32). The ASES scale is a reliable instrument that assesses patients' arthritis-specific self-efficacy, namely, their beliefs about their ability to perform a specific task and cope with OA (33). The full version is composed of 3 subscales: 1) self-efficacy pain scale (5 items), 2) function scale (9 items), and 3) other symptoms scale (6 items). Participants indicate to what extent they feel confident they can do the tasks reported in the items from 10 (very uncertain) to 100 (very certain). In the SOAR, only 1) and 3) were adopted and combined as suggested in the scale instruction (33).

Statistical analysis. Descriptive statistics are reported as mean ± SD and absolute and percentage frequencies. A multivariable exploratory analysis was performed to identify which independent variables were independently associated with exercise adherence in the SOAR (34). Multivariable exploratory analyses detect patterns and identify relationships between the independent variables and the outcome (34–36).

Since the proportional odds assumption was not met, an ordered logistic regression could not be performed. Hence, we ran a multinomial logistic regression with a listwise deletion (Stata function mlogit) to determine the association between the independent variables and the adherence to exercise. No missing data were reported in the outcome (adherence). Less than 1% of the data on socioeconomic characteristics was missing, primarily due to an error during the data upload process in LISA. Missing data on demographic and lifestyle and disease-related characteristics in the SOAR are most likely a result of a mistake by the physiotherapists responsible for uploading the data at the local unit. Hence missing data in both registers could be considered missing completely at random, introducing no or minimal bias in our analysis.

The selection of the variables in the model was informed by previous literature on exercise adherence in other chronic pain conditions (12–17) and the evidence for action on adherence by the WHO (8). Then, the variables were clustered in demographic and lifestyle, socioeconomic, and disease-related groups, following the dimensions proposed by the WHO (8). The multicollinearity assumption between continuous variables was tested, and none of the continuous variables was highly correlated. The relative risk ratio (RRR) of being in medium level of adherence or high level of adherence with respect to low level of adherence and 95% confidence intervals (95% CIs) were estimated for each covariate in the model. For the variables HRQoL and arthritis-specific self-efficacy, the RRR is presented as a 10-point change in their scales.

Finally, the ability of the models to explain the variability of exercise adherence was calculated with the McFadden R^2 statistic (Stata function fitstat). McFadden R^2 measures the ability of a model to explain the variance of dependent variables on a convenient 0–100% scale. In particular, this value highlights how much of the variance in the dependent variable (adherence) can be explained by the independent variables collectively. We calculated McFadden R^2 for the model with all variables included (full model).

Afterward, we excluded 1 set of variables from the model and calculated the difference between McFadden R^2 with the full model. A higher difference would indicate a higher contribution of the variables set into the explanatory power of the full model. The analysis was done through Stata 17.

RESULTS

Between January 1, 2012 and December 31, 2015, 46,905 people with OA were recorded in the SOAR. However, we excluded 7 participants who had joints other than hip and knee as their first cause of pain, 27,147 who did not perform any supervised exercise sessions, and 1 for attending the program more than once. Hence, 19,750 participants with knee (69%) and hip (31%) OA were included in this study (73% female, mean ± SD age 67 ± 8.9 years). Figure 1 shows the participants' selection process. Table 1 presents the characteristics of the entire sample and stratified by the levels of adherence. Specifically, 5,862 (30%) reached a low level of adherence, 3,947 (20%) a medium level, and 9,941 (50%) a high level.

After the listwise deletion, the multinomial logistic regression was run on 16,685 individuals (85%), using low levels of adherence as the reference category (Table 2). Overall, excluded participants (n = 3,065) had similar characteristics to the ones included in the analysis (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25135). We found that female sex (RRR 1.13 [95% CI 1.02–1.27]), living with someone (RRR 1.21 [95% CI 1.10–1.32]), and an increase of 1 number of joints with OA (RRR 1.06 [95% CI 1.01–1.10]) were positively associated with achieving medium levels of adherence. Conversely, an increase in an hour of weekly physical activity (RRR 0.98 [95% CI 0.96–0.99]), living in an urban area (RRR 0.87 [95% CI 0.78–0.98]), and being employed (RRR 0.82 [95% CI 0.72–0.93]) were negatively associated with achieving medium levels of adherence.

An increase of 1 year in age (RRR 1.01 [95% CI 1.01–1.02]), having frequent pain (RRR 1.13 [95% CI 1.02–1.25]), having walking difficulties (RRR 1.12 [95% CI 1.01–1.24]), and having a 10-point increase on the ASES (RRR 1.04 [95% CI 1.02–1.07]) were positively associated with high levels of adherence. By contrast, female sex (RRR 0.82 [95% CI 0.75–0.89]), an increase of 1 point in BMI (RRR 0.99 [95% CI 0.98–0.99]), living in a suburban (RRR 0.79 [95% CI 0.73–0.86]) or an urban area (RRR 0.78 [95% CI 0.71–0.86]), being employed (RRR 0.71 [95% CI 0.64–0.78]),

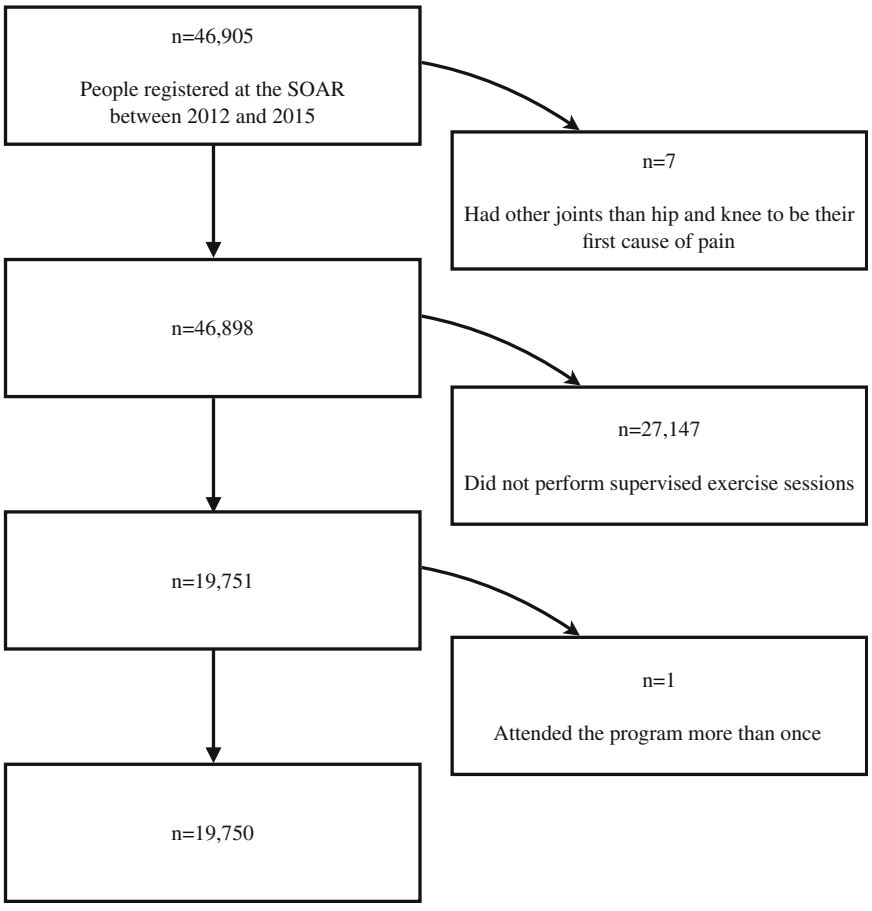

Figure 1. Selection of the study population. SOAR = Swedish Osteoarthritis Registry.

Table 1. Descriptive statistics*

Variables	Total sample (n = 19,750)	Level of adherence		
		Low (n = 5,862)	Medium (n = 3,947)	High (n = 9,941)
Demographic and lifestyle characteristics				
Assigned sex at birth	n = 19,750	n = 5,862	n = 3,947	n = 9,941
Male	5,421 (27.45)	1,519 (25.91)	925 (23.44)	2,977 (29.95)
Female	14,329 (72.55)	4,343 (74.09)	3,022 (76.65)	6,964 (70.05)
Age	n = 19,750	n = 5,862	n = 3,947	n = 9,941
Mean ± SD	66.86 ± 8.94	65.87 ± 9.39	66.47 ± 9.01	67.60 ± 8.57
Body mass index	n = 19,381	n = 5,735	n = 3,867	n = 9,779
Mean ± SD	27.56 ± 4.76	27.73 ± 4.90	27.75 ± 4.89	27.43 ± 4.63
HRQoL (EQ-5D VAS, 0–100)	n = 17,933	n = 5,317	n = 3,592	n = 9,024
Mean ± SD	65.82 ± 19.22	65.84 ± 19.37	65.74 ± 19.35	65.85 ± 19.07
Weekly physical activity, hours	n = 18,050	n = 5,364	n = 3,606	n = 9,080
Mean ± SD	4.11 ± 2.53	4.14 ± 2.53	4.03 ± 2.49	4.13 ± 2.54
Socioeconomic characteristics				
Institutionally based education level	n = 19,699	n = 5,862	n = 3,938	n = 9,918
Low	4,331 (21.99)	1,170 (20.02)	795 (20.19)	2,366 (23.86)
Medium	9,843 (49.97)	2,962 (50.69)	2,007 (50.96)	4,874 (49.14)
High	5,525 (28.05)	1,711 (29.28)	1,136 (28.85)	2,678 (27.00)
Income quartile	n = 19,738	n = 5,858	n = 3,945	n = 9,935
Lowest	4,942 (25.04)	1,345 (22.96)	1,022 (25.91)	2,575 (25.92)
Second	4,936 (25.01)	1,393 (23.78)	982 (24.89)	2,561 (25.78)
Third	4,929 (24.97)	1,517 (25.90)	976 (24.74)	2,436 (24.52)
Highest	4,931 (24.98)	1,603 (27.36)	965 (24.46)	2,363 (23.78)
Area of living	n = 19,738	n = 5,858	n = 3,945	n = 9,935
Rural	6,047 (30.64)	1,667 (28.46)	1,180 (29.91)	3,200 (32.21)
Suburban	8,252 (41.81)	2,435 (41.57)	1,708 (43.30)	4,109 (41.36)
Urban	5,439 (27.56)	1,756 (29.98)	1,057 (26.79)	2,626 (26.43)
Employment	n = 19,738	n = 5,858	n = 3,945	n = 9,935
Unemployed	12,244 (62.03)	3,275 (55.91)	2,394 (60.68)	6,575 (66.18)
Employed	7,494 (37.97)	2,583 (44.09)	1,551 (39.32)	3,360 (33.82)
Living alone	n = 19,738	n = 5,858	n = 3,945	n = 9,935
Living alone	7,754 (39.28)	2,411 (41.16)	1,457 (36.93)	3,886 (39.11)
Living with someone	11,984 (60.72)	3,447 (58.84)	2,488 (63.07)	6,049 (60.89)
Disease-related characteristics				
Worst joint	n = 19,750	n = 5,862	n = 3,947	n = 9,941
Hip	6,049 (30.63)	1,708 (29.14)	1,188 (30.10)	3,153 (31.72)
Knee	13,701 (69.37)	4,154 (70.86)	2,759 (69.90)	6,788 (68.28)
Pain intensity (NRS 0–10)	n = 19,686	n = 5,843	n = 3,935	n = 9,908
Mean ± SD	5.25 ± 1.83	5.23 ± 1.85	5.24 ± 1.87	5.26 ± 1.80
Pain frequency	n = 19,700	n = 5,842	n = 3,940	n = 9,918
Infrequent	3,436 (17.44)	1,100 (18.83)	723 (18.35)	1,613 (16.26)
Frequent	16,264 (82.56)	4,742 (81.17)	3,217 (81.65)	8,305 (87.34)
Number of painful joints	n = 19,750	n = 5,862	n = 3,947	n = 9,941
Mean ± SD	1.94 ± 1.29	1.95 ± 1.28	2.00 ± 1.32	1.91 ± 1.27
Charnley score	n = 19,735	n = 5,855	n = 3,946	n = 9,934
A	6,814 (34.53)	2,000 (34.16)	1,340 (33.96)	3,474 (34.97)
B	3,437 (17.42)	1,009 (17.23)	686 (17.38)	1,742 (17.54)
C	9,484 (48.06)	2,946 (48.61)	1,920 (48.66)	4,718 (47.49)
Walking difficulties	n = 19,651	n = 5,835	n = 3,932	n = 9,884
No	3,472 (17.67)	1,105 (18.94)	731 (18.59)	1,636 (16.55)
Yes	16,179 (82.33)	4,730 (81.06)	3,201 (81.41)	8,248 (83.45)
Fear of movement	n = 19,651	n = 5,821	n = 3,928	n = 9,902
No	16,562 (84.28)	4,871 (83.68)	3,303 (84.09)	8,388 (84.71)
Yes	3,089 (15.72)	950 (16.32)	625 (15.91)	1,514 (15.29)
Desire for surgery	n = 19,558	n = 5,798	n = 3,906	n = 9,854
No	14,936 (76.37)	4,441 (76.60)	3,017 (77.24)	7,478 (75.89)
Yes	4,622 (23.63)	1,357 (23.40)	889 (22.76)	2,376 (24.11)
ASES pain and symptoms (0–100)	n = 19,149	n = 5,660	n = 3,834	n = 9,655
Mean ± SD	65.54 ± 16.43	65.44 ± 16.54	65.51 ± 16.62	65.61 ± 16.28

* Values are the number (%) unless indicated otherwise. To calculate the missing values, subtract the number of participants listed in the second column (Total sample) from the total sample size of 19,750. ASES = Arthritis Self-Efficacy Scale; EQ-5D VAS = EuroQol 5-domain instrument visual analog scale; HRQoL = health-related quality of life; NRS = numeric rating scale.

Table 2. Association between exercise adherence and investigated factors (n = 16,685)*

Variables	*P*	RRR (95% CI for EXP[B])
Medium levels of adherence		
Assigned sex at birth		
Male (base category)	–	–
Female	0.03	1.13 (1.02–1.27)
Age	0.14	1.00 (0.99–1.01)
Body mass index	0.37	0.99 (0.99–1.01)
HRQoL (EQ-5D VAS, 0–100)†	0.57	0.99 (0.97–1.02)
Weekly physical activity, hours	0.02	0.98 (0.96–0.99)
Institutionally based education level		
Low (base category)	–	–
Medium	0.88	0.99 (0.88–1.12)
High	0.63	0.97 (0.84–1.11)
Income quartile		
Lowest (base category)	–	–
Second	0.71	0.98 (0.86–1.11)
Third	0.63	0.97 (0.84–1.11)
Highest	0.41	0.94 (0.81–1.09)
Area of living		
Rural (base category)	–	–
Suburban	0.27	0.94 (0.85–1.05)
Urban	0.02	0.87 (0.78–0.98)
Employment		
Unemployed (base category)	–	–
Employed	<0.01	0.82 (0.72–0.93)
Living alone		
Living alone (base category)	–	–
Living with someone	<0.01	1.21 (1.10–1.32)
Worst joint		
Hip (base category)	–	–
Knee	0.35	0.95 (0.86–1.05)
Pain intensity (NRS 0–10)	0.49	1.01 (0.98–1.04)
Pain frequency		
Infrequent (base category)	–	–
Frequent	0.80	0.98 (0.87–1.11)
Number of painful joints	0.01	1.06 (1.01–1.10)
Charnley score		
A (base category)	–	–
B	0.99	0.99 (0.97–1.15)
C	0.13	0.91 (0.81–1.03)
Walking difficulties		
No (base category)	–	–
Yes	0.93	0.99 (0.88–1.13)
Fear of movement		
No (base category)	–	–
Yes	0.49	1.04 (0.92–1.18)
Desire for surgery		
No (base category)	–	–
Yes	0.26	0.94 (0.83–1.05)
ASES pain and symptoms (0–100)†	0.29	1.02 (0.99–1.05)
High levels of adherence		
Assigned sex at birth		
Male (base category)	–	–
Female	<0.01	0.82 (0.75–0.89)
Age	<0.01	1.01 (1.01–1.02)
Body mass index	0.01	0.99 (0.98–0.99)
HRQoL (EQ-5D VAS, 0–100)†	0.18	0.98 (0.96–1.01)
Weekly physical activity, hours	0.79	0.99 (0.98–1.01)
Institutionally based education level		
Low (base category)	–	–
Medium	0.02	0.89 (0.81–0.98)
High	<0.01	0.84 (0.76–0.94)
Income quartile		
Lowest (base category)	–	–
Second	0.79	1.01 (0.91–1.13)
Third	0.61	1.03 (0.92–1.15)
Highest	0.95	1.00 (0.89–1.14)
Area of living		
Rural (base category)	–	–
Suburban	<0.01	0.79 (0.73–0.86)
Urban	<0.01	0.78 (0.71–0.86)
Employment		
Unemployed (base category)	–	–
Employed	<0.01	0.71 (0.64–0.78)
Living alone		
Living alone (base category)	–	–
Living with someone	0.29	1.04 (0.97–1.12)
Worst joint		
Hip (base category)	–	–
Knee	0.03	0.92 (0.85–0.99)
Pain intensity (NRS 0–10)	0.12	1.02 (0.99–1.04)
Pain frequency		
Infrequent (base category)	–	–
Frequent	0.02	1.13 (1.02–1.25)
Number of painful joints	0.50	1.01 (0.98–1.05)
Charnley score		
A (base category)	–	–
B	0.74	1.02 (0.91–1.14)
C	0.11	0.93 (0.84–1.02)
Walking difficulties		
No (base category)	–	–
Yes	0.03	1.12 (1.01–1.24)
Fear of movement		
No (base category)	–	–
Yes	0.93	1.00 (0.91–1.11)
Desire for surgery		
No (base category)	–	–
Yes	0.44	0.96 (0.88–1.06)
ASES pain and symptoms (0–100)†	<0.01	1.04 (1.02–1.07)

* Low levels of adherence are the reference. 95% CI = 95% confidence interval; ASES = Arthritis Self-Efficacy Scale; EQ-5D VAS = EuroQol 5-domain instrument visual analog scale; HRQoL = health-related quality of life; NRS = numeric rating scale; RRR = relative risk ratio.
† RRR is reported as an increase of 10 points in the scale.

having a medium (RRR 0.89 [95% CI 0.81–0.98]) or a high level of institutionally based education (RRR 0.84 [95% CI 0.76–0.94]), and having the knee as the worst joint (RRR 0.92 [95% CI 0.85–0.99]) were negatively associated with high levels of adherence.

Finally, the McFadden R^2 of the full model suggested that participants' demographic and lifestyle characteristics, socioeconomic characteristics, and disease-related characteristics can explain approximately 1.2% of the variation in adherence. After we removed participants' demographic and lifestyle characteristics, socioeconomic characteristics, and disease-related characteristics alternatively, there was a difference in the McFadden R^2 with respect to the full model of 0.3%, 0.4%, and 0.2%, respectively. Disease-related characteristics had the most

explanatory power, albeit the total explanatory ability of the full model was very small.

DISCUSSION

This study is the first to try to understand the relationship between demographic and lifestyle, socioeconomic, and disease-related factors, with the level of adherence to a face-to-face supervised exercise program for OA in a large sample of participants with this disease. Of the total sample, approximately 30% had low adherence levels, 20% had medium adherence levels, and 50% had high adherence levels. The distribution of adherence levels in our sample is consistent with that of participants in a similar Danish intervention (37) but differs from the distribution observed in an online version of the same intervention, which had a higher proportion of people with high levels of adherence than our sample (38). While several factors were associated with adherence, the full model could explain only 1% of the variability, which suggests that these factors are unlikely to have a tangible impact on adherence.

Regarding demographic and lifestyle factors, female sex was negatively associated with a high level of adherence. Previous evidence has indicated that women (with or without OA) might face societal expectations of household and caregiving responsibilities, experiencing greater difficulty finding time to exercise (39–43). However, in the digital version of this intervention, female sex suggested a positive association with high levels of exercise adherence (38), suggesting that digital interventions may be more convenient for females. Despite these findings, addressing the root causes of these disparities in exercise adherence is crucial, rather than focusing on exercise delivery mode to reduce this sex gap. However, our study only collected information on participants' assigned sex at birth, limiting the generalizability of our results to those individuals who are not cisgender. Therefore, further research is needed to explore the relationship between gender identity, sex, and exercise adherence in individuals with OA. In addition, participants' older age was positively associated with reaching a high level of adherence. Considering how exercise is delivered in this program, our result aligns with previous evidence where older adults adhered more to self-paced rather than moderate-intensity exercise (44). Finally, BMI was negatively associated with reaching high levels of adherence, which is consistent with previous evidence where people with high BMI are less keen on engaging in physical exercise (38,45).

Among the socioeconomic factors, people who lived in an urban or suburban area, were employed, and had medium or high levels of institutionally based education tended to exercise less than their counterparts. Similar results were found in the digital version of this intervention, where lower institutionally based education and living outside the largest Swedish cities were associated with higher adherence (38). These results contrast with the previous literature, where socioeconomic categories typically representing higher socioeconomic positions tended to adhere more to exercise (46,47). However, it is essential to consider that most of the data on adherence were retrieved from secondary analyses of randomized controlled trials (RCTs) (48). First, these studies were not designed to study adherence. RCTs per se tend to enhance adherence to treatment, which might create an overestimation of the factors related to adherence (49).

Second, in RCTs, people are volunteers who are selected following specific inclusion and exclusion criteria, which may fail to mirror the socioeconomic variability of the underlying population from which the sample is drawn (50). Moreover, we might not have reached the more socioeconomically disadvantaged groups, considering the higher socioeconomic positions of the SOAR sample compared to the general Swedish population (29). Finally, another explanation of this tendency is that people in lower socioeconomic positions seemed exposed to a more detrimental OA-disease burden than their higher counterparts (51). Severe symptoms can act as a motivator and drive exercise adherence (46,52). Those who experience a higher disease burden might be more motivated to follow exercise regimens. This phenomenon was also highlighted in our study when looking at the disease-related factors, as having frequent pain and walking difficulties were associated with high levels of adherence.

Moreover, self-efficacy was associated with exercise adherence, as per previous evidence (53), but with a modest RRR. Self-efficacy is characterized by a curvilinear (U-shaped) relationship between this construct and task accomplishment (54). People with low self-efficacy are likely to doubt their chance to accomplish a task, and those with a high-self efficacy might be characterized by complacency, inadequate preparation, and a focus on achieving task-related targets (54). Therefore, low and high levels of self-efficacy can lead to a similar outcome, namely, low adherence to a task (e.g., exercise). Considering the large cohort of our study, the effect of self-efficacy might be diluted due to the high variety of our population.

However, our model could explain just 1% of the variability, as indicated by the McFadden R^2. Thus, if we wanted to design an exercise intervention and understand which strategies to adopt to increase adherence, we should accept that demographic and lifestyle, socioeconomic, and disease-related factors are unlikely to improve adherence significantly, considering how little they explain adherence variability. This conclusion is further supported by the limited ability of similar factors to explain exercise adherence in the digital version of the intervention (38). Therefore, other factors should be taken into account.

The SOAR gathers real-world data from >500 different units throughout Sweden, with considerable variability among them. These contexts are characterized by specific contextual factors (e.g., structures' facilities, clinicians' communication style and ability to motivate patients, etc.) that affect people's outcomes via a placebo (or nocebo) response if positively (placebo) or negatively (nocebo) encoded by the brain via the so-called "mindsets" (55). Mindsets are "core assumptions about a domain or category that

orient individuals to a particular set of attributions, expectations, and goals" (56,57). Preliminary evidence indicated that improving mindsets about exercise increased its adherence (57). Moreover, booster sessions, reminders, and behavioral change techniques can improve exercise adherence by increasing motivation to partake in exercise (58,59). These strategies seem to ground their efficacy on contextual factors as well (e.g., communication with the clinicians, feeling taken care of by them, etc.). Therefore, we can argue that contextual factors and the mindsets responsible for interpreting them are worth exploring in future studies to understand their relationship with exercise adherence.

Some limitations of this study need to be discussed. First, the observational nature of the study does not allow us to establish causality and draw any definitive conclusion on the relationship between exercise adherence and the investigated factors. Second, a few variables were not reported. However, as explained in the methods section, the missingness of our data could be considered to be completely at random, primarily due to an error during the data upload process in the registers, introducing no or minimal bias in our results. However, we recommend interpreting our results cautiously, as we could not verify the reason for the data missingness. Third, our results might not be reliably applied to other forms of exercise (e.g., unsupervised home exercise) due to the specific research question of our study. Finally, physical activity hours, the number of painful joints, and living alone were found to be associated with medium but not high levels of adherence. However, this result may be influenced by chance and could also be attributed to the ad hoc adherence categorization adopted in the SOAR. Bearing in mind the limits of this study, it is worth highlighting that we reported the results of roughly 20,000 people with OA, followed by physiotherapists in the Swedish national health care system who tailored their intervention to patients' needs and characteristics. The size and data quality of our study strengthen its clinical importance and relevance for research.

To conclude, strategies based on demographic and lifestyle, socioeconomic, and disease-related factors are unlikely to improve exercise adherence significantly. Other elements, such as mindsets and contextual factors, need to be investigated. Moreover, as booster sessions, reminders, and behavioral-change techniques seem to improve exercise adherence (58,59), we should also understand how they motivate people to partake in exercise. Considering the complexity of adherence and the types of treatments that have succeeded in improving it so far, there is a call for solutions that go beyond a one-size-fits-all approach, to accept human variability and uncertainty, and to foster tailored interventions for individuals.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Battista, Kiadaliri, Jönsson, Dell'Isola.
Acquisition of data. Battista, Jönsson, Dell'Isola.
Analysis and interpretation of data. Battista, Kiadaliri, Gustafsson, Englund, Testa, Dell'Isola.

REFERENCES

1. Bannuru RR, Osani MC, Vaysbrot EE, et al. OARSI guidelines for the non-surgical management of knee, hip, and polyarticular osteoarthritis. Osteoarthritis Cartilage 2019;27:1578–89.
2. National Institute for Health Care Excellence. Osteoarthritis in over 16s: diagnosis and management. NICE Guideline 226. London: National Institute for Health Care Excellence; 2022.
3. Zeng CY, Zhang ZR, Tang ZM, et al. Benefits and mechanisms of exercise training for knee osteoarthritis. Front Physiol 2021;12:794062.
4. Raposo F, Ramos M, Lúcia Cruz A. Effects of exercise on knee osteoarthritis: a systematic review. Musculoskeletal Care 2021;19:399–435.
5. Skou ST, Pedersen BK, Abbott JH, et al. Physical activity and exercise therapy benefit more than just symptoms and impairments in people with hip and knee osteoarthritis. J Orthop Sports Phys Ther 2018;48:439–47.
6. De Groot IB, Bussmann JB, Stam HJ, et al. Actual everyday physical activity in patients with end-stage hip or knee osteoarthritis compared with healthy controls. Osteoarthritis Cartilage 2008;16:436–42.
7. Dunlop DD, Song J, Semanik PA, et al. Physical activity levels and functional performance in the Osteoarthritis Initiative: a graded relationship. Arthritis Rheum 2011;63:127–36.
8. World Health Organization. Adherence to long term therapies: evidence for action. Geneva: World Health Organization; 2003. URL: https://apps.who.int/iris/bitstream/handle/10665/42682/9241545992.pdf;jsessionid=159CA36674ED7E190879E9A5521BBC3F?sequence=1.
9. Pisters MF, Veenhof C, Schellevis FG, et al. Exercise adherence improving long-term patient outcome in patients with osteoarthritis of the hip and/or knee. Arthritis Care Res (Hoboken) 2010;62:1087–94.
10. Hawker GA. Osteoarthritis is a serious disease [review]. Clin Exp Rheumatol 2019;37 Suppl 120:3–6.
11. Safiri S, Kolahi AA, Smith E, et al. Global, regional and national burden of osteoarthritis 1990-2017: a systematic analysis of the Global Burden of Disease Study 2017. Ann Rheum Dis 2020;79:819–28.
12. Jefferis BJ, Sartini C, Lee IM, et al. Adherence to physical activity guidelines in older adults, using objectively measured physical activity in a population-based study. BMC Public Health 2014;14:382.
13. Hemmingsson E, Ekelund U. Is the association between physical activity and body mass index obesity dependent? Int J Obes (Lond) 2007;31:663–8.
14. Marmot M. Social determinants of health inequalities. Lancet 2005;365:1099–104.
15. Austin S, Qu H, Shewchuk RM. Association between adherence to physical activity guidelines and health-related quality of life among individuals with physician-diagnosed arthritis. Qual Life Res 2012;21:1347–57.
16. Hunter DJ, Schofield D, Callander E. The individual and socioeconomic impact of osteoarthritis [review]. Nat Rev Rheumatol 2014;10:437–41.
17. Taulaniemi A, Kankaanpää M, Rinne M, et al. Fear-avoidance beliefs are associated with exercise adherence: secondary analysis of a randomised controlled trial (RCT) among female healthcare workers with recurrent low back pain. BMC Sports Sci Med Rehabil 2020;12:28.

18. Battista S, Manoni M, Dell'Isola A, et al. Giving an account of patients' experience: a qualitative study on the care process of hip and knee osteoarthritis. Health Expect 2022;25:1140–56.
19. Battista S, Dell'Isola A, Manoni M, et al. Experience of the COVID-19 pandemic as lived by patients with hip and knee osteoarthritis: an Italian qualitative study. BMJ Open 2021;11:e053194.
20. Merlo J, Mulinari S, Wemrell M, et al. The tyranny of the averages and the indiscriminate use of risk factors in public health: the case of coronary heart disease. SSM Popul Health 2017;3:684–98.
21. Dell'Isola A, Jönsson T, Nero H, et al. Factors associated with the outcome of a first-line intervention for patients with hip or knee osteoarthritis or both: data from the BOA register. Phys Ther 2020;100:1771–81.
22. Svenska Artrosregistret. The Swedish Osteoarthritis registry. URL: https://boa.registercentrum.se/in-english/the-swedish-osteoarthritis-registry/p/By_o8GxVg.
23. Jönsson T, Eek F, Dell'Isola A, et al. The Better Management of Patients with Osteoarthritis Program: outcomes after evidence-based education and exercise delivered nationwide in Sweden. PLoS One 2019;14:e0222657.
24. Dell'Isola A, Jönsson T, Ranstam J, et al. Education, home exercise, and supervised exercise for people with hip and knee osteoarthritis as part of a nationwide implementation program: data from the Better Management of Patients With Osteoarthritis Registry. Arthritis Care Res (Hoboken) 2020;72:201–7.
25. Socialstyrelsen. Nationella riktlinjer för rörelseorganens sjukdomar: reumatoid artrit, axial spondylartrit, psoriasisartrit, artros och osteoporos. 2021. URL: https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/artikelkatalog/nationella-riktlinjer/2021-1-7137.pdf.
26. Statistics Sweden. Longitudinal integration database for health insurance and labour market studies (LISA). URL: https://pubmed.ncbi.nlm.nih.gov/30929112/.
27. Bailey DL, Holden MA, Foster NE, et al. Defining adherence to therapeutic exercise for musculoskeletal pain: a systematic review. Br J Sports Med 2020;54:326–31.
28. Balestroni G, Bertolotti G. EuroQol-5D (EQ-5D): an instrument for measuring quality of life. Monaldi Arch Chest Dis 2012;78:155–9. In Italian.
29. Gustafsson K, Kvist J, Eriksson M, et al. Socioeconomic status of patients in a Swedish national self-management program for osteoarthritis compared with the general population: a descriptive observational study. BMC Musculoskelet Disord 2020;21:10.
30. Hawker GA, Mian S, Kendzerska T, et al. Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). Arthritis Care Res (Hoboken) 2011;63:S240–52.
31. Bjorgul K, Novicoff WM, Saleh KJ, et al. Evaluating comorbidities in total hip and knee arthroplasty: available instruments [review]. J Orthop Traumatol 2010;11:203–9.
32. Lomi C. Evaluation of a Swedish version of the Arthritis Self-Efficacy Scale. Scand J Caring Sci 1992;6:131–8.
33. Lorig K, Chastain RL, Ung E, et al. Development and evaluation of a scale to measure perceived self-efficacy in people with arthritis. Arthritis Rheum 1989;32:37–44.
34. Valveny N, Gilliver S. How to interpret and report the results from multivariable analyses. Med Writ 2016;25:37–42.
35. Neely JG, Paniello RC, Cho Lieu JE, et al. Practical guide to understanding multivariable analyses: part A. Otolaryngol Head Neck Surg 2013;148:185–90.
36. Komorowski M, Marshall DC, Salciccioli JD, et al. Exploratory data analysis. In: Secondary analysis of electronic health records. Springer; 2016. p. 185–203.
37. Skou ST, Roos EM. Good Life with osteoArthritis in Denmark (GLA: DTM): evidence-based education and supervised neuromuscular exercise delivered by certified physiotherapists nationwide. BMC Musculoskelet Disord 2017;18:72.
38. Kiadaliri A, Dell'isola A, Lohmander LS, et al. Assessing the importance of predictors of adherence to a digital self-management intervention for osteoarthritis. J Orthop Surg Res 2023;18:97.
39. Gommans LN, Scheltinga MR, van Sambeek MR, et al. Gender differences following supervised exercise therapy in patients with intermittent claudication. J Vasc Surg 2015;62:681–8.
40. Oosenbrug E, Marinho RP, Zhang J, et al. Sex differences in cardiac rehabilitation adherence: a meta-analysis [review]. Can J Cardiol 2016;32:1316–24.
41. McArthur D, Dumas A, Woodend K, et al. Factors influencing adherence to regular exercise in middle-aged women: a qualitative study to inform clinical practice. BMC Womens Health 2014;14:49.
42. Verhoef M, Love EJ, Rose SA. Women's social roles and their exercise participation. Women Health 1992;19:15–29.
43. Cheung C, Wyman JF, Peden-McAlpine C. Long-term yoga and aerobic/strength exercise adherence in older women with knee osteoarthritis: a mixed methods approach. Int J Yoga Therap 2022;32:article 4.
44. Lee HH, Dunsiger S, Connell Bohlen L, et al. Age moderates the effect of self-paced exercise on exercise adherence among overweight adults. J Aging Health 2020;32:154–61.
45. Baillot A, Chenail S, Polita NB, et al. Physical activity motives, barriers, and preferences in people with obesity: a systematic review. PLoS One 2021;16:e0253114.
46. Tuakli-Wosornu YA, Selzer F, Losina E, et al. Predictors of exercise adherence in patients with meniscal tear and osteoarthritis. Arch Phys Med Rehabil 2016;97:1945–52.
47. Areerak K, Waongenngarm P, Janwantanakul P. Factors associated with exercise adherence to prevent or treat neck and low back pain: a systematic review. Musculoskelet Sci Pract 2021;52:102333.
48. Kahan BC, Rehal S, Cro S. Risk of selection bias in randomised trials [review]. Trials 2015;16:405.
49. Van Onzenoort HA, Menger FE, Neef C, et al. Participation in a clinical trial enhances adherence and persistence to treatment: a retrospective cohort study. Hypertension 2011;58:573–8.
50. Collado-Mateo D, Lavín-Pérez AM, Peñacoba C, et al. Key factors associated with adherence to physical exercise in patients with chronic diseases and older adults: an umbrella review. Int J Environ Res Public Health 2021;18:2023.
51. Unevik E, Abbott A, Fors S, et al. Sociodemographic inequalities in outcomes of a Swedish nationwide self-management program for osteoarthritis: results from 22,741 patients between years 2008–2017. J Clin Med 2020;9:2294.
52. Essery R, Geraghty AW, Kirby S, et al. Predictors of adherence to home-based physical therapies: a systematic review. Disabil Rehabil 2017;39:519–34.
53. Daniali SS, Darani FM, Eslami AA, et al. Relationship between self-efficacy and physical activity, medication adherence in chronic disease patients. Adv Biomed Res 2017;6:63.
54. Bachrach DG, Rapp TL, Rapp AA, et al. "Too much" self-efficacy? Understanding the curvilinear consequences of between-person self-efficacy through a moderated-mediation model of perceived proximity and employee effort. Group Organ Manag 2022;2022:1–38.

55. Testa M, Rossettini G. Enhance placebo, avoid nocebo: how contextual factors affect physiotherapy outcomes [review]. Man Ther 2016; 24:65–74.
56. Crum AJ, Salovey P, Achor S. Rethinking stress: the role of mindsets in determining the stress response. J Pers Soc Psychol 2013;104: 716–33.
57. Boswell MA, Evans KM, Zion SR, et al. Mindset is associated with future physical activity and management strategies in individuals with knee osteoarthritis. Ann Phys Rehabil Med 2022;65:101634.
58. Cinthuja P, Krishnamoorthy N, Shivapatham G. Effective interventions to improve long-term physiotherapy exercise adherence among patients with lower limb osteoarthritis: a systematic review. BMC Musculoskelet Disord 2022;23:1–11.
59. Willett M, Duda J, Gautrey C, et al. Effectiveness of behavioural change techniques in physiotherapy interventions to promote physical activity adherence in patients with hip and knee osteoarthritis: a systematic review protocol. BMJ Open 2017;7: e015833.

Arthritis Care & Research
Vol. 75, No. 10, October 2023, pp 2127–2133
DOI 10.1002/acr.25125

American College of Rheumatology
Empowering Rheumatology Professionals

Neuropathic Pain Associated With First Metatarsophalangeal Joint Osteoarthritis: Frequency and Associated Factors

Hylton B. Menz,[1] Jamie J. Allan,[1] Andrew K. Buldt,[1] Karl B. Landorf,[1] Flavia M. Cicuttini,[2] Edward Roddy,[3] and Shannon E. Munteanu[1]

Objective. To determine whether neuropathic pain is a feature of first metatarsophalangeal (MTP) joint osteoarthritis (OA).

Methods. A total of 98 participants (mean ± SD age 57.4 ± 10.3 years) with symptomatic radiographic first MTP joint OA completed the PainDETECT questionnaire (PD-Q), which has 9 questions regarding the intensity and quality of pain. The likelihood of neuropathic pain was determined using established PD-Q cutoff points. Participants with unlikely neuropathic pain were then compared to those with possible/likely neuropathic pain in relation to age, sex, general health (Short Form 12 [SF-12] health survey), psychological well-being (Depression, Anxiety and Stress Scale), pain characteristics (self-efficacy, duration, and severity), foot health (Foot Health Status Questionnaire [FHSQ]), first MTP dorsiflexion range of motion, and radiographic severity. Effect sizes (Cohen's *d* coefficient) were also calculated.

Results. A total of 30 (31%) participants had possible/likely neuropathic pain (19 possible [19.4%], 11 likely [11.2%]). The most common neuropathic symptoms were sensitivity to pressure (56%), sudden pain attacks/electric shocks (36%) and burning (24%). Compared to those with unlikely neuropathic pain, those with possible/likely neuropathic pain were significantly older (d = 0.59, P = 0.010), had worse SF-12 physical scores (d = 1.10, $P < 0.001$), pain self-efficacy scores (d = 0.98, $P < 0.001$), FHSQ pain scores (d = 0.98, $P < 0.001$), and FHSQ function scores (d = 0.82, $P < 0.001$), and had higher pain severity at rest (d = 1.01, $P < 0.001$).

Conclusion. A significant proportion of individuals with first MTP joint OA report symptoms suggestive of neuropathic pain, which may partly explain the suboptimal responses to commonly used treatments for this condition. Screening for neuropathic pain may be useful in the selection of targeted interventions and may improve clinical outcomes.

INTRODUCTION

Pain is the most common and disabling symptom of osteoarthritis (OA) and has primarily been attributed to local tissue damage leading to mechanical and/or inflammatory stimulation of peripheral sensory neurons (nociceptors) in joint tissue (1). However, the suboptimal and variable response to treatment of OA-related pain has led to reappraisal of its underlying cause, and the contribution of non-nociceptive pathways is increasingly recognized (2,3). In particular, neuropathic pain, defined as pain arising as a direct consequence of a lesion or disease affecting the somatosensory system (4), may be responsible for symptoms such as tingling, numbness, burning, and electrical shock sensations (5), which are experienced by one-third of individuals with knee or hip OA (6).

The presence of neuropathic symptoms increases the individual burden of knee OA, as it is associated with more severe pain (7–10), greater impairment in physical function (9–13), worse quality of life (10,11,13,14), and poorer sleep quality (10). Several person-level factors are associated with neuropathic pain in individuals with knee OA, including increased age (13), higher body mass index (BMI) (13), female sex (8), multiple comorbidities (8), pain at multiple sites (7,12), referred pain (7), and hyperalgesia

Supported by the National Health and Medical Research Council of Australia (grant 1105244). Dr. Menz's work is supported by a National Health and Medical Research Council Senior Research fellowship (no. 1135995).

[1]Hylton B. Menz, DSc, Jamie J. Allan, MPodPrac, Andrew K. Buldt, PhD, Karl B. Landorf, PhD, Shannon E. Munteanu, PhD: La Trobe University, Melbourne, Victoria, Australia; [2]Flavia M. Cicuttini, PhD: Monash University and Alfred Hospital, Melbourne, Victoria, Australia; [3]Edward Roddy, DM: Keele University, Keele, Staffordshire, UK, and Haywood Hospital, Burslem, Staffordshire, UK.

Author disclosures and a graphical abstract are available online at https://onlinelibrary.wiley.com/doi/10.1002/acr.25125.

Address correspondence via email to Hylton B. Menz, DSc, at h.menz@latrobe.edu.au.

Submitted for publication November 15, 2022; accepted in revised form March 30, 2023.

SIGNIFICANCE & INNOVATIONS

- This is the first study to evaluate neuropathic pain in individuals with foot osteoarthritis (OA).
- One in 3 individuals with first metatarsophalangeal joint OA had evidence of possible or likely neuropathic pain.
- Those with neuropathic pain were older, had worse general physical health, worse foot health, and greater pain severity at rest.
- Screening for neuropathic pain may be useful in the selection of appropriate interventions.

(9). Knee joint–specific correlations with neuropathic pain include meniscal lesions (15) and prior surgery (10), although reported associations with radiographic severity are inconsistent (12,13,16).

To the best of our knowledge, no studies have examined neuropathic pain related to OA affecting the joints of the foot. This is important since foot OA has a similar prevalence compared to knee OA (17), is considered disabling in 75% of patients (17), and is a common reason for consultation in primary care (18). Foot OA most commonly affects the first metatarsophalangeal (MTP) joint and is characterized by the formation of a dorsal exostosis (19), limited range of motion (20), and altered walking patterns (21). Interventions such as footwear and orthoses have been shown to alter the biomechanical function of the foot in individuals with first MTP joint OA (22,23) but show only modest reductions in pain (24,25), suggesting that non-mechanical factors may contribute to symptoms.

Therefore, the objectives of this study were to determine whether neuropathic pain is a feature of first MTP joint OA and to explore person- and foot-level factors associated with the presence of neuropathic pain in participants enrolled in a recent randomized clinical trial.

PATIENTS AND METHODS

Participants. Participants for this study were recruited from a randomized trial that evaluated the effectiveness of shoe-stiffening inserts for first MTP joint OA. Full details of the trial have been described previously (25,26). Participants were recruited using advertisements in local newspapers, posters placed in senior citizens' centers and retirement villages, mailed advertisements to health care practitioners, mailed advertisements to individuals currently accessing podiatry services at the La Trobe University Health Sciences Clinic, and through social media. To be included in the trial, participants needed to be ≥18 years old, have pain in the first MTP joint on most days for at least 12 weeks, rated ≥30 mm on a 100-mm visual analog scale, have pain upon palpation of the dorsal aspect of the first MTP joint, have restricted first MTP joint dorsiflexion, and be able to walk household distances without the use of a walking aid. Participants were excluded if they had previous first MTP joint surgery, were currently pregnant, or had hallux valgus, a systemic inflammatory condition, or cognitive impairment.

Ethical approval was provided by the La Trobe University Human Ethics Committee (approval no. HEC15-128), and written informed consent was obtained from all participants. In this study, the sample size was determined by the requirements of the randomized trial, which was powered to detect a minimum clinically important difference in the primary outcome measure, the Foot Health Status Questionnaire (FHSQ) pain subscale (25,26).

Demographic, general health, and pain assessments. A structured questionnaire was used to collect data regarding participant demographic characteristics (age and sex), general health (the Short Form 12 questionnaire [27]), psychological well-being (the Depression, Anxiety and Stress Scale [28]), pain characteristics (including the Pain Self-Efficacy Questionnaire [PSEQ] [29], pain duration, and pain severity at rest and while walking [26]), and foot health (the FHSQ pain and function subscales [30]). Only baseline data were used in this analysis.

Clinical and radiographic assessments. Height and weight were measured using a stadiometer and digital scales, and BMI was calculated as weight/height (kg/m^2). Clinical features associated with first MTP joint OA (pain on palpation, a dorsal exostosis, joint effusion, pain on motion, hard end-feel, and crepitus) and passive, non–weight-bearing first MTP joint dorsiflexion range of motion were documented using established techniques (19). The presence of radiographic first MTP joint OA was determined using the La Trobe University Radiographic Atlas, which uses weight-bearing dorsiplantar and lateral radiographs to document the presence of OA based on the observation of osteophytes and joint space narrowing (JSN) (31). Radiographic OA was documented as present or absent based on the La Trobe University Radiographic Atlas case definition (at least one score of 2 for osteophytes or JSN on either the dorsiplantar or lateral view) (32), and radiographic OA severity was documented as mild (no scores for osteophytes or JSN on either view >1), moderate (at least one score of 2 but none >2), or severe (at least one score of 3) (20).

Neuropathic pain assessment. To document the presence of neuropathic pain affecting the first MTP joint, we used the self-reported PainDETECT Questionnaire (PD-Q), which was originally developed to discriminate between nociceptive pain and neuropathic pain in individuals with chronic low back pain (33). The PD-Q comprises 7 items evaluating pain quality (scores from 0 to 5), 1 item evaluating pain pattern (scores from –1 to 1), and 1 item evaluating pain radiation (scores from 0 to 2). The sum of individual question scores was used to calculate a total

score ranging from −1 to 38. Total scores <13 indicate that neuropathic pain is unlikely, scores from 13 to 18 indicate that neuropathic pain is possible, and scores >18 indicate that neuropathic pain is likely (34). The PD-Q has been validated against expert physician diagnosis of neuropathic pain in individuals with low back pain (33) and against quantitative sensory testing for the detection of central sensitization in individuals with knee OA (35).

Statistical analysis. Statistical analysis was undertaken using IBM SPSS Statistics version 26.0. All data were explored for normality and did not require transformation. For continuously scored variables, differences between participants with and those without neuropathic pain were compared using independent samples *t*-tests and effect sizes (Cohen's *d* coefficient). The following interpretation of effect sizes was used: ≤0.01 indicates very small, >0.01 to 0.20 indicates small, >0.20 to 0.50 indicates medium, >0.50 to 0.8 indicates large, >0.80 to 1.2 indicates very large, and >1.20 indicates huge (36). For dichotomous or ordinal variables, differences between groups were calculated using the chi-square statistic.

RESULTS

Participants. A total of 100 participants were recruited for the randomized trial (25). Of these, 98 participants had complete PD-Q data and were included in this analysis (44 men and 54 women, mean ± SD age 57.3 ± 10.3 years). Participant characteristics are shown in Table 1. Data were missing for the following variables: height, weight, and BMI (n = 3), dorsiplantar radiographs (n = 5), and lateral radiographs (n = 6).

Neuropathic pain characteristics. PD-Q responses are shown in Table 2. A total of 69 of the 98 participants (70%) reported at least 1 neuropathic symptom with at least moderate severity, with the most common neuropathic symptoms being sensitivity to pressure (n = 55 [56%]), sudden pain attacks/electric shocks (n = 35 [36%]), and burning (n = 24 [25%]). A total of 37 participants (37.8%) reported pain radiation. Thirty (31%) participants had possible/likely neuropathic pain (n = 19 [19.4%], n = 11 [11.2%], for possible neuropathic pain and likely neuropathic pain, respectively), as defined according to overall PD-Q score.

Differences between participants with and those without neuropathic pain. Participant characteristics in those with and those without neuropathic pain are shown in Table 3. Compared to those with unlikely neuropathic pain, those with possible/likely neuropathic pain were significantly older (d = 0.59, P = 0.010; large effect size), had worse scores on the questionnaires for SF-12 physical function (d = 1.10, P < 0.001; very large effect size), PSEQ (d = 0.98, P < 0.001; very large effect

Table 1. Demographic and clinical characteristics of 98 participants with radiographic first MTP joint OA*

Characteristics	Values
Demographic characteristics and anthropometrics	
Age, mean ± SD years	57.3 ± 10.3
Female sex	54 (55.1)
Height, mean ± SD cm	168.3 ± 8.2
Weight, mean ± SD kg	79.4 ± 13.0
BMI, mean ± SD kg/m^2	28.1 ± 4.6
Clinical features	
Passive, non–weight-bearing first MTP joint maximum dorsiflexion, mean ± SD degrees	45.3 ± 11.2
Pain on palpation	98 (100.0)
Palpable dorsal exostosis	97 (99.0)
Pain on motion of first MTP joint	74 (75.5)
Hard end-feel when dorsiflexed	92 (93.9)
Crepitus	21 (21.4)
Radiographic first MTP joint OA†	84 (90.3)
Radiographic severity‡	
Mild	9 (9.7)
Moderate	38 (40.9)
Severe	46 (49.5)

* Except where indicated otherwise, values are the number (%). MTP = metatarsophalangeal; OA = osteoarthritis.
† Indicates at least one score of 2 for osteophytes or joint space narrowing on either view using the case definition from the La Trobe Radiographic Atlas (31).
‡ Mild indicates no scores >1; moderate indicates at least one score of 2 but none >2; severe: at least one score of 3 for osteophytes or joint space narrowing on either view, using the La Trobe Radiographic Atlas (31).

Table 2. PainDETECT responses in 98 participants with radiographic first MTP joint OA*

Characteristics	Values
Pain severity, mean ± SD (score 0–10)	
How would you assess your pain now, at this moment?	3.76 ± 2.34
How strong was the strongest pain during the past 4 weeks?	7.03 ± 2.02
How strong was the pain during the past 4 weeks on average?	4.96 ± 1.86
Pain pattern	
Persistent pain with slight variations	32 (32.7)
Persistent pain with pain attacks	33 (33.7)
Pain attacks without pain between them	25 (25.5)
Pain attacks with pain between them	8 (8.2)
Pain radiation	37 (37.8)
Pain quality, moderate or more (score ≥3 [of 5])	
Burning	24 (24.5)
Tingling or prickling	14 (14.3)
Sensitivity to light touch	18 (18.4)
Sudden pain attacks/electric shocks	35 (35.7)
Sensitivity to cold or heat	10 (10.2)
Numbness	12 (12.2)
Sensitivity to pressure	55 (56.1)
Total PainDETECT score, mean ± SD (score 0–38)†	10.7 ± 5.5
Neuropathic pain unlikely	68 (69.4)
Neuropathic pain possible	19 (19.4)
Neuropathic pain likely	11 (11.2)

* Except where indicated otherwise, values are the number (%) of participants. See Table 1 for definitions.
† Total scores <13 indicate that neuropathic pain is unlikely; scores of 13–18 indicate that neuropathic pain is possible; and scores >18 indicate that neuropathic pain is likely (34).

Table 3. Demographic and clinical characteristics in OA participants with and those without neuropathic pain affecting the first MTP joint*

Characteristics	Non-neuropathic (n = 68)	Neuropathic (n = 30)	*d*	*P*
Demographic characteristics and anthropometrics				
Age, mean ± SD years	55.5 ± 11.0	61.3 ± 7.1	0.59	0.003
Female sex	34 (50.0)	20 (66.7)	–	0.186
BMI, mean ± SD kg/m^2	27.6 ± 4.6	29.3 ± 4.5	0.38	0.092
General health (SF-12 scores)†				
Physical	49.2 ± 8.0	39.9 ± 9.7	1.10	<0.001
Mental	53.7 ± 9.4	52.5 ± 8.7	0.13	0.543
Psychological well-being (DASS-21)‡				
Depression	2.9 ± 5.9	4.9 ± 5.7	0.35	0.118
Anxiety	3.2 ± 5.3	3.6 ± 4.9	0.08	0.723
Stress	7.3 ± 7.4	9.3 ± 8.8	0.26	0.287
Pain characteristics				
PSEQ§	54.1 ± 6.0	47.0 ± 9.7	0.98	0.001
Pain duration, months	39 ± 47	60 ± 92	0.35	0.055
Pain severity at rest, VAS¶	2.4 ± 1.6	4.1 ± 1.9	1.01	<0.001
Pain severity while walking, VAS¶	5.0 ± 1.5	5.5 ± 1.6	0.33	0.164
Foot health (FHSQ scores)#				
Pain	51.9 ± 16.1	37.0 ± 13.4	0.98	<0.001
Function	71.7 ± 21.6	53.8 ± 23.1	0.82	<0.001
Clinical features				
Passive non–weight-bearing first MTP joint maximum dorsiflexion, mean ± SD degrees	46.6 ± 10.1	42.2 ± 13.1	0.40	0.108
Pain on palpation	68 (100.0)	30 (100.0)	–	NC
Palpable dorsal exostosis	67 (98.5)	30 (100.0)	–	0.504
Pain on motion of first MTP joint	48 (70.6)	26 (86.7)	–	0.088
Hard end-feel when dorsiflexed	64 (94.1)	28 (93.3)	–	0.881
Crepitus	14 (20.6)	7 (23.3)	–	0.760
Radiographic first MTP joint OA**	59 (90.8)	25 (89.3)	–	0.546
Radiographic severity††				
Mild	6 (9.2)	3 (10.7)	–	0.965
Moderate	27 (41.5)	11 (39.3)	–	–
Severe	32 (49.2)	14 (50.0)	–	–

* Except where indicated otherwise, values are the number (%) of participants. MTP = metatarsophalangeal; NC = not calculable; OA = osteoarthritis.
† For short Form 12 (SF-12) scores, scores ranged from 0 to 100, with higher scores indicating better function.
‡ For 21-item Depression, Anxiety and Stress Scale (DASS-21) scores, scores ranged from 0 to 42, with higher scores indicating worse function.
§ For Pain Self-Efficacy (PSEQ) questionnaire scores, score ranged from 0 to 60, with higher scores indicating greater confidence dealing with pain.
¶ For visual analog scale (VAS) scores, score ranged from 0 to 10, with higher scores indicating worse pain.
# For Foot Health Status Questionnaire (FHSQ)scores, score ranged from 0 to100, with higher scores indicating better function.
** At least one score of 2 for osteophytes or joint space narrowing on either view using the case definition from the La Trobe Radiographic Atlas (31).
†† Mild indicates no scores >1; moderate indicates at least one score of 2 but none >2; severe indicates at least one score of 3 for osteophytes or joint space narrowing on either view, using the La Trobe Radiographic Atlas (31).

size), FHSQ pain (d = 0.98, P < 0.001; very large effect size), and FHSQ function (d = 0.82, P < 0.001; very large effect size), and had higher pain severity at rest (d = 1.01, P < 0.001; very large effect size).

DISCUSSION

In this study, we aimed to determine whether neuropathic pain is a feature of foot OA by using the PD-Q in OA participants with first MTP joint OA who were enrolled in a randomized trial. We found that 70% of participants reported ≥1 moderate symptom indicative of neuropathic pain (such as electric shocks, burning, numbness, and tingling), and that the prevalence of possible/likely neuropathic pain in this group using the established overall PD-Q cutoff score was 31%. Those with possible/likely neuropathic pain were older, had worse general physical health, worse foot health, and greater pain severity at rest. To the best of our knowledge, this study provides the first insights into neuropathic pain related to foot OA.

The prevalence of neuropathic pain observed in this study is similar to previous reports in individuals with knee and hip OA. A systematic review and meta-analysis of 39 studies (36 involving the knee and 3 involving the hip) showed a pooled prevalence estimate of 40% (95% confidence interval [95% CI] 32–48) in knee OA and 29% (95% CI 22–37) in hip OA, using the same case definition of possible/likely neuropathic pain from the PD-Q (6). The prevalence of reporting individual neuropathic symptoms was also high in our study, with 70% reporting ≥1 neuropathic

symptom with at least moderate severity. The most frequently reported symptoms—sensitivity to pressure and sudden electric shocks—are hallmark features of neuropathic pain and are believed to result from central sensitization and spontaneous firing of peripheral nociceptors, respectively (37).

We observed several person-level but few foot-level differences between participants with and those without neuropathic pain. Those with neuropathic pain had worse general health (demonstrated by lower SF-12 scores) and greater pain severity, which is consistent with previous reports related to neuropathic pain in individuals with knee OA using a range of health-related quality of life measures (11,13,14) and pain assessment tools (7,8,10,15). Interestingly, we found that although pain severity at rest was higher in those with neuropathic pain, pain during walking was not. This provides further evidence of a centrally mediated pain process in some participants, since pain severity when walking is typically greater than at rest in first MTP joint OA (25), possibly due to the loads associated with walking leading to mechanical stimulation of sensory neurons in local joint tissue.

The contribution of local, joint-level factors to neuropathic pain in OA is unclear. Although neuropathic pain in individuals with knee OA is more common in those with meniscal lesions (15) or those who have undergone surgery (10), findings related to the association with radiographic severity are inconsistent (12,13,16) and may be confounded by the influence of disease duration. We found no difference between the non-neuropathic group and neuropathic group in relation to measures of disease severity, including clinical features (such as range of motion, crepitus, or presence of a dorsal exostosis) or the presence and severity of radiographic OA. This is a notable finding, since previous studies demonstrated several dose-response relationships between radiographic severity of first MTP joint OA, range of motion, and symptoms, consistent with a longitudinal pattern of progression (20).

Taken together, these findings suggest that while local structural factors may play a role in first MTP joint OA disease progression and symptoms more broadly, neuropathic symptoms may be more closely related to systematic factors. However, it is also possible that the initial catalyst for OA symptoms is mechanical, and prolonged nociceptive input subsequently leads to neuropathic pain symptoms via central sensitization (5). Although the relationship was not statistically significant ($P = 0.055$), participants in our study with possible/likely neuropathic pain had a longer duration of symptoms (mean of 60 months versus 39 months).

The key clinical implication of these findings is that there may be some value in screening for neuropathic symptoms in individuals with first MTP joint OA, as this may influence treatment decisions. Emerging evidence suggests that individuals with neuropathic pain associated with knee OA may be less responsive to commonly used treatments such as physical therapy (38) or joint replacement surgery (39). Although no studies have explored this in relation to foot OA, the presence of neuropathic pain may at least partly explain why only modest improvements of symptoms have been observed in clinical trials of footwear and foot orthoses, interventions that address mechanical deficits associated with first MTP joint OA (24,25,40). In individuals with predominantly neuropathic symptoms, centrally acting pharmacologic treatment approaches may be indicated (2), although only duloxetine, a serotonin–norepinephrine reuptake inhibitor, has sufficient evidence to support its use in OA (41).

Strengths of this study include the well-characterized sample with validated clinical and radiographic measures of first MTP joint OA and a broad array of general health measures. However, our findings need to be considered in the context of several inherent limitations of the study design. First, our participants were recruited from a randomized trial rather than a population-based cohort, so the sample size was relatively small and may not be reflective of the broader population with first MTP joint OA. Second, our case definition for neuropathic pain was based on the PD-Q. Although this is a commonly used tool with some evidence of validity, there is currently no gold standard to definitively identify OA-associated neuropathic pain. We also used the original PD-Q rather than the modified version, the latter of which may have better validity, since it requests participants focus on neuropathic symptoms in or around the affected joint rather than their main area of pain, and the pain radiation question was reworded to improve clarity (35). We consider misclassification of neuropathic pain location in our study to be unlikely, as all symptom-related questions in the baseline survey specifically referred to the big toe joint. However, it is possible that some participants misunderstood the pain radiation question, since some non-adjacent radiation patterns were reported. Third, we did not perform any quantitative sensory testing, which would have provided greater insights into the contribution of central sensitization (42).

In conclusion, in this analysis of data from a randomized trial of individuals with first MTP joint OA, 1 in 3 individuals reported symptoms suggestive of neuropathic pain. Those with possible or likely neuropathic pain were older, had worse general physical health, worse foot health, and greater pain severity at rest. Screening for neuropathic pain may be helpful in the optimum selection of interventions in clinical practice and may be worthy of consideration when designing clinical trials.

ACKNOWLEDGMENTS

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

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. Dr. Menz had full access to all of

the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Menz, Landorf, Cicuttini, Roddy, Munteanu.

Acquisition of data. Allan, Buldt.

Analysis and interpretation of data. Menz, Landorf, Cicuttini, Roddy, Munteanu.

REFERENCES

1. Fu K, Robbins SR, McDougall JJ. Osteoarthritis: the genesis of pain. Rheumatology (Oxford) 2018;57 Suppl 4:iv43–50.
2. Thakur M, Dickenson AH, Baron R. Osteoarthritis pain: nociceptive or neuropathic? Nat Rev Rheumatol 2014;10:374–80.
3. Dimitroulas T, Duarte RV, Behura A, et al. Neuropathic pain in osteoarthritis: a review of pathophysiological mechanisms and implications for treatment. Semin Arthritis Rheum 2014;44:145–54.
4. Treede RD, Jensen TS, Campbell JN, et al. Neuropathic pain: redefinition and a grading system for clinical and research purposes. Neurology 2008;70:1630–5.
5. Hochman JR, French MR, Bermingham SL, et al. The nerve of osteoarthritis pain. Arthritis Care Res (Hoboken) 2010;62:1019–23.
6. Zolio L, Lim KY, McKenzie JE, et al. Systematic review and meta-analysis of the prevalence of neuropathic-like pain and/or pain sensitization in people with knee and hip osteoarthritis. Osteoarthritis Cartilage 2021;29:1096–116.
7. Hochman JR, Gagliese L, Davis AM, et al. Neuropathic pain symptoms in a community knee OA cohort. Osteoarthritis Cartilage 2011;19:647–54.
8. Ben Tekaya A, Rouached L, Maaoui R, et al. Neuropathic pain in patients with knee osteoarthritis: relation with comorbidities and functional status. Curr Rheumatol Rev doi: https://doi.org/10.2174/1573397118666220527140626. 2022. E-pub ahead of print.
9. Moss P, Benson HA, Will R, et al. Patients with knee osteoarthritis who score highly on the PainDETECT questionnaire present with multimodality hyperalgesia, increased pain, and impaired physical function. Clin J Pain 2018;34:15–21.
10. Valdes AM, Suokas AK, Doherty SA, et al. History of knee surgery is associated with higher prevalence of neuropathic pain-like symptoms in patients with severe osteoarthritis of the knee. Semin Arthritis Rheum 2014;43:588–92.
11. Aşkın A, Özkan A, Tosun A, et al. Quality of life and functional capacity are adversely affected in osteoarthritis patients with neuropathic pain. Kaohsiung J Med Sci 2017;33:152–8.
12. Van Helvoort EM, Welsing PM, Jansen MP, et al. Neuropathic pain in the IMI-APPROACH knee osteoarthritis cohort: prevalence and phenotyping. RMD Open 2021;7:e002025.
13. Güngör Demir U, Demir AN, Toraman NF. Neuropathic pain in knee osteoarthritis. Adv Rheumatol 2021;61:67.
14. Blikman T, Rienstra W, van Raay JJ, et al. Neuropathic-like symptoms and the association with joint-specific function and quality of life in patients with hip and knee osteoarthritis. PLoS One 2018;13:e0199165.
15. Roubille C, Raynauld JP, Abram F, et al. The presence of meniscal lesions is a strong predictor of neuropathic pain in symptomatic knee osteoarthritis: a cross-sectional pilot study. Arthritis Res Ther 2014;16:507.
16. Ohtori S, Orita S, Yamashita M, et al. Existence of a neuropathic pain component in patients with osteoarthritis of the knee. Yonsei Med J 2012;53:801–5.
17. Roddy E, Thomas MJ, Marshall M, et al. The population prevalence of symptomatic radiographic foot osteoarthritis in community-dwelling older adults: the Clinical Assessment Study of the Foot. Ann Rheum Dis 2015;74:156–63.
18. Paterson KL, Harrison C, Britt H, et al. Management of foot/ankle osteoarthritis by Australian general practitioners: an analysis of national patient-encounter records. Osteoarthritis Cartilage 2018;26:888–94.
19. Zammit GV, Munteanu SE, Menz HB. Development of a diagnostic rule for identifying radiographic osteoarthritis in people with first metatarsophalangeal joint pain. Osteoarthritis Cartilage 2011;19:939–45.
20. Menz HB, Roddy E, Marshall M, et al. Demographic and clinical factors associated with radiographic severity of first metatarsophalangeal joint osteoarthritis: cross-sectional findings from the Clinical Assessment Study of the Foot. Osteoarthritis Cartilage 2015;23:77–82.
21. Menz HB, Auhl M, Tan JM, et al. Centre of pressure characteristics during walking in individuals with and without first metatarsophalangeal joint osteoarthritis. Gait Posture 2018;63:91–6.
22. Menz HB, Auhl M, Tan JM, et al. Biomechanical effects of prefabricated foot orthoses and rocker-sole footwear in individuals with first metatarsophalangeal joint osteoarthritis. Arthritis Care Res (Hoboken) 2016;68:603–11.
23. McClelland JA, Allan JJ, Auhl M, et al. Effects of shoe-stiffening inserts on lower extremity kinematics in individuals with first metatarsophalangeal joint osteoarthritis. Arthritis Care Res (Hoboken) 2022;74:1849–56.
24. Menz HB, Auhl M, Tan JM, et al. Effectiveness of foot orthoses versus rocker-sole footwear for first metatarsophalangeal joint osteoarthritis: randomized trial. Arthritis Care Res (Hoboken) 2016;68:581–9.
25. Munteanu SE, Landorf KB, McClelland JA, et al. Shoe-stiffening inserts for first metatarsophalangeal joint osteoarthritis: a randomised trial. Osteoarthritis Cartilage 2021;29:480–90.
26. Munteanu SE, Landorf KB, McClelland JA, et al. Shoe-stiffening inserts for first metatarsophalangeal joint osteoarthritis (the SIMPLE trial): study protocol for a randomised controlled trial. Trials 2017;18:198.
27. Ware J Jr, Kosinski M, Keller SD. A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996;34:220–33.
28. Brown TA, Chorpita BF, Korotitsch W, et al. Psychometric properties of the Depression Anxiety Stress Scales (DASS) in clinical samples. Behav Res Ther 1997;35:79–89.
29. Nicholas MK. The pain self-efficacy questionnaire: taking pain into account. Eur J Pain 2007;11:153–63.
30. Bennett PJ, Patterson C, Wearing S, et al. Development and validation of a questionnaire designed to measure foot-health status. J Am Podiatr Med Assoc 1998;88:419–28.
31. Menz HB, Munteanu SE, Landorf KB, et al. Radiographic classification of osteoarthritis in commonly affected joints of the foot. Osteoarthritis Cartilage 2007;15:1333–8.
32. Menz HB, Munteanu SE, Marshall M, et al. Identification of radiographic foot osteoarthritis: sensitivity of views and features using the La Trobe radiographic atlas. Arthritis Care Res (Hoboken) 2022;74:1369–73.
33. Freynhagen R, Baron R, Gockel U, et al. painDETECT: a new screening questionnaire to identify neuropathic components in patients with back pain. Curr Med Res Opin 2006;22:1911–20.
34. Freynhagen R, Tölle TR, Gockel U, et al. The painDETECT project: far more than a screening tool on neuropathic pain. Curr Med Res Opin 2016;32:1033–57.
35. Hochman JR, Davis AM, Elkayam J, et al. Neuropathic pain symptoms on the modified painDETECT correlate with signs of central sensitization in knee osteoarthritis. Osteoarthritis Cartilage 2013;21:1236–42.

36. Sawilowsky SS. New effect size rules of thumb. J Mod Appl Stat Methods 2009;8:597–9.

37. Baron R. Mechanisms of disease: neuropathic pain: a clinical perspective. Nat Clin Pract Neurol 2006;2:95–106.

38. O'Leary H, Smart KM, Moloney NA, et al. Pain sensitization associated with nonresponse after physiotherapy in people with knee osteoarthritis. Pain 2018;159:1877–86.

39. Phillips JR, Hopwood B, Arthur C, et al. The natural history of pain and neuropathic pain after knee replacement: a prospective cohort study of the point prevalence of pain and neuropathic pain to a minimum three-year follow-up. Bone Joint J 2014;96-B:1227–33.

40. Paterson KL, Hinman RS, Metcalf BR, et al. Effect of foot orthoses vs sham insoles on first metatarsophalangeal joint osteoarthritis symptoms: a randomized controlled trial. Osteoarthritis Cartilage 2022;30:956–64.

41. Kolasinski SL, Neogi T, Hochberg MC, et al. 2019 American College of Rheumatology/Arthritis Foundation guideline for the management of osteoarthritis of the hand, hip, and knee. Arthritis Rheumatol 2020;72:220–33.

42. Gwilym SE, Keltner JR, Warnaby CE, et al. Psychophysical and functional imaging evidence supporting the presence of central sensitization in a cohort of osteoarthritis patients. Arthritis Rheum 2009;61:1226–34.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 75, No. 10, October 2023, pp 2134–2141
DOI 10.1002/acr.25127

Association of Sleep Disturbance With Catastrophizing and Knee Pain: Data From the Osteoarthritis Initiative

Yining Wang,[1] Xiaoxi Li,[1] Youyou Zhang,[1] Yubo Ma,[1] Shengqian Xu,[2] Zongwen Shuai,[2] Faming Pan,[1] and Guoqi Cai[1]

Objective. To investigate the relationship between sleep disturbance, catastrophizing, and knee pain in middle-aged and older individuals.

Methods. Data from the Osteoarthritis Initiative cohort from months 48 to 96 were used, where month 48 was treated as baseline. Knee pain (Western Ontario and McMaster Universities Osteoarthritis Index pain scale score ≥5 [range 0–20]), catastrophizing (extracted from Coping Strategies Questionnaire score ≥3 [range 0–6]), and sleep quality (extracted from Center for Epidemiologic Studies Depression Scale [range 1–4]) were assessed annually. We described the association of sleep disturbance with the presence and risk of knee pain and catastrophizing. The mediation effect of knee pain and catastrophizing on the sleep–catastrophizing and sleep–pain association was evaluated, respectively.

Results. Catastrophizing and knee pain were reported in 346 (10%) and 917 (24%) of the 3,813 participants (mean 64.9 years, 58% female) at baseline. Participants with worse sleep disturbance were more likely to have knee pain (prevalence ratio [PR] 1.4–2.0, *P* for trend <0.001) and catastrophizing (PR 1.4–3.1, *P* for trend <0.001). Sleep disturbance at baseline predicted the risk of knee pain (risk ratio [RR] 1.1, *P* for trend <0.001) and catastrophizing (RR 1.2–1.7, *P* for trend <0.001) during follow-up. No statistically significant interactions between sleep disturbance and knee pain or catastrophizing were observed. Knee pain and catastrophizing mediated the sleep–catastrophizing and sleep–pain association, respectively, at baseline, and knee pain negatively mediated the sleep–catastrophizing association longitudinally.

Conclusion. Sleep disturbance was associated with the presence and risk of catastrophizing and knee pain. Sleep interventions may have a universal and independent effect in preventing incident knee pain.

INTRODUCTION

Knee osteoarthritis (OA) affects ∼260 million people worldwide, and the most prominent problem it causes is pain (1,2). Pain relief is a primary goal for patients with knee OA. In addition to traditional pharmacologic treatments, many studies have shown that psychological treatments, including the management of pain catastrophizing, may be beneficial for pain relief (3–5). Catastrophizing is when someone assumes that the worst will happen, and that one is in a worse situation than the actual situation (6). The relationship between catastrophizing and pain has been shown to be reciprocal, but the cause and effect of their appearance and the order in which they occur remain uncertain.

Sleep disturbance affects ∼50% of older patients with knee OA (7). Previous studies have shown that purely targeting sleep problems only has small effects on pain outcomes (8,9). Sleep disturbance and pain-related experiences are closely related, and poor sleep quality can induce unhealthy emotions including catastrophizing (10,11). However, the associations between sleep disturbance, pain, and catastrophizing, especially the

This article was prepared using an Osteoarthritis Initiative (OAI) public-use data set, and its contents do not necessarily reflect the opinions or views of the OAI Study Investigators, the NIH, or the private funding partners of the OAI. The OAI is a public–private partnership between the NIH (contracts N01-AR-2-2258, N01-AR-2-2259, N01-AR-2-2260, N01-AR-2-2261, and N01-AR-2-2262) and private funding partners (Merck Research Laboratories, Novartis Pharmaceuticals, GlaxoSmithKline, and Pfizer, Inc.) and is conducted by the OAI Study Investigators. Private sector funding for the OAI is managed by the Foundation for the NIH. The authors of this article are not part of the OAI investigative team.

Supported by the National Natural Science Foundation of China (grants 82103933, 81773514, and 82073655).

[1]Yining Wang, BM, Xiaoxi Li, BM, Youyou Zhang, BM, Yubo Ma, MMed, Faming Pan, PhD, Guoqi Cai, PhD: Anhui Medical University and Inflammation and Immune Mediated Diseases Laboratory of Anhui Province, Hefei, China; [2]Shengqian Xu, MD, Zongwen Shuai, MD: The First Affiliated Hospital of Anhui Medical University, Hefei, China.

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.25127&file=acr25127-sup-0001-Disclosureform.pdf.

Address correspondence via email to Faming Pan, PhD at famingpan@ahmu.edu.cn, and Guoqi Cai, PhD at Guoqi.Cai@utas.edu.au.

Submitted for publication April 28, 2022; accepted in revised form April 4, 2023.

SIGNIFICANCE & INNOVATIONS

- Sleep disturbance was associated with the presence and predicted the risk of knee pain and catastrophizing in middle-aged and elderly individuals.
- Knee pain and catastrophizing mediated the association of sleep disturbance with catastrophizing and knee pain, respectively, in cross-sectional analysis at baseline.
- Knee pain inversely mediated the association between sleep disturbance and risk of catastrophizing longitudinally.
- The associations of sleep disturbance with the presence and risk of knee pain were not modified by catastrophizing status.

underlying pathways, are to be clarified. Several cross-sectional studies have confirmed that catastrophizing or pain were more likely to occur in patients with sleep disturbances (10–13). Interestingly, Tighe et al suggested that pain catastrophizing mediated the association between sleep disturbance and OA symptom severity (14), and Wilt et al showed that pain mediated the relationship between sleep disturbance and pain catastrophizing (15). However, these studies were small in sample size, and the temporal association between sleep disturbance and catastrophizing or pain cannot be determined. In a 24-month prospective study, the authors showed that insomnia promoted the spreading of chronic pain (16). Therefore, the aims of this study were to examine the role of sleep disturbance in the presence and risk of knee pain and catastrophizing, and to evaluate whether the sleep–pain and sleep–catastrophizing relations were modified by and mediated through catastrophizing and knee pain, respectively.

PATIENTS AND METHODS

Participants and procedure. This study was reported according to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement (17). Data used in this study were derived from the Osteoarthritis Initiative (OAI), a multicenter, longitudinal, prospective observational study of participants with or at increased risk of knee OA. The OAI cohort included 4,796 participants ages 45–79 years at the time of recruitment. Ethics approvals were obtained from the institutional review board at each of the 4 clinical centers that recruited OAI participants. All participants provided informed consent. Data of the OAI from 48–96 months were used in this study, where month 48 was considered as baseline. A total of 3,813 participants who had data on sleep disturbance and either knee pain or catastrophizing at 48 months and did not have any knee replacement surgery prior to month 48 were included in the study (Figure 1).

Primary predictor (sleep disturbance). Sleep quality during the past week was assessed by the Center for Epidemiologic Studies Depression Scale (CES-D) (18) every 12 months from months 48 to 96. There are 4 responses to the eleventh item of the CES-D, "My sleep was restless": Rarely or None of the time (<1 day); Some of the time (1–2 days); Much of the time (3–4 days); and Most or All of the time (5–7 days). Accordingly, we classified sleep disturbance into 4 scales from 1–4, with the larger number indicating worse sleep disturbance (19,20).

Outcome measure (catastrophizing). Catastrophizing was assessed annually based on 2 items from the Coping Strategies Questionnaire (CSQ) (21,22): when I feel pain, "It is terrible and I feel it is never going to get any better," and "I feel I can't stand it anymore." Both items were scored on a scale from 0 to 6, where 0 = never do, 3 = sometimes do, and 6 = always do. Scores for the 2-item scale were averaged such that the catastrophizing score ranged from 0 to 6. Based on the responses of the CSQ scale, we chose a score of 3 as the cutoff value for catastrophizing (i.e., ≥3). The validity of the 2-item CSQ scale has been demonstrated for the assessment of catastrophizing using data from the OAI (22).

Outcome measure (knee pain). Knee pain was measured annually using the 5-item Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain subscale (23) using the past 7 days as the timeframe. Each item was evaluated by a 5-point Likert scale (0 = none, 1 = mild, 2 = moderate, 3 = severe, 4 = extreme), and knee pain score was the sum of individual item scores (range 0–20). Previous studies defined symptomatic knees using a WOMAC pain score of ≥5 on a scale of 0–20 (24,25), which represents the upper tertile of all knees with any knee pain (WOMAC pain score >0). We adopted this definition in the present study, and knee pain was considered present if any knee had a WOMAC pain score of ≥5.

Covariates. Covariates were selected based on previous literature (26–28). They included the following variables: sex, age (year), body mass index (BMI; kg/m^2), race (White, Black, Other), level of education, the use of medications (antidepressants and antianxiety drugs, nonsteroidal antiinflammatory drugs, steroids, painkillers, and acetaminophen), and radiographic OA. Radiographic OA was defined as a Kellgren/Lawrence grade of ≥2 using a fixed-flexion knee radiograph (no, unilateral, and bilateral) (29). For covariates that were not available at month 48 (i.e., the baseline of this study), we used data measured at month 0 (i.e., the baseline of the OAI).

Statistical analysis. Baseline characteristics were described using the mean ± SD or percentage, split by sleep disturbance. Analysis of variance and chi-square tests were used to test between-group differences. Data analyses were performed

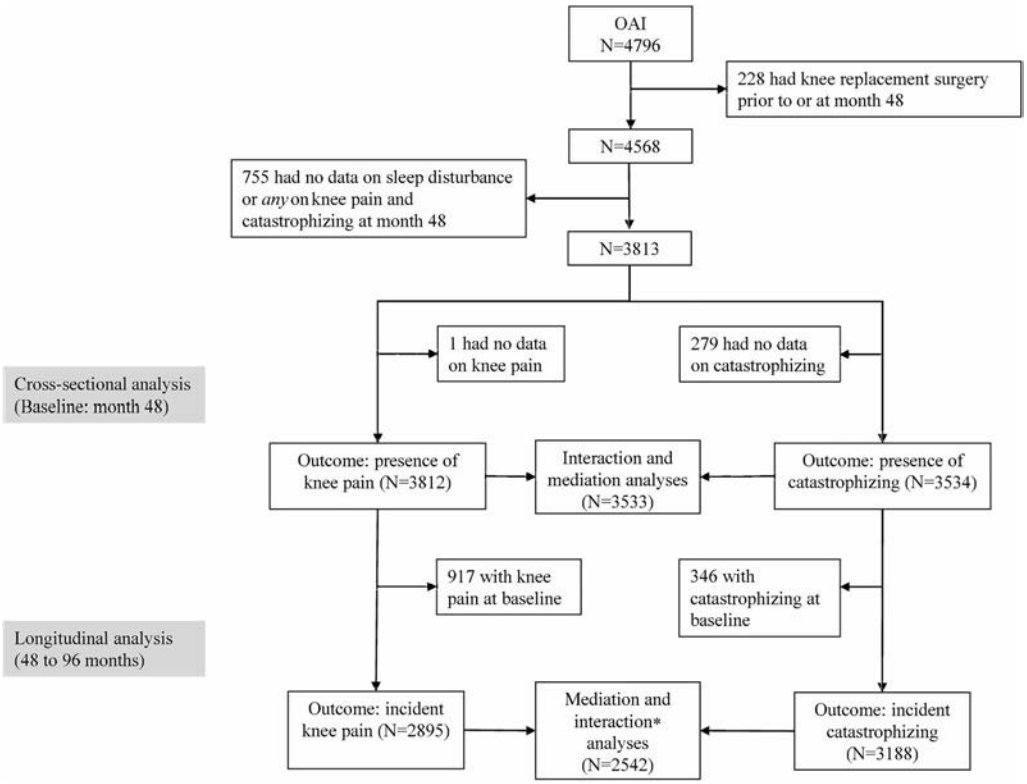

Figure 1. Study flow chart. * = Longitudinal interaction analyses were conducted between sleep disturbance and catastrophizing at month 48 for incident knee pain from 48 to 96 months (n = 2,895), and between sleep disturbance and knee pain at month 48 for incident catastrophizing from 48 to 96 months (n = 3,188). OAI = Osteoarthritis Initiative.

using R, version 4.1.1. A 2-sided *P* value less than 0.05 was considered statistically significant.

In cross-sectional analyses, we used log-binomial regression models to analyze the associations of sleep disturbance with the presence of knee pain and catastrophizing at baseline, and the results were shown as the prevalence ratio (PR) with 95% confidence intervals (95% CIs). In longitudinal analyses, log-binomial regression models were also used to evaluate the associations of sleep disturbance at baseline and the risk of knee pain and catastrophizing during follow-up, which was done for participants with no knee pain (WOMAC pain score <5) or catastrophizing (2-item CSQ score <3) at baseline, respectively. For participants who had died or relocated during follow-up, they were considered lost to follow-up, and only available data were used in the data analyses. The results were shown as the risk ratio (RR) with 95% CI. If the log-binomial regression models failed to converge, the PR or RR was estimated using a Poisson distribution and robust SEs (30). The interactions between sleep disturbance and knee pain or catastrophizing at baseline were assessed when the presence or risk of catastrophizing and knee pain were the outcome measures, respectively.

We conducted cross-sectional and longitudinal mediation analyses using the Karlson-Holm-Breen decomposition method (31) to determine whether knee pain mediated the sleep–catastrophizing association and whether catastrophizing mediated the sleep–pain association. In cross-sectional mediation analyses, data on sleep disturbance, knee pain, and catastrophizing at baseline were used. Longitudinal mediation analyses were conducted for participants with neither knee pain nor catastrophizing at baseline (i.e., WOMAC pain score <5 and 2-item CSQ score <3). Data from participants were removed since the time of implementation of any knee replacement surgery during follow-up. Specifically, sleep disturbance at baseline was considered the exposure, any occurrence of knee pain or catastrophizing during follow-up was considered the outcome, respectively, and mediator was defined as any presence of catastrophizing or knee pain prior to outcome, respectively. For example, when knee pain was the outcome and first presented at 84 months, catastrophizing (the mediator) would be considered present if a participant had a 2-item CSQ score of ≥3 at any follow-up visit for month 60, 72, or 84. This was to guarantee the temporal order between the exposure and the mediator. The interactions between exposure (sleep disturbance) and mediator (knee pain or catastrophizing) were assessed, and no statistically significant interaction was observed (all $P > 0.05$).

For cross-sectional and longitudinal analyses but not mediation analyses, missing data on covariates (0.05–9.4% missing) were addressed using multiple imputation with chained equations. Ten imputations were conducted using complete covariates and nonmissing values of the predictor and outcome

measures at baseline, assuming missing at random. Three models to adjust for covariates were conducted. Apart from univariable model, model 1 adjusted for age, sex, race, BMI, and education level at baseline, and model 2 further adjusted for the use of medications and radiographic OA at baseline. Complete case analyses were conducted as a sensitivity analysis for the associations of sleep disturbance with the presence and risk of knee pain and catastrophizing. The false discovery rate (FDR) approach was conducted to control the family-wise Type I error (32).

RESULTS

Participants. Table 1 summarizes the baseline characteristics of the study participants. Among 3,813 participants included in this study, 917 (24%) and 346 (9%) were defined as having knee pain and catastrophizing at baseline, respectively. Participants with worse sleep disturbance were younger, had higher BMI and education level, and were more likely to have knee pain and catastrophizing and to be female and White.

Association of sleep disturbance with knee pain and catastrophizing at baseline. In both univariable and multivariable models, participants with sleep disturbance were more likely to have knee pain and catastrophizing, and these associations were stronger with worse sleep disturbance (Table 2). The interaction between sleep disturbance and knee pain for the presence of catastrophizing was not statistically significant, nor was the interaction between sleep disturbance and catastrophizing for the presence of knee pain (all P for interaction >0.05) (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25127). Mediation analyses showed that the association between sleep disturbance and knee pain was mediated through catastrophizing, and that the association between sleep disturbance and catastrophizing was also mediated through knee pain (Table 3).

Sleep disturbance predicted the risk of knee pain and catastrophizing. During follow-up, 919 of 2,895 (32%) and 643 of 3,188 (20%) participants without knee pain and catastrophizing at baseline developed incident knee pain and catastrophizing, respectively. Sleep disturbance was associated with increased risks of both knee pain and catastrophizing in univariable and multivariable models (Table 4). The interaction between sleep disturbance and catastrophizing for incident knee pain was not statistically significant, nor was the interaction between

Table 1. Baseline characteristics of study participants*

		Sleep disturbance in the past week (n = 3,813)				
	Total	<1 day (n = 1,529)	1–2 days (n = 1,634)	3–4 days (n = 399)	5–7 days (n = 253)	*P*
Age, mean ± SD years	64.9 ± 9.0	65.1 ± 9.1	65.3 ± 9.0	63.7 ± 9.2	63.4 ± 9.5	0.00127
Body mass index, mean ± SD kg/m^2	28.5 ± 5.0	28.2 ± 4.7	28.61 ± 5.0	28.8 ± 5.3	29.7 ± 5.3	<0.001
Sex						
Male	1,598 (41.9)	671 (43.9)	685 (41.9)	148 (37.2)	94 (37.2)	0.03758
Female	2,215 (58.1)	858 (56.1)	948 (58.0)	250 (62.8)	159 (62.9)	–
Race						
White	3,108 (81.5)	1,222 (79.9)	1,344 (82.3)	332 (83.4)	210 (83.0)	0.02104
Black	615 (16.1)	276 (18.1)	253 (15.5)	50 (12.6)	36 (14.2)	–
Other	87 (2.3)	31 (2.0)	33 (2.0)	16 (4.0)	7 (2.8)	–
Education level						
Less than high school	102 (2.7)	29 (1.9)	44 (2.7)	13 (3.3)	16 (6.3)	0.007046
High school	437 (11.5)	174 (11.4)	178 (10.9)	42 (10.6)	43 (17.0)	–
Some college	860 (22.6)	339 (22.2)	367 (22.5)	95 (23.9)	59 (23.3)	–
College graduate	844 (22.1)	359 (23.5)	358 (21.9)	79 (19.9)	48 (19.0)	
Some graduate school	322 (8.4)	129 (8.4)	139 (8.5)	36 (9.1)	18 (7.1)	–
Graduate degree (master's or PhD)	1,230 (32.3)	492 (32.2)	540 (33.1)	131 (32.9)	67 (26.5)	–
Knee pain (n = 3,812)	917 (24.1)	268 (17.5)	417 (25.5)	123 (30.9)	109 (43.1)	<0.001
Catastrophizing (n = 3,534)	346 (9.8)	92 (6.0)	140 (8.6)	51 (12.8)	63 (24.9)	<0.001
Medications						
Depression and anxiety	573 (15.0)	189 (12.4)	249 (15.2)	82 (20.6)	53 (21.0)	<0.001
Painkiller	127 (3.3)	35 (2.3)	42 (2.6)	24 (6.0)	26 (10.3)	<0.001
Steroids	116 (3.0)	29 (1.9)	58 (3.6)	20 (5.0)	9 (3.6)	0.003332
Acetophenone	341 (8.9)	93 (6.1)	159 (9.7)	44 (11.1)	45 (17.8)	<0.001
NSAIDs	844 (22.1)	260 (17.0)	403 (24.7)	114 (28.6)	67 (26.5)	<0.001
Radiographic osteoarthritis						
Neither knee	1,416 (37.1)	558 (36.5)	607 (37.2)	158 (39.7)	93 (36.8)	0.3735
Unilateral	876 (23.0)	343 (22.4)	398 (24.4)	84 (21.1)	51 (20.2)	–
Bilateral	1,155 (30.3)	482 (31.5)	479 (29.3)	110 (27.6)	84 (33.2)	–
Multijoint pain	472 (12.4)	123 (8.0)	212 (13.0)	74 (18.6)	63 (24.9)	<0.001

* Values are the number (%) unless indicated otherwise. NSAIDs = nonsteroidal antiinflammatory drugs.

Table 2. Cross-sectional associations of sleep disturbance with knee pain and catastrophizing*

	Univariable	Model 1†	Model 2‡
Knee pain (n = 3,812)			
Rarely or none of the time (<1 day)	–	–	–
Some of the time (1–2 days)	1.46 (1.27–1.67)§	1.45 (1.27–1.66)§	1.36 (1.19–1.54)§
Much of the time (3–4 days)	1.76 (1.47–2.12)§	1.74 (1.47–2.07)§	1.56 (1.32–1.85)§
Most or all of the time (5–7 days)	2.46 (2.05–2.94)§	2.30 (1.93–2.75)§	2.01 (1.69–2.39)§
P for trend	<0.001§	<0.001§	<0.001§
Catastrophizing (n = 3,534)			
Rarely or none of the time (<1 day)	–	–	–
Some of the time (1–2 days)	1.45 (1.12–1.86)§	1.47 (1.15–1.89)§	1.35 (1.05–1.73)§
Much of the time (3–4 days)	2.21 (1.60–3.05)§	2.19 (1.60–3.01)§	1.92 (1.40–2.63)§
Most or all of the time (5–7 days)	4.14 (3.10–5.53)§	3.77 (2.84–5.00)§	3.05 (2.30–4.06)§
P for trend	<0.001§	<0.001§	<0.001§

* Values are the prevalence ratio (95% confidence interval) unless indicated otherwise.
† Adjusted for sex, age, education level, race, and body mass index.
‡ Adjusted for model 1 plus medication use and radiographic osteoarthritis.
§ Significant.

sleep disturbance and knee pain for incident catastrophizing (all *P* for interaction >0.05) (see Supplementary Table 1, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25127). Mediation analysis illustrated no statistically significant indirect effect through catastrophizing for the association of sleep disturbance with incident knee pain, but knee pain, as the mediator, showed a negative indirect effect for the association of sleep disturbance with catastrophizing (Table 3).

Sensitivity analyses. Complete case analyses did not materially change the main findings that showed that sleep disturbance was associated with the presence and predicted the risk of knee pain and catastrophizing (see Supplementary Tables 2 and 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25127). The results were not materially changed after FDR correction (see Supplementary Table 4, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25127).

DISCUSSION

Using data from the OAI, we found that sleep disturbance was associated with the presence and predicted the risk of both knee pain and catastrophizing, and that knee pain and

Table 3. Mediation effects of catastrophizing and knee pain for the cross-sectional and longitudinal associations of sleep disturbance with knee pain and catastrophizing*

	Cross-sectional		Longitudinal	
Sleep disturbance	Catastrophizing†	Knee pain‡	Catastrophizing§	Knee pain¶
Rarely or none of the time (<1 day)#	Reference	Reference	Reference	Reference
Some of the time (1–2 days)				
Total effect	1.58 (1.29–1.95)**	1.41 (1.03–1.92)**	1.33 (1.16–1.52)**	1.17 (0.97–1.41)
Direct effect	1.54 (1.25–1.89)**	1.32 (0.97–1.81)	1.33 (1.17–1.52)**	1.22 (1.01–1.47)**
Indirect effect	1.03 (1.00–1.06)**	1.06 (1.03–1.10)**	1.00 (0.99–1.01)	0.96 (0.93–0.99)**
Mediation percentage††	6.20**	18.20**	–	–‡‡
Much of the time (3–4 days)				
Total effect	2.15 (1.58–2.93)**	2.11 (1.38–3.22)**	1.62 (1.35–1.94)**	1.27 (0.94–1.71)
Direct effect	1.98 (1.45–2.70)**	1.90 (1.24–2.90)**	1.61 (1.34–1.94)**	1.36 (1.01–1.84)**
Indirect effect	1.09 (1.04–1.14)**	1.11 (1.05–1.18)**	1.00 (0.99–1.02)	0.93 (0.89–0.98)**
Mediation percentage††	10.70**	14.30**	0.36	–
Most or all of the time (5–7 days)				
Total effect	3.37 (2.37–4.79)**	4.20 (2.74–6.44)**	1.76 (1.39–2.22)**	1.72 (1.23–2.40)**
Direct effect	2.76 (1.94–3.93)**	3.45 (2.24–5.30)**	1.78 (1.41–2.25)**	1.79 (1.28–2.50)**
Indirect effect	1.22 (1.14–1.30)**	1.22 (1.12–1.32)**	0.99 (0.96–1.01)	0.96 (0.91–1.02)
Mediation percentage††	16.40**	13.70**	–	–

* Values are the risk ratio (95% confidence interval) unless indicated otherwise.
† Adjusted for sex, age, education, race, body mass index, use of medications, and radiographic osteoarthritis.
‡ Catastrophizing mediated the association between sleep disturbance and the presence of knee pain at baseline.
§ Knee pain mediated the association between sleep disturbance and the presence of catastrophizing at baseline.
¶ Catastrophizing mediated the association between sleep disturbance and the risk of knee pain during follow-up.
# Knee pain mediated the association between sleep disturbance and the risk of catastrophizing during follow-up.
** Significant.
†† The mediation percentage = β of indirect effect / β of total effect.
‡‡ The indirect effect and total effect were in different directions.

Table 4. Predictive effect of sleep disturbance on incident knee pain and catastrophizing during follow-up*

Sleep disturbance	Univariable model	Model 1†	Model 2‡
Knee pain (n = 2,895)			
Rarely or none of the time (<1 day)	Reference	Reference	Reference
Some of the time (1–2 days)	1.34 (1.17–1.52)§	1.07 (1.04–1.10)§	1.06 (1.03–1.10)§
Much of the time (3–4 days)	1.57 (1.30–1.90)§	1.13 (1.07–1.18)§	1.12 (1.07–1.18)§
Most or all of the time (5–7 days)	1.76 (1.41–2.21)§	1.13 (1.06–1.21)§	1.14 (1.06–1.22)§
P for trend	<0.001§	<0.001§	<0.001§
Catastrophizing (n = 3,188)			
Rarely or none of the time (<1 day)	Reference	Reference	Reference
Some of the time (1–2 days)	1.20 (1.00–1.45)§	1.18 (0.98–1.43)	1.18 (0.97–1.43)
Much of the time (3–4 days)	1.35 (1.01–1.80)§	1.40 (1.05–1.87)§	1.25 (0.92–1.70)
Most or all of the time (5–7 days)	1.76 (1.27–2.43)§	1.74 (1.27–2.39)§	1.71 (1.22–2.40)§
P for trend	<0.001§	<0.001§	<0.001§

* Values are the risk ratio (95% confidence interval) unless indicated otherwise.
† Adjusted for sex, age, education level, race, and body mass index.
‡ Adjusted for model 1 plus medication use and radiographic osteoarthritis.
§ Significant.

catastrophizing did not modify the associations of sleep disturbance with the presence and risk of catastrophizing and knee pain, respectively. Cross-sectional mediation analyses at baseline showed that catastrophizing mediated the association between sleep disturbance and knee pain (6.2–16.4%) and that knee pain mediated the association between sleep disturbance and catastrophizing (13.7–18.2%). However, longitudinal mediation analyses suggested that knee pain inversely mediated the association between sleep disturbance and incident knee pain. The predictive effect of sleep disturbance on incident knee pain was universal in participants with or without catastrophizing and in participants with or without knee pain when predicting incident catastrophizing. These findings suggest that although sleep disturbance, catastrophizing, and knee pain are closely related, the effect of sleep interventions on knee pain may not be influenced by catastrophizing status.

Cross-sectionally, we observed a higher prevalence of both knee pain and catastrophizing at baseline among knee OA patients with worse sleep disturbance in a dose-response manner. This is consistent with existing evidence showing a role of sleep disturbance in pain and catastrophizing (10–13). We found no statistically significant interactions between sleep disturbance and knee pain or catastrophizing when the presence of catastrophizing and knee pain were the outcome measures, respectively. Considering that sleep interventions may improve OA and spinal pain (32,33), our findings suggested a universal effect of sleep interventions on pain for patients with or without catastrophizing. Cross-sectional mediation analyses suggested that knee pain and catastrophizing served as a mediator for the sleep–catastrophizing and sleep–pain associations at baseline, respectively. Wilt et al indicated that pain mediated the association between sleep disturbance and catastrophizing, but catastrophizing was not a mediator for the association between sleep disturbance and pain (15). This inconsistency may be due to the small sample size and large age difference in the study population (15). However, an important issue of these cross-sectional mediation analyses is that it is not possible to establish a temporal link, and the statistically significant mediation effects may only be a reflection of the coexistence of sleep disturbance, knee pain, and catastrophizing.

In the longitudinal analyses, we found that worse sleep disturbance at baseline was associated with higher risks of both knee pain and catastrophizing during the 4-year follow-up. Previous studies also showed that sleep disturbance contributed to the worsening of knee symptoms in knee OA patients (14). The effect of sleep disturbance on pain may be due to its negative role in innate immunity and inflammatory biology (34–36), which has been found to be associated with the progression of both knee symptoms and structural damage (37). Moreover, sleep disturbance may contribute to abnormalities in the central nervous system pathways that regulate pain (38). The reasons that sleep disturbance predicted catastrophizing may be primarily due to the close relationship between pain and catastrophizing, as catastrophizing stems from a misperception of some painful events (6), which can be enhanced by poor sleep quality (39,40). Another reason may be that sleep disturbance leads to increased concern and anxiety about getting enough sleep, a process that adds to the cognitive load and thus produces more catastrophizing emotions (10,41). Despite this, we did not observe a statistically significant interaction between sleep disturbance and catastrophizing for incident knee pain or between sleep disturbance and knee pain for incident catastrophizing. This suggests that among middle-aged and older individuals without knee pain, the management of sleep disturbance plays a role in preventing incident knee pain, irrespective of the status of catastrophizing.

Importantly, our longitudinal mediation analyses, in contrast to the cross-sectional mediation analyses, showed that catastrophizing did not mediate the associations of sleep disturbance at baseline with the risk of knee pain, but knee pain negatively mediated the association of sleep disturbance and catastrophizing. The negative mediation effect was due to the inverse association between new-onset knee pain and incident catastrophizing, which contrasts

with the findings of a previous longitudinal study showing a positive association between them (15). The reason for the inconsistency is unclear, but it may be due to the fact that the definitions used for pain and catastrophizing in that study were different than those in the current study. Moreover, the OAI was followed-up annually and did not specify the time anchor when assessing catastrophizing, and this may have introduced recall biases. Therefore, further longitudinal studies are warranted. This, combined with the findings of the longitudinal analyses, suggest that while sleep disturbance may have a detrimental effect on the risk of both knee pain and catastrophizing, the effect was independent of one another. Therefore, sleep interventions may be beneficial for both knee pain and catastrophizing, albeit the mechanisms of action may differ.

The strengths of this study include the large sample size with a long-term follow-up. Moreover, we performed both cross-sectional and longitudinal mediation analyses to better determine the associations among sleep disturbance, knee pain, and catastrophizing. Limitations of the current study are worth noting. First, catastrophizing was evaluated using 2 simple questions from the CSQ. Although the validity of catastrophizing using the CSQ has been proven (22), its assessment using a more specialized tool, such as the pain catastrophizing scale (42), would provide a better estimation. Second, the assessments of knee pain and catastrophizing were implemented per annum, and this may have missed some variations of these subjective measures. Third, as an observational study, residual confounding cannot be excluded. However, we have included multiple confounders based on previous literature to minimize risk of bias. Moreover, we did not check the sequential ignorability assumption for the causal mediation analysis.

In conclusion, sleep disturbance was associated with the presence and predicted the risk of catastrophizing and knee pain. While pain relief may benefit from the integrated management of both sleep problems and catastrophizing, sleep interventions may have a universal and independent effect in preventing incident knee pain, irrespective of catastrophizing.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. Drs. Cai and Pan had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Wang, Pan, Cai.

Acquisition of data. Wang, Cai.

Analysis and interpretation of data. Wang, Li, Zhang, Ma, Xu, Shuai, Pan, Cai.

REFERENCES

1. GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018;392:1789–858.
2. Peat G, McCarney R, Croft P. Knee pain and osteoarthritis in older adults: a review of community burden and current use of primary health care. Ann Rheum Dis 2001;60:91–7.
3. Smith MT, Finan PH, Buenaver LF, et al. Cognitive-behavioral therapy for insomnia in knee osteoarthritis: a randomized, double-blind, active placebo-controlled clinical trial. Arthritis Rheumatol 2015;67:1221–33.
4. Lerman SF, Finan PH, Smith MT, et al. Psychological interventions that target sleep reduce pain catastrophizing in knee osteoarthritis. Pain 2017;158:2189–95.
5. Cohen E, Lee YC. A mechanism-based approach to the management of osteoarthritis pain. Curr Osteoporos Rep 2015;13:399–406.
6. Van Damme S, Crombez G, Bijttebier P, et al. A confirmatory factor analysis of the Pain Catastrophizing Scale: invariant factor structure across clinical and non-clinical populations. Pain 2002;96:319–24.
7. Finan PH, Goodin BR, Smith MT. The association of sleep and pain: an update and a path forward. J Pain 2013;14:1539–52.
8. Bohra MH, Espie CA. Is cognitive behavioural therapy for insomnia effective in treating insomnia and pain in individuals with chronic non-malignant pain? Br J Pain 2013;7:138–51.
9. Flink IK, Linton SJ. Pain, sleep and catastrophizing: the conceptualization matters: comment on Wilt JA et al. "A multilevel path model analysis of the relations between sleep, pain, and pain catastrophizing in chronic pain rehabilitation patients." Scand J Pain 2016;10:119–21.
10. Harvey AG, Greenall E. Catastrophic worry in primary insomnia. J Behav Ther Exp Psychiatry 2003;34:11–23.
11. Vega R, Miró J, Esteve R, et al. Sleep disturbance in individuals with physical disabilities and chronic pain: the role of physical, emotional and cognitive factors. Disabil Health J 2019;12:588–93.
12. Ponce Martinez C, Edwards KA, Roos CR, et al. Associations among sleep disturbance, pain catastrophizing, and pain intensity for methadone-maintained patients with opioid use disorder and chronic pain. Clin J Pain 2020;36:641–7.
13. Campbell CM, Buenaver LF, Finan P, et al. Sleep, pain catastrophizing, and central sensitization in knee osteoarthritis patients with and without insomnia. Arthritis Care Res (Hoboken) 2015;67:1387–96.
14. Tighe CA, Youk A, Ibrahim SA, et al. Pain catastrophizing and arthritis self-efficacy as mediators of sleep disturbance and osteoarthritis symptom severity. Pain Med 2020;21:501–10.
15. Wilt JA, Davin S, Scheman J. A multilevel path model analysis of the relations between sleep, pain, and pain catastrophizing in chronic pain rehabilitation patients. Scand J Pain 2016;10:122–9.
16. Wiklund T, Gerdle B, Linton SJ, et al. Insomnia is a risk factor for spreading of chronic pain: a Swedish longitudinal population study (SwePain). Eur J Pain 2020;24:1348–56.
17. Von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007;370:1453–7.
18. Radloff LS. The CES-D Scale: a self-report depression scale for research in the general population. Appl Psychol Meas 1977;1:385–401.
19. Bethea TN, Zhai W, Zhou X, et al. Associations between longitudinal changes in sleep disturbance and depressive and anxiety symptoms during the COVID-19 virus pandemic among older women with and without breast cancer in the thinking and living with breast cancer study. Cancer Med 2022;11:3352–63.
20. Leggett A, Assari S, Burgard S, et al. The effect of sleep disturbance on the association between chronic medical conditions and depressive symptoms over time. Longit Life Course Stud 2017;8:138–51.

21. Sandborgh M, Lindberg P, Denison E. Pain belief screening instrument: development and preliminary validation of a screening instrument for disabling persistent pain. J Rehabil Med 2007;39:461–6.
22. Riddle DL, Jensen MP. Construct and criterion-based validity of brief pain coping scales in persons with chronic knee osteoarthritis pain. Pain Med 2013;14:265–75.
23. Bellamy N, Buchanan WW, Goldsmith CH, et al. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988;15:1833–40.
24. Bellamy N, Bell MJ, Pericak D, et al. BLISS index for analyzing knee osteoarthritis trials data. J Clin Epidemiol 2007;60:124–32.
25. Wang Y, Teichtahl AJ, Abram F, et al. Knee pain as a predictor of structural progression over 4 years: data from the Osteoarthritis Initiative, a prospective cohort study. Arthritis Res Ther 2018;20:250.
26. Cai G, Li X, Zhang Y, et al. Knee symptom but not radiographic knee osteoarthritis increases the risk of falls and fractures: results from the Osteoarthritis Initiative. Osteoarthritis Cartilage 2022;30:436–42.
27. Jehu DA, Davis JC, Falck RS, et al. Risk factors for recurrent falls in older adults: a systematic review with meta-analysis. Maturitas 2021; 144:23–8.
28. Geng Y, Lo JC, Brickner L, et al. Racial-ethnic differences in fall prevalence among older women: a cross-sectional survey study. BMC Geriatr 2017;17:65.
29. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis 1957;16:494–502.
30. Zou G. A modified Poisson regression approach to prospective studies with binary data. Am J Epidemiol 2004;159:702–6.
31. Karlson KB, Holm A. Decomposing primary and secondary effects: a new decomposition method. Res Soc Stratif Mobil 2011;29:221–37.
32. Vitiello MV, Rybarczyk B, Von Korff M, et al. Cognitive behavioral therapy for insomnia improves sleep and decreases pain in older adults with co-morbid insomnia and osteoarthritis. J Clin Sleep Med 2009; 5:355–62.
33. Ho KK, Ferreira PH, Pinheiro MB, et al. Sleep interventions for osteoarthritis and spinal pain: a systematic review and meta-analysis of randomized controlled trials. Osteoarthritis Cartilage 2019;27:196–218.
34. Irwin MR, Olmstead R, Carroll JE. Sleep disturbance, sleep duration, and inflammation: a systematic review and meta-analysis of cohort studies and experimental sleep deprivation. Biol Psychiatry 2016;80: 40–52.
35. Irwin MR, Opp MR. Sleep health: reciprocal regulation of sleep and innate immunity. Neuropsychopharmacology 2017;42:129–55.
36. Irwin MR. Why sleep is important for health: a psychoneuroimmunology perspective. Annu Rev Psychol 2015;66:143–72.
37. Woodell-May JE, Sommerfeld SD. Role of inflammation and the immune system in the progression of osteoarthritis. J Orthop Res 2020;38:253–7.
38. Boyden SD, Hossain IN, Wohlfahrt A, et al. Non-inflammatory causes of pain in patients with rheumatoid arthritis. Curr Rheumatol Rep 2016;18:30.
39. Mun CJ, Davis MC, Campbell CM, et al. Linking nonrestorative sleep and activity interference through pain catastrophizing and pain severity: an intraday process model among individuals with fibromyalgia. J Pain 2020;21:546–56.
40. Schuh-Hofer S, Wodarski R, Pfau DB, et al. One night of total sleep deprivation promotes a state of generalized hyperalgesia: a surrogate pain model to study the relationship of insomnia and pain. Pain 2013; 154:1613–21.
41. Perlis M, Aloia M, Millikan A, et al. Behavioral treatment of insomnia: a clinical case series study. J Behav Med 2000;23:149–61.
42. Sullivan MJ, Bishop SR, Pivik J. The pain catastrophizing scale: development and validation. Psychol Assess 1996;7:524–32.

Arthritis Care & Research
Vol. 75, No. 10, October 2023, pp 2142–2150
DOI 10.1002/acr.25114

American College of Rheumatology
Empowering Rheumatology Professionals

Diagnostic Yield of Computed Tomography for Cancer Detection in a Tertiary Referral Population of Idiopathic Inflammatory Myositis Patients

Christopher A. Mecoli,[1] Brant Chee,[2] Mengkun Chen,[1] XingYao Wang,[1] Jemima Albayda,[1] Julie J. Paik,[1] Eleni Tiniakou,[1] Brittany Adler,[1] Will Kelly,[1] Andrew L. Mammen,[3] Elizabeth A. Platz,[4] Livia Casciola-Rosen,[1] Lisa Christopher-Stine,[1] and Ami A. Shah[1]

Objective. To inform guidance for cancer detection in patients with idiopathic inflammatory myopathy (IIM), we evaluated the diagnostic yield of computed tomography (CT) imaging for cancer screening/surveillance within distinct IIM subtypes and myositis-specific autoantibody strata.

Methods. We conducted a single-center, retrospective cohort study in IIM patients. Overall diagnostic yield (number of cancers diagnosed/number of tests performed), percentage of false positives (number of biopsies performed not leading to cancer diagnosis/number of tests performed), and test characteristics were determined on CT of the chest and abdomen/pelvis.

Results. Within the first 3 years since IIM symptom onset, a total of 9 of 1,011 (0.9%) chest CT scans and 12 of 657 (1.8%) abdomen/pelvis CT scans detected cancer. Diagnostic yields for both CT of the chest and CT of the abdomen/pelvis were highest in dermatomyositis, specifically anti–transcription intermediary factor 1γ (2.9% and 2.4% for CT of the chest and abdomen/pelvis, respectively). The highest percentage of false positives was in patients with antisynthetase syndrome (ASyS) (4.4%) and immune-mediated necrotizing myopathy (4.4%) on CT of the chest, and ASyS (3.8%) on CT of the abdomen/pelvis. Patients ages <40 years old at IIM onset had both low diagnostic yields (0% and 0.5%) and high false-positive rates (1.9% and 4.4%) for CT of the chest and abdomen/pelvis, respectively.

Conclusion. In a tertiary referral cohort of IIM patients, CT imaging has a wide range of diagnostic yield and frequency of false positives for contemporaneous cancer. These findings suggest that cancer detection strategies targeted according to IIM subtype, autoantibody positivity, and age may maximize cancer detection while minimizing the harms and costs of over-screening.

INTRODUCTION

The association between specific subgroups of idiopathic inflammatory myopathy (IIM) and contemporaneous cancer is well established (1,2). Over the past few decades, the ability to risk stratify patients for cancer-associated myositis has improved through the application of clinically and biologically relevant filters, such as IIM subtype and myositis-specific autoantibodies (MSAs). For instance, patients with dermatomyositis (DM) and anti–transcription intermediary factor 1γ (anti-TIF1γ) antibodies have a high risk of cancer within 3 years of IIM symptom onset. However, these risk stratification filters are imperfect; for example, not all patients with anti-TIF1γ antibodies develop cancer (3), and patients without anti-TIF1γ antibodies can still develop cancer

Supported by the Jerome L. Greene Foundation, the Donald B. and Dorothy L. Stabler Foundation, the Huayi and Siuling Zhang Discovery Fund, Dr. Peter Buck, and the NIH (grants P30-AR-070254 and P30-CA-006973).

Drs. Mecoli, Paik, Tiniakou, and Shah's work was supported by the NIH (grants 1K23-AR-0758980, K23-AR-0739, K08-AR-0777732, and K24-AR-080217, respectively). Dr. Mammen's work was supported by the Intramural Research Program of the NIH/National Institute of Arthritis and Musculoskeletal and Skin Diseases. Drs. Casciola-Rosen and Shah's work was supported by the NIH (grant R01-AR-073208).

[1]Christopher A. Mecoli, MD, MHS, Mengkun Chen, MD, MHS, XingYao Wang, BS, Jemima Albayda, MD, Julie J. Paik, MD, MHS, Eleni Tiniakou, MD, Brittany Adler, MD, Will Kelly, BS, Livia Casciola-Rosen, PhD, Lisa Christopher-Stine, MD, MPH, Ami A. Shah, MD, MHS: Johns Hopkins University School of Medicine, Baltimore, Maryland; [2]Brant Chee, PhD: Johns Hopkins University School of Medicine, Baltimore, Maryland, and Johns Hopkins University Applied Physics Laboratory, Laurel, Maryland; [3]Andrew L. Mammen, MD, PhD: Johns Hopkins University School of Medicine, Baltimore, Maryland, and the NIH/National Institute of Arthritis and Musculoskeletal and Skin Diseases, Bethesda, Maryland; [4]Elizabeth A. Platz, ScD, MPH: Johns Hopkins Bloomberg School of Public Health and Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins University, Baltimore, Maryland.

Drs. Christopher-Stine and Shah contributed equally to this work.

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.25114&file=acr25114-sup-0001-Disclosureform.pdf.

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

Submitted for publication October 11, 2022; accepted in revised form March 9, 2023.

SIGNIFICANCE & INNOVATIONS

- Computed tomography (CT) is a commonly used test for cancer screening/surveillance in newly diagnosed idiopathic inflammatory myopathy (IIM) patients.
- In a tertiary referral IIM population, age <40 years at IIM onset was associated with lower diagnostic yield and higher frequency of false-positive results on CT imaging, even in dermatomyositis patients.
- These findings suggest that cancer detection strategies targeted according to IIM subtype, autoantibody positivity, and age may maximize cancer detection while minimizing the harms and costs of over-screening.

(4,5). Therefore, the current consensus is that most newly diagnosed IIM patients should be evaluated for cancer, yet few studies exist to help inform the optimal approach for cancer assessment.

Complicating matters further, a common theme in published cohort studies of IIM and cancer is the large variety of cancer types reported, including ovarian, breast, lung, lymphoma, cervical, and gastrointestinal malignancies (2,6). Consequently, computed tomography (CT) of the chest, abdomen, and pelvis has become common to assess IIM patients for contemporaneous cancer. However, minimal data exists regarding the diagnostic yield of CT imaging in IIM, particularly by IIM subtype and MSA positivity. In this study, we demonstrate cancer diagnostic yield using a standardized approach for cancer assessment by CT in a large, single-center, tertiary referral IIM population. The study specifically focuses on IIM subtype, age, and myositis-specific and associated autoantibodies, as these have been proven to be useful tools in the risk stratification of contemporaneous cancer (4,5,7,8).

PATIENTS AND METHODS

We reviewed the electronic medical records of all patients who previously provided informed consent for enrollment in our myositis research registry from 2003 to 2020 in order to identify those who met the following criteria: 1) probable or definite DM according to definitions by Bohan and Peter (9), 2) immune-mediated necrotizing myopathy (IMNM) according to the 2003 European Neuromuscular Center Criteria (10), 3) classic DM rash (Gottron's/heliotrope) and consistent histopathology on skin biopsy, and 4) antisynthetase syndrome (ASyS) defined as the presence of an ASyS autoantibody with ≥1 of the following features: an inflammatory myopathy, interstitial lung disease (ILD), arthritis, Raynaud's syndrome, fever, or mechanic's hands (11). Patients with clear diagnoses of muscular dystrophy, metabolic or mitochondrial myopathy, or inclusion body myositis were excluded, even if they technically met criteria for probable/definite polymyositis (PM). A flow chart of patient inclusion is shown (Figure 1).

MSAs and myositis-associated autoantibodies were assayed on the first available banked serum sample from each patient (median disease duration 1.7 years) using Euroimmun line blot (16 antigen IgG, unless otherwise noted for Mi-2, TIF1γ, and nuclear matrix protein 2 [NXP2]), enzyme-linked immunosorbent assay (anti–Mi-2 and anti-TIF1γ [MBL]), in-house immunoprecipitation (anti–NXP-2), and anti–hydroxymethylglutaryl-coenzyme A reductase (anti-HMGCR) (Inova Diagnostics) as previously described (5). Anti-TIF1γ was considered positive based on 2 thresholds: >32 units according to the manufacturer's instructions, and >6 units (3 SDs more than the healthy control cohort mean as previously described [12]). Patients were determined to have seronegative DM/IMNM if they met DM/IMNM criteria as described above but were negative for all assayed autoantibodies. We systematically reviewed each patient chart and outside medical records for any CT of the chest or abdomen/pelvis ordered for cancer surveillance/screening.

The strategy for malignancy surveillance/screening in the Johns Hopkins Myositis Center includes CT of the chest, abdomen, and pelvis with IV contrast upon IIM diagnosis. Furthermore, clinicians recommend age- and sex-appropriate cancer screening tests, including PAP smear, colonoscopy, mammogram, prostate-specific antigen, and routine dermatologist skin evaluations. Additional testing (CA 125, transvaginal ultrasound, positron emission tomography/CT, as well as repeat CT imaging) is ordered at the discretion of the treating clinician. Tests for cancer were marked as performed if: 1) they were performed at Johns Hopkins, 2) if records were scanned from an outside institution or accessible via interoperability platforms such as Care Everywhere within the Johns Hopkins electronic medical record, or 3) the treating physician documented the specifics of the test in the encounter note. Given that some CT scans had multiple indications for ordering (e.g., ILD), scans were only included in these analyses if documentation existed that cancer surveillance/screening was at least one of the motivators for ordering. Furthermore, while our center uses IV contrast for all CT cancer imaging, CT imaging performed outside of our institution was included even if IV contrast status was not known. With regard to cancer ascertainment, every patient was contacted by research staff by either phone or online survey to update their cancer status using a systematic script (5).

The Standards for Reporting of Diagnostic Accuracy Studies 2015 guidelines were followed in this study (13). Diagnostic yield was calculated as the number of cancers diagnosed divided by the number of tests performed. True positives were defined as a CT scan finding that led to a biopsy-confirmed cancer diagnosis. Premalignant lesions and non-melanoma skin cancers were excluded. False positives were defined as a CT scan finding that was ultimately biopsied and deemed noncancerous. True negatives were defined as a negative CT scan for cancer and the patient not developing cancer within 3 years of IIM symptom onset. False negatives were defined as a CT scan negative for

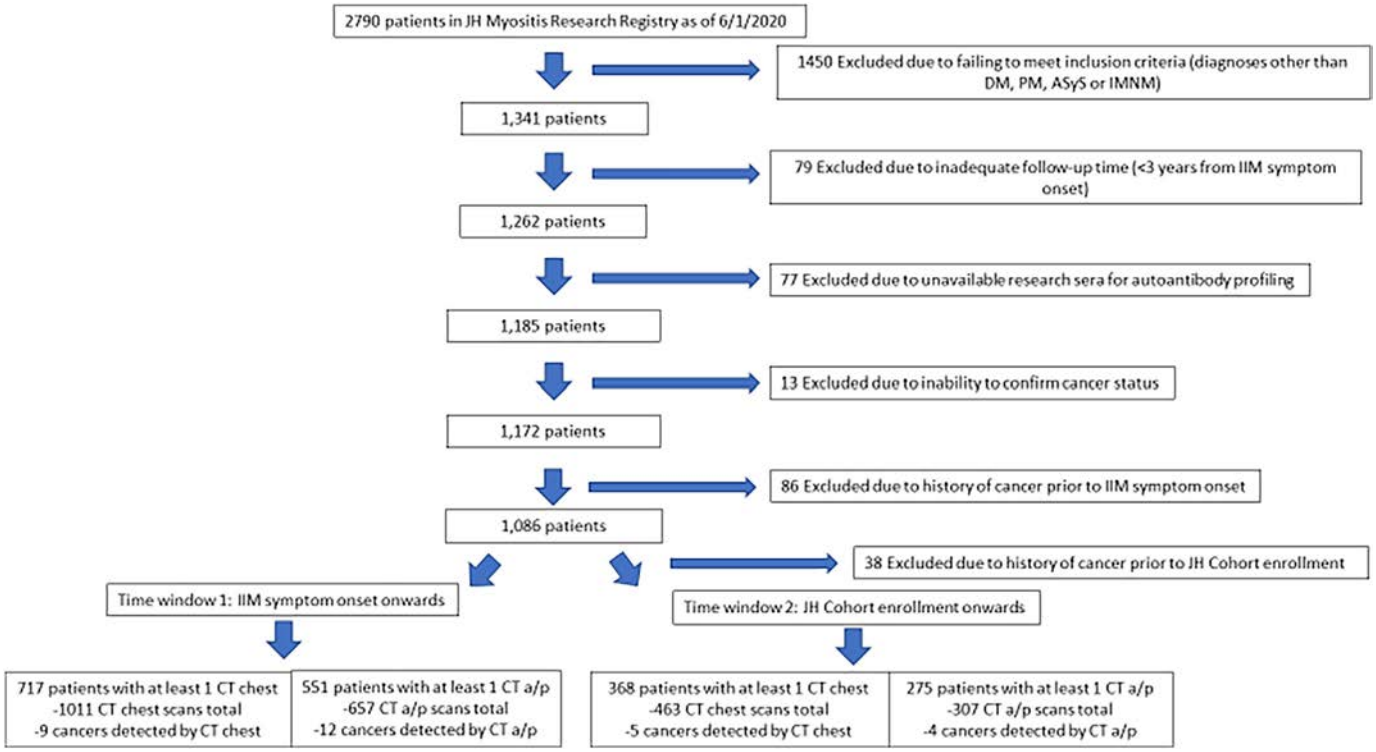

Figure 1. Flow diagram describing patient inclusion and computed tomography (CT) imaging performed. a/p = abdomen/pelvis; ASyS = antisynthetase syndrome; DM = dermatomyositis; IIM = idiopathic inflammatory myopathy; IMNM = immune-mediated necrotizing myopathy; JH = Johns Hopkins; PM = polymyositis.

cancer, but cancer was diagnosed within 12 months of the scan (all cancers were included, not just those restricted to the location that was scanned). A second investigator checked data abstraction for accuracy in a randomly selected study population (20% of all patients). We calculated the positive likelihood ratio (LR) as (sensitivity/1 – specificity) and calculated the negative LR as (1 – sensitivity/specificity). We also calculated 95% confidence intervals (95% CIs) for sensitivity, specificity, positive, and negative LRs as previously described (14).

Analyses were stratified according to IIM subtype (DM, PM, IMNM, ASyS, seronegative DM, seronegative IMNM), age at IIM onset (<40 or ≥40 years old), and MSAs and myositis-associated autoantibodies. Two periods were analyzed: cancer surveillance/screening tests performed 1) after IIM symptom onset and 2) after myositis cohort enrollment, both through June 30, 2020. Patients with a history of cancer prior to IIM onset/cohort enrollment were excluded from their respective analyses, since these patients likely differed biologically and thus were likely screened more intensely for cancer by clinical providers. All eligible CT scans were included in the analysis until any of the following three events occurred: 1) first cancer was diagnosed, 2) a CT scan was positive for cancer, or 3) patient had IIM disease duration of 3 years, whichever came first. Last, sensitivity analyses were performed, examining the time window of 0–12 months after IIM onset to replicate the clinical practice environment outside of tertiary referral centers.

RESULTS

Patient cohort. Patient selection based on inclusion and exclusion criteria is shown in Figure 1. A total of 1,086 patients were the focus of this study. Patient demographic and disease characteristics are shown in Supplementary Table 1 (available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25114/abstract). The median age of IIM onset was 49 years (interquartile range [IQR] 38–58 years); 71% of patients were female; and 68% of the cohort was White, 21% were Black, 3.6% were Asian, and 7.4% were other/unknown. The time from IIM symptom onset to cohort enrollment was a median of 1.7 years (IQR 0.7–3.9 years), and the median follow-up duration was 5.3 years (IQR 0.75–3.90 years). Of 1,086 patients, 62 had cancer diagnosed within the first 3 years since IIM onset, of which, 30 diagnoses occurred after cohort entry. Cancer types and stages are shown in Supplementary Table 2 (available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25114/abstract). Breast cancer was the most common cancer (19%), followed by melanoma (13%) and cervical/uterine (10%).

Percentage of patients undergoing CT imaging within the first 3 years of IIM symptom onset. Of the 1,086 patients, a total of 1,011 chest CT scans (in 717 unique patients) and 657 abdomen/pelvis CT scans (in 551 unique patients) were performed from IIM onset onward (Figure 1). In our cohort, ~66% of all IIM patients underwent CT of the chest within 3 years of IIM symptom onset, whereas 51% underwent CT of the abdomen/pelvis (Supplementary Tables 3 and 4, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25114/abstract). The range of these percentages varied greatly depending on the IIM subtype and autoantibody subgroup examined, with chest CT completion ranging from 47% to 91%, and abdomen/pelvis CT completion ranging from 25% to 73%. Patients who did not undergo CT imaging were significantly younger (median age 45.5 years old for chest CT scan and 47.0 years old for abdomen/pelvis CT scan) than patients who did undergo screening (chest 50.5 [$P < 0.001$], abdomen/pelvis 49.6 [$P = 0.016$]).

Chest CT scans—IIM symptom onset onward. The number of cancers diagnosed, chest CT scans performed, and diagnostic yield of chest CT scans within the first 3 years of IIM symptom onset is shown in Table 1. The number of cancers diagnosed within the first 3 years of disease was 62. The percentage of chest CT scans leading to a cancer diagnosis was 9 of 1,011 (0.9%). Of note, a similar diagnostic yield was observed when examining per patient screened rather than per CT performed: 6 cancers detected on initial chest CT scans performed divided by 717 patients screened = 0.8%.

The subgroup with the highest diagnostic yield was DM patients ages ≥40 years old at IIM onset (1.4% [7 of 509]), and this was largely driven by anti-TIF1γ–positive patients (Supplementary Table 5, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25114/abstract). No patient with PM or ASyS had a positive chest CT scan for cancer at ages <40 years, and no IMNM patient had a positive chest CT scan at any age (Table 1). With regard to false-positive CT of the chest, the number of scans leading to a noncancerous biopsy was 28 of 1,011 (2.8%) within 3 years of IIM symptom onset. The subgroups with highest frequency of false-positive CT scans include PM, IMNM and ASyS (3.9% [13 of 331], 4.4% [7 of 160], and 4.4% [13 of 293], respectively). Patients with anti–Jo-1 autoantibodies had a diagnostic yield of 0 and false-positive frequency of 4.1% (7 of 169) (Supplementary Table 5). The most commonly biopsied lesions leading to false-positive results were pulmonary nodules, thyroid nodules, and nonspecific lymphadenopathy. Based on our data, CT imaging of the chest in ASyS and IMNM patients are associated with the most harm from a cancer screening perspective. Additional data stratified by MSAs and myositis-associated autoantibodies is shown in Supplementary Tables 5 and 6 (available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25114/abstract).

The sensitivity, specificity, and LRs for CT of the chest are shown in Table 2. The positive LR for a positive chest CT scan in all IIM patients was 8.2 (95% CI 4.2–16.2) in DM patients ages >40 years old at IIM onset was 12.9 (95% CI 5.6–29.6) and in patients with anti-TIF1γ was, 25.6 (95% CI 5.3–123.1). In all groups, sensitivity for CT of the chest was low (<0.50), indicating that CT imaging did not detect many cancers that were diagnosed in our cohort.

Chest CT scans—cohort enrollment onward. In the analysis for cohort enrollment onward, only patient data collected after enrollment in our cohort study was included (Supplementary Tables 7 and 8, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25114/abstract). This time window was analyzed to understand the yield and outcomes of CT imaging obtained in

Table 1. Cancer and chest CT scan data collected within the first 3 years since IIM symptom onset*

Population, age at IIM onset	No. of patients	No. cancers within 0–3 years (total/asymptomatic)	Positive CT scans	Total CT scans	Diagnostic yield, %	False-positive CT scans	False positives, %
All	1,086	62/40	9	1,011	0.9	28	2.8
<40 years	328	8/6	0	260	0.0	5	1.9
≥40 years	758	54/34	9	751	1.2	23	3.1
Dermatomyositis	678	46/32	7	680	1.0	15	2.2
<40 years	204	7/6	0	171	0.0	3	1.8
≥40 years	474	39/26	7	509	1.4	12	2.4
Polymyositis	408	16/8	2	331	0.6	13	3.9
<40 years	124	1/0	0	89	0.0	2	2.2
≥40 years	284	15/8	2	242	0.8	11	4.5
IMNM	234	16/9	0	160	0.0	7	4.4
<40 years	64	1/0	0	44	0.0	0	0.0
≥40 years	170	15/9	0	116	0.0	7	6.0
Antisynthetase	234	4/1	1	293	0.3	13	4.4
<40 years	74	0/0	0	85	0.0	3	3.5
≥40 years	160	4/1	1	208	0.5	10	4.8

* Patients with cancer diagnoses without positive computed tomography (CT) imaging reflect cancer diagnoses made using other methods (e.g., mammogram, skin biopsy, etc.). IIM = idiopathic inflammatory myopathy; IMNM = immune-mediated necrotizing myopathy.

Table 2. Chest CT scan characteristics within the first 3 years since IIM symptom onset*

Population, age at IIM onset	True positives†	False positives‡	True negatives§	False negatives¶	Sensitivity (95% CI)	Specificity (95% CI)	Positive LR (95% CI)#	Negative LR (95% CI)**
All	9	28	945	29	0.24 (0.13–0.39)	0.97 (0.96–0.98)	8.2 (4.2–16.2)	0.8 (0.7–0.9)
<40 years	0	5	248	7	0.00 (0.0–0.35)	0.98 (0.95–0.99)	–	–
≥40 years	9	23	697	22	0.29 (016–0.47)	0.97 (0.95–0.98)	9.1 (4.6–18)	0.7 (0.6–0.9)
Dermatomyositis	7	15	636	22	0.24 (0.12–0.42)	0.98 (0.96–0.99)	10.5 (4.6–23.7)	0.8 (0.6–1.0)
<40 years	0	3	161	7	0.00 (0.0–0.35)	0.98 (0.95–0.99)	–	–
≥40 years	7	12	475	15	0.32 (0.16–0.53)	0.98 (0.96–0.99)	12.9 (5.6–29.6)	0.7 (0.5–0.9)
Polymyositis	2	13	309	7	0.22 (0.06–0.55)	0.96 (0.93–0.98)	5.5 (1.5–20.9)	0.8 (0.6–1.2)
<40 years	0	2	87	0	–	–	–	–
≥40 years	2	11	222	7	0.22 (0.06–0.55)	0.95 (0.92–0.97)	4.7 (1.2–18.2)	0.8 (0.6–1.2)
IMNM	0	7	146	7	0.00 (0.0–0.35)	0.95 (0.91–0.98)	–	–
<40 years	0	0	44	0	–	–	–	–
≥40 years	0	7	102	7	0.00 (0.0–0.35)	0.94 (0.87–0.97)	–	–
Antisynthetase	1	13	278	1	0.50 (0.09–0.91)	0.96 (0.93–0.97)	11.2 (2.5–49.4)	0.5 (0.1–2.1)
<40 years	0	3	82	0	–	–	–	–
≥40 years	1	10	196	1	0.50 (0.09–0.91)	0.95 (0.91–0.97)	10.3 (2.3–46.7)	0.5 (0.1–2.1)

* 95% CI = 95% confidence interval. See Table 1 for other definitions.
† True positive indicates a CT scan revealed findings that led to cancer diagnosis.
‡ False positive indicates a CT scan revealed findings that was ultimately biopsied and noncancerous.
§ True negative indicates a CT scan was unremarkable and the patient did not develop cancer within 3 years of IIM symptom onset.
¶ False negative indicates a CT scan was unremarkable but cancer was diagnosed within 12 months of the scan.
# Positive likelihood ratio (LR): sensitivity/(1 − specificity).
** Negative LR: (1 − sensitivity)/specificity.

patients after their referral to a tertiary care center. In this analysis, the number of cancers within 3 years of IIM symptom onset was 30. The number of true-positive chest CT scans was 5, and the number of false-positive CT scans was 10, corresponding to a diagnostic yield of 1.1% and false-positive frequency of 2.2%. Similar trends were observed for this time window (cohort enrollment onward) compared to IIM symptom onset onward, with the diagnostic yield highest in DM and the highest false-positive frequencies in PM, IMNM, and ASyS. When examining diagnostic yields per patient screened rather than per CT performed, the result was similar: 5 cancers detected on initial CT of the chest performed divided by 368 patients screened = 1.3%.

Abdomen/pelvis CT scans—IIM symptom onset onward. The number of abdomen/pelvis CT scans performed, diagnostic yield, and percent of false positives within the first 3 years of IIM symptom onset is shown in Table 3. The percentage of abdomen/pelvis CT scans leading to a cancer diagnosis was 12 of 657 (1.8%). A similar diagnostic yield was observed when examining per patient screened rather than per CT performed: 10 cancers detected on initial CT of the abdomen/pelvis performed divided by 551 patients screened = 1.8%.

Similar to chest CT scans, yields were highest in DM patients ages >40 years old at IIM onset (2.7% [9 of 334]), and lowest in PM patients (1.0% [2 of 200]) and ASyS patients (0% [0 of 104]). Autoantibody groups with the highest diagnostic yield include anti-TIF1γ (2.4% [3 of 125]) and anti–small ubiquitin-like modifier activating enzyme hetermodimer (14.3% [2 of 14]) (Supplementary Table 9, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25114/abstract). With regard to false-positive abdomen/pelvis CT scans, a total of 14 of 657 (2.1%) were false positives (Table 3). The highest frequencies of false-positive abdomen/pelvis CT scans were in DM patients <40 years old (4.9%, [6 of 123]) and ASyS patients (3.8% [4 of 104]), driven by patients with anti-TIF1γ, anti–Mi-2, anti–Jo-1, and anti–PL-12 autoantibodies. Additional data stratified by MSAs and myositis-associated autoantibodies are shown in Supplementary Tables 9 and 10 (available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25114/abstract). Upon restricting our study population to patients entered into our cohort within the first 12 months since IIM symptom onset, a total of 2 patients had a positive abdomen/pelvis CT scan within this time window. Both patients had DM; one had anti-TIF1γ antibodies, and the other was anti–NXP-2 positive.

The sensitivity, specificity, and LRs for CT of the abdomen/pelvis are shown in Table 4. The highest positive LRs were observed in older patients (ages >40 years old) across multiple IIM subtypes: 29.2 (all IIM patients ages >40 years old), 28.3 (DM patients ages >40 years old), 33.0 (PM patients ages >40 years old), and 26 (IMNM patients ages >40 years old). Similar to CT of the chest, the sensitivity of abdomen/pelvis CT was low (<0.50), indicating that CT imaging did not detect many cancers that were diagnosed in our cohort.

Abdomen/pelvis CT scans—cohort enrollment onward. In the analysis for cohort enrollment onward, only patient data collected after enrollment into our cohort study was included (Supplementary Tables 11 and 12, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25114/abstract). In these analyses, the number of cancers within 3 years of IIM symptom onset was 30; the number of true-positive abdomen/pelvis CT scans was 4, and the number of false-positive CT scans was 7, corresponding to a diagnostic yield of 1.3% and a false-positive frequency of 2.3%. Similar trends were observed in this time window (cohort enrollment onward) compared to IIM symptom onset onward, with the diagnostic yield highest in DM and DM-specific autoantibodies (anti-TIF1γ, anti-SAE, anti–Mi-2,

Table 3. Cancer and abdomen/pelvis CT scan data collected within the first 3 years since IIM symptom onset*

Population, age at IIM onset	No. of patients	No. cancers within 0–3 years (total/asymptomatic)	Positive CT scans	Total CT scans	Diagnostic yield, %	False-positive CT scans	False positives, %
All	1,086	62/40	12	657	1.8	14	2.1
<40 years	328	8/6	1	183	0.5	8	4.4
≥40 years	758	54/34	11	474	2.3	6	1.3
Dermatomyositis	678	46/32	10	457	2.2	11	2.4
<40 years	204	7/6	1	123	0.8	6	4.9
≥40 years	474	39/26	9	334	2.7	5	1.5
Polymyositis	408	16/8	2	200	1.0	3	1.5
<40 years	124	1/0	0	60	0.0	2	3.3
≥40 years	284	15/8	2	140	1.4	1	0.7
IMNM	234	16/9	3	122	2.5	2	1.6
<40 years	64	1/0	0	35	0.0	1	2.9
≥40 years	170	15/9	3	87	3.4	1	1.1
Antisynthetase	234	4/1	0	104	0.0	4	3.8
<40 years	74	0/0	0	33	0.0	2	6.1
≥40 years	160	4/1	0	71	0.0	2	2.8

* Patients with cancer diagnoses without positive CT imaging reflect cancer diagnoses made using other methods (e.g., mammogram, skin biopsy, etc.). See Table 1 for definitions.

Table 4. Abdomen/pelvis CT scan characteristics within the first 3 years since IIM symptom onset*

Population, age at IIM onset	True positives†	False positives‡	True negatives§	False negatives¶	Sensitivity (95% CI)	Specificity (95% CI)	Positive LR (95% CI)#	Negative LR (95% CI)**
All	12	14	611	20	0.38 (0.23–0.55)	0.98 (0.96–0.99)	16.7 (8.4–33.2)	0.6 (0.5–0.8)
<40 years	1	8	171	3	0.25 (0.05–0.70)	0.96 (0.91–0.98)	5.6 (0.9–34.8)	0.8 (0.4–1.4)
≥40 years	11	6	440	17	0.39 (0.24–0.58)	0.99 (0.97–0.99)	29.2 (11.7–73.2)	0.6 (0.5–0.8)
Dermatomyositis	10	11	422	14	0.42 (0.24–0.61)	0.97 (0.96–0.99)	16.4 (7.7–34.8)	0.6 (0.4–0.8)
<40 years	1	6	113	3	0.25 (0.05–0.70)	0.95 (0.89–0.98)	5.0 (0.8–32.1)	0.8 (0.4–1.4)
≥40 years	9	5	309	11	0.45 (0.26–0.66)	0.98 (0.96–0.99)	28.3 (10.4–76.5)	0.6 (0.4–0.8)
Polymyositis	2	3	189	6	0.25 (0.07–0.59)	0.98 (0.96–0.99)	16.0 (3.1–82.8)	0.8 (0.5–1.1)
<40 years	0	2	58	0	–	–	–	–
≥40 years	2	1	131	6	0.25 (0.07–0.59)	0.99 (0.96–1.0)	33.0 (3.3–326.5)	0.8 (0.5–1.1)
IMNM	3	2	111	6	0.33 (0.12–0.65)	0.98 (0.94–1.0)	18.8 (3.6–98.6)	0.7 (0.4–1.1)
<40 years	0	1	34	0	–	–	–	–
≥40 years	3	1	77	6	0.33 (0.12–0.65)	0.99 (0.93–1.0)	26.0 (3.0–224.4)	0.7 (0.4–1.1)
Antisynthetase	0	4	99	1	0.00 (0.0–0.79)	0.96 (0.90–0.98)	–	–
<40 years	0	2	31	0	–	–	–	–
≥40 years	0	2	68	1	0.00 (0.0–0.79)	0.97 (0.90–0.99)	–	–

* 95% CI = 95% confidence interval. See Table 1 for other definitions.
† True positive indicates a CT scan revealed findings that led to cancer diagnosis.
‡ False positive indicates a CT scan revealed findings that was ultimately biopsied and noncancerous.
§ True negative indicates a CT scan was unremarkable and the patient did not develop cancer within 3 years of IIM symptom onset.
¶ False negative indicates a CT scan was unremarkable but cancer was diagnosed within 12 months of the scan.
# Positive likelihood ratio (LR): sensitivity/(1 − specificity).
** Negative LR: (1 − sensitivity)/specificity.

anti–NXP-2) and highest false-positive frequencies in PM and ASyS. When examining diagnostic yields per patient screened rather than per CT performed, the result was similar: 4 cancers detected on initial CT abdomen/pelvis performed divided by 275 patients screened = 1.5%.

Yield of repeat CT imaging. At the discretion of the treating provider, additional CT scans may be ordered for the purpose of cancer surveillance. In these cases, information regarding each individual CT scan was collected, which enabled us to determine the yield of serial CT imaging; that is, in patients who had an initial negative CT scan, the number/percentage with a subsequent positive CT scan for cancer. After IIM symptom onset, 3 of 291 chest CT scans were positive following the initial negative CT scan of the chest. The 3 cancers detected were thyroid, metastatic undifferentiated carcinoma (biopsied in the lung), and non–small-cell lung carcinoma—these patients were anti-SAE, anti-TIF1γ, and antibody-negative IMNM. Similarly, 2 of 104 abdomen/pelvis CT scans were positive after the initial negative CT scan of the abdomen/pelvis. The two cancers detected were cholangiocarcinoma and metastatic undifferentiated carcinoma (biopsied in the liver); these patients were anti-HMGCR and anti-SAE positive.

DISCUSSION

We observed a large range in the diagnostic yield of CT imaging to detect contemporaneous cancer in IIM patients. The diagnostic yield is greatly dependent on the IIM subtype (DM, PM, IMNM, ASyS), autoantibody, and age at IIM onset. Overall, CT of the chest and abdomen/pelvis had the highest diagnostic yield in DM patients, specifically anti-TIF1γ–positive patients, but the lowest yield in patients with PM and those with ASyS. Furthermore, we demonstrated that false-positive scans are much more common in specific subgroups, highlighting the potential harms of CT imaging. Taken together, these findings provide new information of potential utility in informing clinical decision-making with regard to cancer screening in newly diagnosed IIM patients.

For both CT of the chest and abdomen/pelvis, the diagnostic yield in patients with DM and DM-specific autoantibodies (anti-TIF1γ) is ∼2–4%, corresponding to a need for 30–40 CT scans to diagnose 1 cancer in this subgroup. Because both CT of the chest and abdomen/pelvis have similar diagnostic yield, this likely reflects the heterogeneity of cancer types that occur in DM patients (5,15). Conversely, patients with PM, ASyS, or IMNM had high percentages of false positives for both CT of the chest and abdomen/pelvis. However, the relatively low yield of CT chest scans for cancer in ASyS patients must be considered in the context of the high prevalence of ILD in this subgroup. That is, while CT of the chest may appear to have a low diagnostic yield from a cancer detection perspective, this does not consider the potential value of assessing and monitoring ILD longitudinally. It is notable, however, that no ASyS patient had a positive CT of the abdomen/pelvis for cancer in our cohort.

It is helpful when viewing these data to provide appropriate context—a 2% diagnostic yield for CT of the chest might be viewed as high by some and low by others. In lung cancer screening trials (enriched for older adults with prominent smoking history), the diagnostic yields were ∼1% (number of low-dose chest CT scans needed to detect 1 lung cancer was ∼100) (16,17). In randomized controlled trials examining CT colonography for detecting colorectal malignancy, studies demonstrate a 0.5% yield (5 of 982). In IIM cancer screening studies specifically, there is a lack of data on CT imaging (18). One of the most comprehensive studies to date is by Leatham et al, where 29 cancers were diagnosed in 27 patients, 6 using CT of the chest, abdomen, or pelvis. However, the total number of CT scans performed was not reported, and thus a diagnostic yield could not be calculated (19).

The range of LR varied widely in our study, with the highest LRs observed in DM patients. For context, a negative LR ranges between 0 and 1, and a positive LR ranges from >1 to infinity. The further LRs are from 1, the stronger the evidence of the presence (>1) or absence (<1) of disease (e.g., cancer). In general, LRs are considered most useful if the positive LR is >10 or negative LR is <0.10 (20,21). In our study, multiple subgroups had LRs >10, most consistently found in DM and DM-specific autoantibodies. Essentially, no subgroup had LRs <0.10.

A crucial unanswered question from these data is whether CT imaging impacts clinical outcomes—both cancer-related outcomes (number of scans needed to prevent 1 cancer-associated death) as well as IIM-related outcomes (does earlier detection/elimination of cancer lead to a more favorable IIM clinical trajectory). Ultimately this level of data will be needed in IIM patients to guide clinical decision-making, and well-designed studies aimed at answering these questions should be a top priority in the field.

Our study has several limitations. The yield of CT of the chest and abdomen/pelvis in IIM cohorts is largely dependent on 1) overall prevalence of cancer in the cohort and 2) cancer types/sites within the cohort. Given our decision to exclude IIM patients with a prior cancer history, along with the fact that many cancers diagnosed in our cohort are not conventionally detected on CT scan (breast, melanoma), the yield of CT imaging overall may be lower in our study than in other IIM–cancer cohorts. Our definition of a false-positive test was conservative (e.g., requiring biopsy). Other less stringent definitions (e.g., requiring serial imaging/additional testing, etc.) are not captured in our data set. Defining false-positive tests with a biopsy most certainly reduces their number; however, the relative frequency of false positives is likely preserved when comparing IIM subgroups (i.e., ASyS patients have higher frequency of false positives compared to DM patients ages >40 years old). Furthermore, this information is clinically

useful in patient discussions regarding the risks and benefits of imaging.

Additional limitations of our study include that not all patients had a CT of the chest and abdomen/pelvis. The possibility of selection bias therefore exists in our results, as other nonmeasured factors may have influenced clinicians to order (or not order) CT imaging. This may artificially inflate the diagnostic yield of CT imaging. In addition, while every effort was made to obtain all outside hospital and office visit notes, we likely did not completely capture all cancer testing, particularly for studies performed prior to entry in our cohort. This may impact our calculations of diagnostic yield and percent of false positives. Last, since our study spans 17 years, the possibility of changing cancer screening practices exists, particularly since subgroups of IIM were shown to be at higher risk of cancer during this period (e.g., anti-TIF1γ–positive DM patients). During this 17-year timeframe, our chest/abdomen/pelvis CT screening practice remained largely the same; however, some high-risk patients may have been more likely to undergo additional CT imaging (serial scans).

In conclusion, our findings demonstrate that CT of the chest and abdomen/pelvis has the highest diagnostic yield in patients with DM and those with DM-specific MSAs (anti-TIF1γ, anti–NXP-2, anti-SAE, and anti–Mi-2), while it has the lowest diagnostic yield in patients with ASyS and those with PM. Importantly, DM patients <40 years old at IIM onset have a low diagnostic yield. Furthermore, we demonstrate the potential harms of CT imaging: false-positive studies are much more common in PM and IMNM, specifically patients with anti-HMGCR (CT of the chest) and antisynthetase antibodies (CT of the chest and abdomen/pelvis). In our cohort, the yield of repeat CT imaging was low. These data provide evidence that can help inform clinical decision-making to maximize cancer detection while minimizing the harms and costs of over-screening in this patient population.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Mecoli, Platz, Christopher-Stine, Shah.
Acquisition of data. Mecoli, Chen, Wang, Albayda, Paik, Tiniakou, Adler, Mammen, Casciola-Rosen, Christopher-Stine, Shah.
Analysis and interpretation of data. Mecoli, Chee, Albayda, Paik, Tiniakou, Adler, Kelly, Mammen, Platz, Casciola-Rosen, Christopher-Stine, Shah.

REFERENCES

1. Sigurgeirsson B, Lindelof B, Edhag O, et al. Risk of cancer in patients with dermatomyositis or polymyositis: a population-based study. N Engl J Med 1992;326:363–7.
2. Hill CL, Zhang Y, Sigurgeirsson B, et al. Frequency of specific cancer types in dermatomyositis and polymyositis: a population-based study. Lancet 2001;357:96–100.
3. Fiorentino DF, Mecoli CA, Rosen MC, et al. Immune responses to CCAR1 and other dermatomyositis autoantigens are associated with attenuated cancer emergence. J Clin Invest 2022;132:e150201.
4. Yang H, Peng Q, Yin L, et al. Identification of multiple cancer-associated myositis-specific autoantibodies in idiopathic inflammatory myopathies: a large longitudinal cohort study. Arthritis Res Ther 2017; 19:259.
5. Mecoli CA, Igusa T, Chen M, et al. Subsets of idiopathic inflammatory myositis enriched for contemporaneous cancer relative to the general population. Arthritis Rheumatol 2023;75:620–9.
6. Dani L, Che WI, Lundberg IE, et al. Overall and site-specific cancer before and after diagnosis of idiopathic inflammatory myopathies: a nationwide study 2002–2016. Semin Arthritis Rheum 2021;51:331–7.
7. Betteridge Z, McHugh N. Myositis-specific autoantibodies: an important tool to support diagnosis of myositis. J Intern Med 2016;280: 8–23.
8. Chinoy H, Fertig N, Oddis CV, et al. The diagnostic utility of myositis autoantibody testing for predicting the risk of cancer-associated myositis. Ann Rheum Dis 2007;66:1345–9.
9. Bohan A, Peter JB. Polymyositis and dermatomyositis (first of two parts). N Engl J Med 1975;292:344–7.
10. Hoogendijk JE, Amato AA, Lecky BR, et al. 119th ENMC international workshop: trial design in adult idiopathic inflammatory myopathies, with the exception of inclusion body myositis, 10–12 October 2003, Naarden, The Netherlands. Neuromuscul Disord 2004;14:337–45.
11. Selva-O'Callaghan A, Pinal-Fernandez I, Trallero-Araguas E, et al. Classification and management of adult inflammatory myopathies. Lancet Neurol 2018;17:816–28.
12. Fiorentino DF, Gutierrez-Alamillo L, Hines D, et al. Distinct dermatomyositis populations are detected with different autoantibody assay platforms. Clin Exp Rheumatol 2019;37:1048–51.
13. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ 2015;351:h5527.
14. Simel DL, Samsa GP, Matchar DB. Likelihood ratios with confidence: sample size estimation for diagnostic test studies. J Clin Epidemiol 1991;44:763–70.
15. Oldroyd A, Sergeant JC, New P, et al. The temporal relationship between cancer and adult onset anti-transcriptional intermediary factor 1 antibody-positive dermatomyositis. Rheumatology (Oxford) 2019;58:650–5.
16. De Koning HJ, van der Aalst CM, de Jong PA, et al. Reduced lung-cancer mortality with volume CT screening in a randomized trial. N Engl J Med 2020;382:503–13.
17. Aberle DR, Adams, AM, Berg CD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med 2011;365:395–409.
18. Stoop EM, de Haan MC, de Wijkerslooth TR, et al. Participation and yield of colonoscopy versus non-cathartic CT colonography in population-based screening for colorectal cancer: a randomised controlled trial. Lancet Oncol. 2012;13:55–64.
19. Leatham H, Schadt C, Chisolm S, et al. Evidence supports blind screening for internal malignancy in dermatomyositis: data from 2 large US dermatology cohorts. Medicine (Baltimore) 2018;97: e9639.
20. Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. BMJ 2004; 329:168–9.
21. McGee S. Simplifying likelihood ratios. J Gen Intern Med 2002;17: 646–9.

Arthritis Care & Research
Vol. 75, No. 10, October 2023, pp 2151–2157
DOI 10.1002/acr.25130

American College of Rheumatology
Empowering Rheumatology Professionals

Impact of Television Depictions of Gout on Perceptions of Illness: A Randomized Controlled Trial

Rachel Murdoch, Borislav Mihov, Anne M. Horne, Keith J. Petrie, Gregory D. Gamble, and Nicola Dalbeth

Objective. Gout is a chronic disease that can be effectively managed with long-term urate-lowering therapy. However, it is frequently portrayed on screen as an acute disease caused by a poor diet that should be managed with lifestyle changes. This study was undertaken to investigate the impact of a fictional television depiction of gout on perceptions of the disease and its management.

Methods. In a randomized controlled single-blind study, 200 members of the public watched either a 19-minute commercial television comedy episode that depicted gout as an acute disease caused by poor diet and managed with lifestyle changes, or a control episode from the same television series that did not mention gout or other diseases. Participants completed a survey regarding their perceptions of gout, its likely causes, and management strategies.

Results. Participants randomized to watch the gout-related episode believed gout had greater consequences (mean score of 7.1 versus 6.2 on an 11-point Likert scale; $P < 0.001$) and were more likely to rank the most important cause as poor eating habits compared to the control group (70% versus 38%; $P < 0.001$). They were also less likely to believe it is caused by genetic factors or chance. Participants watching the gout-related episode believed a change in diet would be a more effective management strategy (9.0 versus 8.4; $P = 0.004$) and long-term medication use would be less effective (6.9 versus 7.6; $P = 0.007$) compared to participants in the control group.

Conclusion. Television depictions of gout can perpetuate inaccurate beliefs regarding causes of the disease and underemphasize effective medical strategies required in chronic disease management.

INTRODUCTION

Gout is a common rheumatic disease caused by the deposition of monosodium urate (MSU) crystals in the setting of hyperuricemia (1,2). Although gout typically presents as recurrent episodes of acute inflammatory arthritis (gout flares), it is a chronic condition of MSU crystal deposition. Biologic factors, including genetic variants (3) and chronic kidney disease (4), play an important role in the development of hyperuricemia and gout. For recurrent gout flares, rheumatology guidelines recommend long-term urate-lowering therapy (ULT) (5,6), which leads to crystal dissolution and suppression of gout flares (7–9). In contrast, dietary change is only recommended as adjunctive therapy (5,6): although certain foods may play a role in triggering disease flares disease, dietary interventions alone are not sufficient to control serum urate or disease activity in gout (10,11). Despite the known efficacy of long-term medication, gout remains an under-treated disease with low rates of ULT (12). Beliefs that gout is an acute arthritis that can be managed through dietary change rather than long-term ULT may contribute to low treatment rates (13,14).

Historically, gout has been viewed as related to overindulgence (15). In modern society, it is viewed as caused by an individual's behavior, through poor diet and overconsumption of alcohol (13,16). Contemporary cultural depictions of gout focus on dietary rather than biologic factors. In newspaper articles about gout, overindulgence in food and alcohol was the most reported cause (17). In a recent content analysis of gout in contemporary film and television, dietary choices and alcohol were the most common causes depicted, and the disease was frequently portrayed as humorous and embarrassing. The most common management strategies described in those on-screen depictions were change in diet (36%) and short-term pain relief (32%), with ULT rarely mentioned (18).

ANZCTR no.: 383779.
Supported by the New Zealand Rheumatology Association.
The New Zealand Rheumatology Association did not alter the study design, data collection, data analysis, data interpretation, or writing of the report.
Rachel Murdoch, MBChB, Borislav Mihov, BPhty, Anne M. Horne, MBChB, Keith J. Petrie, PhD, Gregory D. Gamble, MSc, Nicola Dalbeth, MD: University of Auckland, Auckland, New Zealand.

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.25130&file=acr25130-sup-0001-Disclosureform.pdf.
Address correspondence via email to Nicola Dalbeth, MD, at n.dalbeth@auckland.ac.nz.
Submitted for publication October 25, 2022; accepted in revised form April 4, 2023.

SIGNIFICANCE & INNOVATIONS

- Most popular cultural depictions of gout portray the disease as an acute illness that is caused by dietary discretion and can be managed with lifestyle changes.
- The impact of these depictions on public perceptions of gout has not been formally studied.
- This randomized controlled trial demonstrates that television depictions of gout can reinforce misconceptions about the disease and effective management strategies.

Television is a popular source of entertainment and information for the general public. In one study, the Centers for Disease Control and Prevention found that 64% of respondents were regular watchers of television, and 67% of regular daytime drama viewers reported learning something new about a health issue or disease from a television show in the previous 6 months (19). Television is a preferred source of medical information for many members of the public (20,21), including some patients with gout (22), and compared to print media, television depictions can have a strong influence on patient-reported symptoms (23,24).

There is a substantial need for improved community knowledge regarding gout and its management. To date, the effects of on-screen depictions on perceptions of gout have not been reported. This study aimed to determine whether television depictions of gout contribute to community perceptions of gout and beliefs regarding its causes and management.

PATIENTS AND METHODS

Study design and participants. In this single-blind randomized controlled trial, 200 participants were recruited from the community from February 2022 to July 2022 by advertising for participants for a study about "perceptions of illness" on social media and a job search website in Aotearoa/New Zealand. Participants were not aware that the study focused specifically on gout at the time of recruitment. To meet inclusion criteria, patients had to be >18 years old and able to complete and understand a questionnaire in English. The trial was approved by the University of Auckland Human Participants Ethics Committee (approval no. UAHPEC3277) and registered with the Australian New Zealand Clinical Trials Registry (ANZCTR no. 383779). Written informed consent was obtained from all participants.

Randomization and masking. The randomization schedule was created by a statistician (GDG) using a pseudorandom number generator within Microsoft Excel. A total of 6 blocks of variable size with 10–30 participants each were created. Within each block, participants included in the study were sorted into 2 groups: the random numbers in the lower half of the block were allocated to watch the gout-related episode, and the random numbers in the upper half of the block were allocated to watch the control episode. The randomization schedule was not available to the researcher who carried out the participant surveys.

Procedures. Participants were randomly assigned (in a 1:1 ratio) to watch either a 19-minute commercial television comedy episode depicting gout or a similar-length control episode from the same television series. A researcher emailed participants a link to watch the gout or control episode along with instructions on how to view the episode; this researcher had no further contact with study participants. A separate researcher (RM) who was blinded with regard to the allocation group completed a survey over Zoom with each participant after they had watched the episode. Participants were instructed not to identify the episode they had watched to the researcher administering the survey.

The gout and control episodes were from the television sitcom "Everybody Hates Chris." The gout-related episode, "Everybody Hates Chris, Everybody Hates the Gout" (season 1 episode 16) was identified in the prior content analysis of television episodes depicting gout (18) and involved a character who experienced a gout flare "trying to figure out how he got the gout." He was seen by a doctor who described gout as caused by "a poor diet," including "smothered pork chops," "smothered chicken fried bacon," and "smothered chicken fried bananas." His family blamed him for having gout, encouraged him to eat salad and raw vegetables, and prevented him from eating pancakes. Gout was shown as a humorous and embarrassing disease, with the character experiencing gout referred to as "gout boy." He appeared to be in severe pain and was unable to work. There was no explanation of the biologic causes of gout and no mention of ULT.

This episode was chosen since it was representative of the findings of the prior content analysis depicting gout as caused by lifestyle choices, such as poor diet, and with biologic causes not explored. It also portrayed gout as humorous and embarrassing, in a similar way to the majority of episodes reviewed in the previous review of film and television episodes depicting gout (18). The episode "Everybody Hates Chris, Everybody Hates the Lottery" (season 1 episode 15) was selected as the control episode since it was the episode immediately prior in the same television series, featured the same characters, and did not mention gout or other diseases.

Participant demographic data were collected, including sex, age, ethnicity, and occupation. Prioritized ethnicity was self-reported using the standard ethnicity question from the New Zealand census (25). Participants were asked about their personal or family history of rheumatoid arthritis (RA) or gout.

After viewing the episode, all participants completed an interview with the researcher (RM) who was blinded to their study allocation. Participants confirmed that they had watched the television episode prior to the interview. Illness perceptions were assessed using the Brief Illness Perception Questionnaire (BIPQ) (26), which

Table 1. Demographic characteristics in participants assigned to either watch a gout-related episode of a television program or a control episode of the same program*

Demographic characteristic	All participants (n = 200)	Gout episode group (n = 100)	Control episode group (n = 100)
Sex			
Female	153 (76.5)	75 (75)	78 (78)
Male	47 (23.5)	25 (25)	22 (22)
Age, mean ± SD (range)	29 ± 11 (18–70)	27 ± 10 (18–66)	31 ± 12 (19–70)
Ethnicity			
Māori	15 (7.5)	8 (8)	7 (7)
Pacific peoples	2 (1)	1 (1)	1 (1)
Asian	42 (21)	22 (22)	20 (20)
New Zealand European	113 (56.5)	57 (57)	56 (56)
Other	28 (14)	12 (12)	16 (16)
Diagnosis of gout			
Personal diagnosis	0 (0)	0 (0)	0 (0)
Diagnosis in family member or close friend	36 (18)	16 (16)	20 (20)

* Except where indicated otherwise, values are the number (%) of patients.

assesses key illness perception domains on an 11-point Likert scale, including consequences (how much the illness would affect a person's life), timeline (how long it would continue), personal control (how much control a person would have over the illness), treatment control (how much medication would help), identity (how severe symptoms would be), concern (how concerned a person would be about the illness), emotions (how much the illness would affect a person emotionally), and understanding (how well a person understands the illness). Participants were able to view the questionnaire over Zoom while completing it with the researcher.

Participants were asked to list the 3 most important causes of gout, with their answers grouped into diet, genetics, alcohol, lifestyle, and aging for analysis. The groupings were determined based on the responses by 2 researchers (RM and ND). The responses were grouped by one researcher (RM) with any unclear responses discussed with a second researcher (ND). These determinations were made with researchers blinded with regard to study allocation. The groupings were simplified from those used in a previous study that assessed perceptions of patients with gout using the BIPQ (27). Participants also rated possible causes of gout, such as aging, diet, and hereditary factors, on a 5-point scale from "strongly disagree" to "strongly agree." The survey asked about management strategies, such as the likely efficacy of dietary change, exercise, long-term medication use, alcohol cessation, and losing weight, which participants could rate on an 11-point scale from "would not help it at all" to "very likely to help." Missing and incomplete questionnaires were excluded.

As an internal control and to assess whether the depictions of gout also influenced views regarding other diseases, participants were asked about their perceptions of RA using identical questions from the BIPQ.

Statistical analysis. The sample size of 200 participants was based on a previous study of diseases labeled "urate crystal arthritis" or "gout" (16). This sample size allowed the detection of a ≥2-point difference for each BIPQ item with 90% power to detect a significant difference and an alpha level of 0.05. *P* values were adjusted for multiple comparisons with Bonferroni correction. The primary end point was disease perception, measured using the BIPQ. Differences between the groups in ratings of illness perceptions, causes, and management strategies were tested with *t*-tests for independent samples. With large sample sizes, the use of parametric analysis methods has been validated for any distribution (28).

Table 2. Perceptions of gout using the BIPQ according to whether a participant watched a gout-related episode of a television program or a control episode of the same program*

BIPQ item†	Gout episode group	Control episode group	Mean difference (95% CI)	*P*‡
Consequences	7.1 ± 1.4	6.2 ± 1.6	0.9 (0.5, 1.3)	<0.001
Timeline	5.4 ± 2.2	7.2 ± 2.0	−1.7 (−2.3, −1.2)	<0.001
Personal control	5.8 ± 1.8	5.5 ± 2.0	0.2 (−0.3, 0.8)	0.390
Treatment control	6.3 ± 2.0	6.7 ± 1.6	−0.5 (−1.0, 0.1)	0.077
Identity	7.0 ± 2.1	7.0 ± 1.8	−0.0 (−0.6, 0.5)	0.97
Concern	6.7 ± 2.1	6.8 ± 2.0	−0.0 (−0.6, 0.5)	0.89
Understanding	4.8 ± 2.1	3.6 ± 2.1	1.2 (0.6, 1.8)	<0.001
Emotional response	6.6 ± 1.8	6.1 ± 1.9	0.5 (−0.1, 1.0)	0.083

* Except where indicated otherwise, values are the mean ± SD score. 96% CI = 95% confidence interval; Brief Illness Perception Questionnaire = BIPQ.
† Scale from 0, indicating "no effect at all" to 10, indicating "severely affect my life."
‡ Adjusted for multiple comparisons using Bonferroni correction. *P* values <0.006 were statistically significant.

Table 3. Perception of most likely cause of gout on the BIPQ according to whether a participant watched a gout-related episode of a television program or a control episode of the same program*

Grouped cause	Gout episode group (n = 100)	Control episode group (n = 100)	Percentage difference (95% CI)	P†
Diet	70 (70)	38 (38)	32 (19, 45)	<0.001
Genetics	10 (10)	26 (26)	−16 (−26, −5.9)	0.0040
Alcohol	5 (5)	11 (11)	−6 (−13, 1.5)	0.13
Lifestyle	3 (3)	7 (7)	−4 (−10, 2)	0.22
Aging	3 (3)	7 (7)	−4 (−10, 2)	0.22

* Except where indicated otherwise, values are the number (%) of participants who listed a cause as the most important cause of gout. 95% CI = 95% confidence interval; BIPQ = Brief Illness Perception Questionnaire.
† Adjusted for multiple comparisons using Bonferroni correction. *P* values <0.01 were considered significant.

Data availability. Data requests for anonymized data should be submitted to the corresponding author for consideration.

RESULTS

Between February 21, 2022 and July 31, 2022, 206 participants were screened for eligibility and inclusion in the study. Six participants were unable to attend the study interview and were not enrolled. A total of 200 participants met inclusion criteria, consented to the study, and were randomized to groups to watch the gout-related episode (n = 100) or control episode (n = 100). All 200 participants completed the study interview, and the primary outcome was assessed. Participant demographic characteristics are shown in Table 1. Most participants were of New Zealand European ethnicity (113 [56.5%]) and were female (153 [76.5%]), and the average age was 29 years (range 18–70 years). A total of 15 (7.5%) participants were of Māori ethnicity (Indigenous New Zealanders). No participants had a personal diagnosis of gout, and 36 (18%) had a family member or friend with gout.

Participants who watched the gout-related episode believed they had a greater understanding of gout (mean ± SD 4.8 ± 2.1 versus 3.6 ± 2.1; $P < 0.001$), that gout would last a shorter time (mean ± SD 5.4 ± 2.2 versus 7.2 ± 2.0; $P < 0.001$), and that gout would have a greater effect on their life (mean ± SD 7.1 ± 1.4 versus 6.2 ± 1.6; $P < 0.001$) compared to the control group (Table 2). There was no between-group difference in terms of the other BIPQ items and no difference in perceptions of RA assessed using the BIPQ (see Supplementary Table 1, available on the Arthritis Care & Research website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25130/abstract).

When asked about the most important causes of gout, most participants who viewed the gout-related episode listed diet as the most important cause of gout (70 [70%]) compared to participants in the control group (38 [38%]) ($P < 0.001$) (Table 3). Those who had watched the control episode were more likely to list genetics as the most important cause of gout (26 [26%] versus 10 [10%]; $P = 0.0040$).

Participants were also asked to rate different causes of gout. Those who watched the gout-related episode believed diet was a more important cause (mean ± SD 4.6 ± 0.6 versus 4.2 ± 0.9; $P < 0.001$), and that hereditary factors (mean ± SD 3.3 ± 1.0 versus 3.8 ± 0.8; $P = 0.001$), and chance (mean ± SD 2.4 ± 1.1 versus 2.8 ± 1.1; $P = 0.002$) were less important causes, compared to the control group (Table 4). There was no difference between the 2 groups when rating stress, aging, alcohol use, pollution, a germ or virus, or the behavior of an individual as possible causes of gout.

Participants who watched the gout-related episode were less likely to believe patients would require long-term medication

Table 4. Perception of causes of gout according to whether a participant watched a gout-related episode of a television program or a control episode of the same program*

Cause	Gout episode group	Control episode group	Mean difference (95% CI)	P†
Stress or worry	2.8 ± 0.1	3.0 ± 1.0	−0.3 (−0.5, 0.0)	0.063
Hereditary	3.3 ± 1.0	3.8 ± 0.8	−0.4 (−0.7, −0.2)	0.001
Diet or poor eating habits	4.6 ± 0.6	4.2 ± 0.9	0.4 (0.2, 0.6)	<0.001
Aging	3.7 ± 0.9	3.8 ± 0.8	−0.1 (−0.3, 0.2)	0.61
Alcohol use	3.8 ± 0.8	3.8 ± 1.0	0.1 (−0.2, 0.3)	0.69
Pollution	2.3 ± 0.8	2.5 ± 0.8	−0.2 (−0.4, 0.1)	0.16
Chance or bad luck	2.4 ± 1.1	2.8 ± 1.1	−0.5 (−0.8, −0.2)	0.002
A germ or virus	2.4 ± 1.1	2.5 ± 1.1	−0.1 (−0.4, 0.2)	0.44
Individual behavior	3.7 ± 1.0	3.5 ± 1.0	0.2 (−0.1, 0.4)	0.22

* Except where indicated otherwise, values are the mean ± SD score (scale 1–5, with 1 indicating "strongly disagree" and 5 indicating "strongly agree"). 95% CI = 95% confidence interval.
† Adjusted for multiple comparisons using Bonferroni correction. *P* values <0.006 were considered significant.

Table 5. Views about gout medications according to whether a participant watched a gout-related episode of a television program or a control episode of the same program*

View about medication	Gout episode group	Control episode group	Mean difference (95% CI)	P†
Need for long-term medications	5.6 ± 2.4	6.6 ± 1.8	−1.0 (−1.6, −0.4)	0.001
Motivation to take long-term medications	6.8 ± 2.7	7.0 ± 2.5	−0.2 (−0.9, 0.5)	0.57
Concern about long-term medication use	5.6 ± 2.5	5.9 ± 2.8	−0.3 (−1.0, 0.5)	0.47

* Except where indicated otherwise, values are the mean ± SD score (scale 1–10, with 1 indicating "not at all" and 10 indicating "extremely"). 95% CI = 95% confidence interval.
† Adjusted for multiple comparisons using Bonferroni correction. *P* values <0.02 were considered significant.

than those who watched the control episode (mean ± SD 5.6 ± 2.4 versus 6.6 ± 1.8; $P = 0.001$). There was no difference between the 2 groups in motivation to take long-term medications or concern regarding long-term medication use (Table 5).

Those who watched the gout-related episode were more likely to believe changing to a healthier diet would be an effective management strategy compared to participants in the control group (mean ± SD 9.0 ± 1.4 versus 8.4 ± 1.7; $P = 0.004$) and were less likely to think that taking long-term medication would be effective (mean ± SD 6.9 ± 1.8 versus 7.6 ± 1.7; $P = 0.007$) (Table 6).

DISCUSSION

Findings from this randomized controlled single-blind study demonstrate that viewing a television episode that portrays gout as a humorous disease caused by dietary choices strongly influences public perceptions of gout. Viewers of the gout-related episode had a perception of greater understanding of the disease but were more likely to believe it is an acute rather than chronic disease and that its management is mainly dietary rather than through ULT. Watching a depiction of gout did not appear to influence beliefs about RA, indicating that the influence of television depictions is disease specific. These results align with previous studies of other health conditions, which showed that fictional depictions of disease on television influence viewer understanding of the disease (29,30); although in contrast to previous studies that showed an improvement in understanding with realistic depictions, this study shows the negative impact of inaccurate representations of disease.

This study analyzed the impact of viewing a single television episode. It is possible that repeated exposure to similar depictions has an even larger effect. Previous studies have shown repeated exposure to storylines about medical conditions such as the *BRCA* mutation amplify the effect on viewer knowledge and attitudes (31). A previous content analysis of gout on film and television showed the majority of descriptions focus on lifestyle factors (18). The high frequency of inaccurate depictions of gout is concerning, since repeated exposure is likely to reinforce misconceptions held by the public. Perceptions of disease have clinical relevance, since illness perceptions affect how likely patients are to adhere to ULT (32), and predict musculoskeletal disability after 1 year in patient with gout (27).

It is not known whether the impact of television depictions of gout on public beliefs leads to a change in behavior, such as willingness to take ULT. However, research in other conditions has shown that patients who have seen cardiopulmonary resuscitation (CPR) on television are more likely to express a wish to receive CPR (33), and restricting exposure to smoking on film results in decreased uptake of smoking in adolescents (34–36). Targeting misconceptions about other diseases has been shown to affect behavior (37–39): patients experiencing a myocardial infarction assigned to receive sessions with a psychologist addressing illness perceptions had an improved functional outcome and returned to work more quickly than individuals in a control group receiving

Table 6. Management strategies according to whether a participant watched a gout-related episode of a television program or a control episode of the same program*

Management strategy	Gout episode group	Control episode group	Mean difference (95% CI)	P†
Managing stress	5.9 ± 2.0	6.2 ± 2.1	−0.3 (−0.8, 0.3)	0.36
Regular exercise	6.9 ± 2.1	7.1 ± 2.0	−0.3 (−0.8, 0.3)	0.37
Reducing/stopping alcohol	7.9 ± 1.6	7.8 ± 2.1	0.1 (−0.5, 0.6)	0.79
Long-term medication	6.9 ± 1.8	7.6 ± 1.7	−0.7 (−1.2, −0.2)	0.007
Changing to a healthier diet	9.0 ± 1.4	8.4 ± 1.7	0.6 (0.2, 1.1)	0.004
Alternative medication	5.3 ± 2.1	5.4 ± 2.4	−0.0 (−0.7, 0.6)	0.93
Losing weight	7.2 ± 2.0	7.2 ± 1.9	−0.1 (−0.6, 0.5)	0.80

* Except where indicated otherwise, values are the mean ± SD score (scale 1–10, with 1 indicating "wouldn't help it at all" and 10 indicating "very likely to help"). 95% CI = 95% confidence interval.
† Adjusted for multiple comparisons using Bonferroni correction. *P* values <0.007 were considered significant.

standard cardiac rehabilitation (37). Entertainment-education is a strategy of incorporating educational messages into media that draws on theories such as homophily (that people are more likely to connect with those with similar characteristics) to influence the beliefs and behavior of viewers (40) that has been shown to be effective in other diseases such as HIV (41).

Television depictions could play a role in representing the experience of patients and their communities. The gout-related episode in this study emphasized the main character's pain and inability to work, and participants watching the depiction of gout considered it a more severe disease than the control group. Accurate television depictions of gout could have positive effects, such as increasing public awareness about the impact of gout.

Strengths of this study included that it was a randomized controlled trial, the researcher (RM) completing the study interviews was blinded with regard to study allocation, and the study was powered to detect small changes in the BIPQ. This is the first study demonstrating that fictional television depictions of gout have a substantial impact on public perceptions. The study format optimized the understandability of information for participants through use of a visual medium (television) and individual interviews with patients. Participants in the study tended to be young, with an average age of 29 years, and most participants were female, which may affect generalizability to the wider population. No participants had a diagnosis of gout, which affects the generalizability of the results to patients with gout, although 18% of the participants had a family member or close friend with gout. Examining public perceptions is relevant because it is a disease where patients experience high levels of stigma (42) and patients are often told that having gout is due to dietary choices (43), which is likely to contribute to uptake of ineffective management strategies.

Improving the accuracy of depictions of gout could help the family and community of the patient to support them in their treatment plan rather than suggest ineffective alternatives. Limitations of this study included the short follow-up following a single viewing, and it is unknown how long the differences in illness beliefs persist after viewing a single episode, or the effects of repeated exposure to similar on-screen depictions. A television episode showing gout caused by biologic factors could not be included as a comparison because no suitable shows were identified in the previous content analysis (18), as almost all depictions of gout focused on dietary causes.

In summary, gout is commonly depicted on screen as a humorous disease caused by dietary choices, which can reinforce misconceptions about the disease and its effective management strategies. Increasing the accuracy of depictions of gout could improve understanding of its treatment, increase representation of the difficulties experienced by patients with gout, and lead to better management of gout in the community. Organizational guidelines about depictions of gout, for example from non–governmental organizations (NGOs), arthritis support NGOs, professional societies, and advocacy groups that could be used by medical advisors to media companies may improve the accuracy of television depictions. Clinicians have an opportunity to address common misconceptions by discussing these issues with patients and their families and supporting patients to receive accurate information regarding the disease. Writers and producers of television programs have the opportunity to improve the public's understanding of gout and the uptake of effective therapies by portraying accurate depictions of the disease.

ACKNOWLEDGMENTS

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

AUTHOR CONTRIBUTIONS

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

Study conception and design. Murdoch, Petrie, Gamble, Dalbeth.
Acquisition of data. Murdoch, Mihov, Horne, Dalbeth.
Analysis and interpretation of data. Murdoch, Petrie, Gamble, Dalbeth.

REFERENCES

1. Jin Z, Wang D, Zhang H, et al. Incidence trend of five common musculoskeletal disorders from 1990 to 2017 at the global, regional and national level: results from the global burden of disease study 2017. Ann Rheum Dis 2020;79:1014–22.
2. Loeb JN. The influence of temperature on the solubility of monosodium urate. Arthritis Rheum 1972;15:189–92.
3. Tin A, Marten J, Halperin Kuhns VL, et al. Target genes, variants, tissues and transcriptional pathways influencing human serum urate levels. Nat Genet 2019;51:1459–74.
4. Chonchol M, Shlipak MG, Katz R, et al. Relationship of uric acid with progression of kidney disease. Am J Kidney Dis 2007;50:239–47.
5. FitzGerald JD, Dalbeth N, Mikuls T, et al. 2020 American College of Rheumatology guideline for the management of gout. Arthritis Rheumatol 2020;72:879–95.
6. Richette P, Doherty M, Pascual E, et al. 2018 updated European League Against Rheumatism evidence-based recommendations for the diagnosis of gout. Ann Rheum Dis 2020;79:31–8.
7. Pascual E, Sivera F. Time required for disappearance of urate crystals from synovial fluid after successful hypouricaemic treatment relates to the duration of gout. Ann Rheum Dis 2007;66:1056–8.
8. Perez-Ruiz F, Calabozo M, Pijoan JI, et al. Effect of urate-lowering therapy on the velocity of size reduction of tophi in chronic gout. Arthritis Rheum 2002;47:356–60.
9. Doherty M, Jenkins W, Richardson H, et al. Efficacy and cost-effectiveness of nurse-led care involving education and engagement of patients and a treat-to-target urate-lowering strategy versus usual care for gout: a randomised controlled trial. Lancet 2018;392:1403–12.
10. Juraschek SP, White K, Tang O, et al. Effects of a Dietary Approach to Stop Hypertension (DASH) diet intervention on serum uric acid in African Americans with hypertension. Arthritis Care Res (Hoboken) 2018;70:1509–16.

11. Holland R, McGill NW. Comprehensive dietary education in treated gout patients does not further improve serum urate. Intern Med J 2015;45:189–94.
12. Kuo CF, Grainge MJ, Mallen C, et al. Rising burden of gout in the UK but continuing suboptimal management: a nationwide population study. Ann Rheum Dis 2015;74:661–7.
13. Lindsay K, Gow P, Vanderpyl J, et al. The experience and impact of living with gout: a study of men with chronic gout using a qualitative grounded theory approach. J Clin Rheumatol 2011;17:1–6.
14. Spencer K, Carr A, Doherty M. Patient and provider barriers to effective management of gout in general practice: a qualitative study. Ann Rheum Dis 2012;71:1490–5.
15. Porter R. Gout: framing and fantasizing disease. Bull Hist Med 1994; 68:1–28.
16. Petrie KJ, MacKrill K, Derksen C, et al. An illness by any other name: the effect of renaming gout on illness and treatment perceptions. Health Psychol 2018;37:37–41.
17. Duyck SD, Petrie KJ, Dalbeth N. "You don't have to be a drinker to get gout, but it helps": a content analysis of the depiction of gout in popular newspapers. Arthritis Care Res (Hoboken) 2016;68:1721–5.
18. Derksen C, Murdoch R, Petrie KJ, et al. "An apple pie a day does not keep the doctor away": fictional depictions of gout in contemporary film and television. BMC Rheumatol 2021;5:4.
19. Centers for Disease Control and Prevention. TV drama/comedy viewers and health information: 2005 Porter Novelli HealthStyles Survey. 2005. URL: https://www.cdc.gov/healthcommunication/pdf/healthstyles_2005.pdf.
20. Al-Sane M, Bourisly N, Almulla T, et al. Laypeoples' preferred sources of health information on the emergency management of tooth avulsion. Dent Traumatol 2011;27:432–7.
21. James C, James N, Davies D, et al. Preferences for different sources of information about cancer. Patient Educ Couns 1999;37:273–82.
22. Treharne GJ, Richardson AC, Neha T, et al. Education preferences of people with gout: exploring differences between indigenous and non-indigenous peoples from rural and urban locations. Arthritis Care Res (Hoboken) 2018;70:260–7.
23. MacKrill K, Gamble GD, Petrie KJ. The effect of television and print news stories on the nocebo responding following a generic medication switch. Clin Psychol Eur 2020;2:e2623.
24. Faasse K, Gamble G, Cundy T, et al. Impact of television coverage on the number and type of symptoms reported during a health scare: a retrospective pre-post observational study. BMJ Open 2012;2: e001607.
25. Ministry of Health. Ethnicity data protocols. 2017. URL: https://www.moh.govt.nz/notebook/nbbooks.nsf/0/569DDB6A56F7E726CC2581FC00665BEB/$file/hiso-10001-2017-ethnicity-data-protocols.pdf.
26. Broadbent E, Petrie KJ, Main J, et al. The brief illness perception questionnaire. J Psychosom Res 2006;60:631–7.
27. Dalbeth N, Petrie KJ, House M, et al. Illness perceptions in patients with gout and the relationship with progression of musculoskeletal disability. Arthritis Care Res (Hoboken) 2011;63:1605–12.
28. Lumley T, Diehr P, Emerson S, et al. The importance of the normality assumption in large public health data sets. Annu Rev Public Health 2002;23:151–69.
29. Brodie M, Foehr U, Rideout V, et al. Communicating health information through the entertainment media. Health Aff (Millwood) 2001;20: 192–9.
30. Rosenthal EL, de Castro Buffington S, Cole G. From the small screen to breast cancer screening: examining the effects of a television storyline on awareness of genetic risk factors. J Commun Healthc 2018; 11:140–50.
31. Hether HJ, Huang GC, Beck V, et al. Entertainment-education in a media-saturated environment: examining the impact of single and multiple exposures to breast cancer storylines on two popular medical dramas. J Health Commun 2008;13:808–23.
32. Walsh CP, Prior JA, Chandratre P, et al. Illness perceptions of gout patients and the use of allopurinol in primary care: baseline findings from a prospective cohort study. BMC Musculoskelet Disord 2016; 17:394.
33. Moss AH, Hozayen O, King K, et al. Attitudes of patients toward cardiopulmonary resuscitation in the dialysis unit. Am J Kidney Dis 2001;38:847–52.
34. Distefan JM, Pierce JP, Gilpin EA. Do favorite movie stars influence adolescent smoking initiation? Am J Public Health 2004;94:1239–44.
35. Sargent JD, Beach ML, Dalton MA, et al. Effect of seeing tobacco use in films on trying smoking among adolescents: cross sectional study. BMJ 2001;323:1394–7.
36. Sargent JD, Beach ML, Dalton MA, et al. Effect of parental R-rated movie restriction on adolescent smoking initiation: a prospective study. Pediatrics 2004;114:149–56.
37. Petrie KJ, Cameron LD, Ellis CJ, et al. Changing illness perceptions after myocardial infarction: an early intervention randomized controlled trial. Psychosom Med 2002;64:580–6.
38. Broadbent E, Ellis CJ, Thomas J, et al. Further development of an illness perception intervention for myocardial infarction patients: a randomized controlled trial. J Psychosom Res 2009;67:17–23.
39. Broadbent E, Ellis CJ, Thomas J, et al. Can an illness perception intervention reduce illness anxiety in spouses of myocardial infarction patients? A randomized controlled trial. J Psychosom Res 2009;67: 11–5.
40. Moyer-Gusé EJ. Toward a theory of entertainment persuasion: explaining the persuasive effects of entertainment-education messages. Commun Theory 2008;18:407–25.
41. Pappas-DeLuca KA, Kraft JM, Galavotti C, et al. Entertainment-education radio serial drama and outcomes related to HIV testing in Botswana. AIDS Educ Prev 2008;20:486–503.
42. Kleinstäuber M, Wolf L, Jones AS, et al. Internalized and anticipated stigmatization in patients with gout. ACR Open Rheumatol 2020;2: 11–7.
43. Te Karu L, Bryant L, Elley CR. Maori experiences and perceptions of gout and its treatment: a kaupapa Maori qualitative study. J Prim Health Care 2013;5:214–22.

Arthritis Care & Research
Vol. 75, No. 10, October 2023, pp 2158–2165
DOI 10.1002/acr.25115

American College of Rheumatology
Empowering Rheumatology Professionals

Evaluation of Measurement Properties and Differential Item Functioning in the English and French Versions of the University of California, Los Angeles, Loneliness Scale-6: A Scleroderma Patient-Centered Intervention Network (SPIN) Study

Chelsea S. Rapoport,[1] Alyssa K. Choi,[1] Linda Kwakkenbos,[2] Marie-Eve Carrier,[3] Richard S. Henry,[4] Luc Mouthon,[5] Scott C. Roesch,[6] Brett D. Thombs,[4] Vanessa L. Malcarne,[6] the SPIN COVID-19 Patient Advisory Team, and the SPIN Investigators

Objective. Loneliness has been associated with poorer health-related quality of life but has not been studied in patients with systemic sclerosis (SSc). The current study was undertaken to examine and compare the psychometric properties of the English and French versions of the University of California, Los Angeles, Loneliness Scale-6 (ULS-6) in patients with SSc during the COVID-19 pandemic.

Methods. This study used baseline cross-sectional data from 775 adults enrolled in the Scleroderma Patient-Centered Intervention Network (SPIN) COVID-19 Cohort. Reliability and validity of ULS-6 scores overall and between languages were evaluated using confirmatory factor analysis (CFA), differential item functioning (DIF) through the multiple-indicator multiple-cause (MIMIC) model, omega/alpha calculation, and correlations of hypothesized convergent relationships.

Results. CFA for the total sample supported the single-factor structure (comparative fit index [CFI] 0.96, standardized root mean residual [SRMR] 0.03), and all standardized factor loadings for items were large (0.60–0.86). The overall MIMIC model with language as a covariate fit well (CFI 0.94, SRMR 0.04, root mean square error of approximation 0.11). Statistically significant DIF was found for 3 items across language (β_{item2} = 0.14, $P < 0.001$; β_{item4} = –0.07, $P = 0.01$; β_{item6} = 0.13, $P < 0.001$), but these small differences were without practical measurement implications. Analyses demonstrated high internal consistency with no language-based convergent validity differences.

Conclusion. Analyses demonstrated evidence of acceptable reliability and validity of ULS-6 scores in English- and French-speaking adults with SSc. DIF analysis supported use of the ULS-6 to examine comparative experiences of loneliness without adjusting for language.

INTRODUCTION

Loneliness is a pervasive and distressing experience involving a person's perception that their social relationships do not fulfill their social needs (1). It is an especially important problem among individuals with chronic illnesses, as their symptoms may subject them to greater challenges integrating in social and work settings (2). Specifically, patients with autoimmune rheumatic

Supported by the Canadian Institutes of Health Research (grants VR4-172745 and GA4-177764) and the Mi4 McGill Interdisciplinary Initiative in Infection and Immunity Emergency COVID-19 Research Fund. The Scleroderma Patient-Centered Intervention Network (SPIN) has received funding from the Canadian Institutes of Health Research, Arthritis Society Canada, and the Canadian Institute for Outcomes in Rheumatology Care. SPIN has also received contributions from the Lady Davis Institute for Medical Research of the Jewish General Hospital, the Jewish General Hospital Foundation, McGill University, the Scleroderma Society of Ontario, Scleroderma Canada, Sclérodermie Québec, Scleroderma Manitoba, Scleroderma Atlantic, the Scleroderma Association of BC, Scleroderma Saskatchewan, Scleroderma Australia, Scleroderma New South Wales, Scleroderma Victoria, and Scleroderma Queensland.

[1]Chelsea S. Rapoport, BA, Alyssa K. Choi, BA: San Diego State University/University of California San Diego Joint Doctoral Program in Clinical Psychology, San Diego, California; [2]Linda Kwakkenbos, PhD: Radboud University, Nijmegen, The Netherlands; [3]Marie-Eve Carrier, MSc: Jewish General Hospital, Montreal, Quebec, Canada; [4]Richard S. Henry, PhD, Brett D. Thombs, PhD: Jewish General Hospital and McGill University, Montreal, Quebec, Canada; [5]Luc Mouthon, MD: Reference Center for Rare Systemic Autoimmune Diseases of Ile de France, Hôpital Cochin, AP-HP, Université Paris Descartes, Paris, France; [6]Scott C. Roesch, PhD, Vanessa L. Malcarne, PhD: San Diego State University/University of California San Diego Joint Doctoral Program in Clinical Psychology and San Diego State University, San Diego, California.

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.25115&file=acr25115-sup-0001-Disclosureform.pdf.

Address correspondence via email to Vanessa L. Malcarne, PhD, at vmalcarne@sdsu.edu.

Submitted for publication November 29, 2022; accepted in revised form March 16, 2023.

SIGNIFICANCE & INNOVATIONS

- The COVID-19 pandemic has caused higher levels of loneliness globally, but this has not been explored in patients with autoimmune rheumatic diseases. Patients with systemic sclerosis (SSc) have increased risks of COVID-19–related pulmonary involvement, immunosuppressive medication use, and general frailty and may face higher rates of loneliness and its subsequent physical and mental health consequences.
- No studies have explored loneliness in patients with SSc, and no measures for loneliness had been validated prior to this study.
- There were not measurement differences that affected scores between the English and French versions of the University of California, Los Angeles, Loneliness Scale-6, supporting the combined use of English and French data for analysis and comparison in future research on loneliness in SSc.

diseases may experience high symptom burden, which can lead to disability and isolation from others (3).

The COVID-19 pandemic has caused unprecedented challenges for the global population due to social distancing and isolation, and a systematic review found small post-pandemic increases in loneliness compared to pre-pandemic (34 studies, n = 215,026; standardized mean difference 0.27 [95% confidence interval (95% CI) 0.14, 0.40]); however, only 1 study of patients with chronic health conditions was included (4). Systemic sclerosis (SSc) is a rare autoimmune disorder that damages the skin and connective tissue, and SSc-related symptoms, such as chronic fatigue and pain, reduce health-related quality of life (5). Patients with SSc in particular are at higher risk of poor COVID-19 outcomes given their general frailty and immunosuppressive medication use, and because interstitial lung disease is found in ∼40% of patients (6,7). There is scant research, however, on loneliness in autoimmune rheumatic diseases and none in SSc (2). No measures to assess loneliness have been evaluated in autoimmune rheumatic diseases.

The University of California, Los Angeles (UCLA), Loneliness Scale-6 (ULS-6) is a 6-item short form of the revised UCLA Loneliness Scale (R-ULS), a 20-item self-report measure that has been used in multiple populations, including patients with chronic illnesses (2,8). The 6 items for the ULS-6 were selected in a sample of 286 Portuguese adolescents based on an exploratory factor analysis of the R-ULS that found that they loaded onto an "isolation and withdrawal" factor, which was determined to capture the essence of the loneliness construct (9). A subsequent confirmatory factor analysis (CFA) of older Portuguese adults (n = 1,154) found that the ULS-6 showed acceptable fit with the predicted single-factor model (10). No studies have assessed the measurement properties of the ULS-6 in English or French or in individuals with chronic illnesses.

The aim of the present study was to assess the measurement properties of ULS-6 scores during the COVID-19 pandemic for patients with SSc overall and separately in English and French. The specific objectives were to evaluate structural validity and to determine whether there was differential item functioning (DIF) between English- and French-language responses, internal consistency, and convergent validity overall and within and between language samples.

PATIENTS AND METHODS

This was a cross-sectional study that analyzed data from participants enrolled in the Scleroderma Patient-Centered Intervention Network (SPIN) COVID-19 Cohort. The SPIN COVID-19 Cohort study was approved by the Research Ethics Committee of the CIUSSS du Centre-Ouest-de-l'Île-de-Montréal. This report was documented in accordance with COSMIN guidelines (11). See Appendix A for additional members of the SPIN COVID-19 Patient Advisory Team and the SPIN investigators and their affiliations.

Participants and procedure. Participants were recruited from the ongoing SPIN Cohort and additionally via social media and patient organization advertisements (12). The SPIN Cohort includes over 1,800 participants from 47 centers in Canada, the US, the UK, France, Spain, Mexico, and Australia who complete regular 3-month online assessments. SPIN Cohort participants must be age ≥18 years, fluent in English, French, or Spanish, and meet the 2013 American College of Rheumatology/EULAR criteria for SSc, verified by a SPIN physician (13). SPIN Cohort participants provide informed consent for participation and for future contact about additional SPIN studies. SPIN site personnel submit an online medical form post-consent to enroll participants, who then receive instructions via email to activate SPIN accounts and complete measures in English, French, or Spanish. Participants complete assessments every 3 months.

English and French-speaking SPIN Cohort participants were recruited from April 9 to April 27, 2020 via email and popups during SPIN Cohort online assessments to enroll in the SPIN COVID-19 Cohort. Potential participants were also invited through recruitment announcements on social media (e.g., SPIN's Facebook page and Twitter account) and patient organization advertisements in English and French in countries with large English and French-speaking populations, including Canada, the US, France, the UK, Australia, New Zealand, and the Philippines. SPIN COVID-19 Cohort participants completed measures using the Qualtrics online survey package.

Measures. Basic demographic and disease variables were self-reported at baseline, including age, gender, years of education, marital status, ethnicity, and current country. Loneliness was assessed via the ULS-6, a 6-item measure with responses

ranging from 1 (never) to 4 (often), with higher scores indicating greater loneliness (9). The SPIN researchers administered the English version of the ULS-6 and the French version of the ULS-6. The English ULS-6 was drawn from 6 items of the English R-ULS that aligned with Neto's selected ULS-6 items from the Portuguese R-ULS (9). The French version of those items from the R-ULS French translation was used for French-speaking participants (14).

Symptoms of depression were measured via the Patient Health Questionnaire 8 (PHQ-8), an 8-item measure evaluating depressive symptoms over the last 2 weeks. Responses range from 0 (not at all) to 3 (nearly every day), and total scores range from 0 to 24, with higher scores indicating greater depressive symptoms (15). The PHQ-8 is available in English and French and demonstrates an equivalent performance to the PHQ-9, which has been validated in patients with SSc (16,17). Social support was evaluated through the Oslo Social Support Scale 3 (OSSS-3), a 3-item self-report measure without a timeframe specification (18). The first response ranges from 1 to 4, and second and third responses range from 1 to 5; the total score ranges from 3 to 14, with higher scores indicating greater social support. The OSSS-3 has demonstrated sufficient internal reliability and structural validity, although it has not been validated in patients with SSc (18). The SPIN research team translated the OSSS-3 into French using the World Health Organization's well-accepted forward-backward translation method. Participants were also asked to self-report number of individuals currently living in their household and number of one-on-one and group interactions over the phone or through videoconferencing software in the past week.

Statistical analyses. Descriptive statistics, including means and SDs for each ULS-6 item and the total, were first calculated (SPSS software, version 27). Cohen's *d* standardized mean difference effect sizes between English and French ULS-6 total scores were compared with 95% CIs (19). The magnitude of effect size was interpreted as small ($0.20 \leq d < 0.50$), medium ($0.50 \leq d < 0.80$), and large ($d \geq 0.80$). CFA was used to evaluate the previously identified single-factor structure of the ULS-6, following the recommendations of Bentler (20). The following indicators of good model fit were used: 1) the chi-square test; 2) a comparative fit index (CFI) of >0.95; 3) a root mean square error of approximation (RMSEA) of <0.08; and 4) a standardized root mean residual (SRMR) of <0.08. The chi-square test was not used as a sole indicator of model fit, given its sensitivity to large sample sizes; therefore, the additional descriptive fit indices were employed, which do not depend on sample size (20).

The multiple-indicator multiple-cause (MIMIC) model was used to examine differential item functioning (DIF) for the English versus French versions of the ULS-6. The base MIMIC model is comprised of the CFA model and the direct effect of language group on the latent loneliness factor, which controls for group differences on the level of the latent factor (21). To assess for DIF, each item on the ULS-6 was regressed, one at a time, on language group. After items with statistically significant DIF were identified, MIMIC models that adjust and do not adjust for DIF were compared to evaluate the degree to which DIF may influence comparisons between groups.

Internal consistency reliability was calculated using Cronbach's alpha coefficient and McDonald's omega. Convergent validity was evaluated via Pearson's product-moment correlations of the ULS-6 with measures of depression (PHQ-8), social support (OSSS-3), number of people currently in the household, and frequency of social interactions. The magnitude of correlations was interpreted as small ($|r| \leq 0.30$), moderate ($0.30 < |r| < 0.50$), or large ($|r| \geq 0.50$) (22). Based on previous findings, for overall, English, and French samples, we predicted a large positive correlation between loneliness and depression and a large negative correlation between loneliness and social support (8,10,23). We anticipated a moderate negative correlation of loneliness to number of people currently in the household and frequency of social interactions (8,10,23). We predicted a small nonsignificant correlation with gender, given previous findings suggesting that loneliness levels do not depend on gender, and a moderate negative correlation with marital status, with married individuals scoring lower than nonmarried individuals (10,24). We also predicted a moderate positive correlation between age and ULS-6 scores (10,24). Correlation differences across language were calculated by transforming correlations to Fisher's Z values and using univariate generalized linear modeling. We predicted no correlation differences across language.

Regarding sample size for sufficiently powered analyses, a 1-factor CFA with 6 indicators would require a minimum sample size between 60 and 190 for factor loadings between 0.50 and 0.80 (25). For MIMIC models in the context of DIF, a total sample size of ≥600 allows for detection of even very small mediation effects and controlling the Type I error rate (21). A Pearson correlation of ≥0.30 with 95% confidence and a precision of 0.10 requires a sample size of ≥403 (25). There were no missing data for the CFA or MIMIC models. For the Pearson correlations, there was a range of 0 to 7 missing participant responses, accounting for ≤0.9% of the sample.

RESULTS

The initial sample had 800 participants, but 25 participants did not complete any ULS items and were therefore removed from analyses. Of the included 775 adults with SSc, 315 (42%, 16 missing) had diffuse SSc, 697 (90%, 4 missing) were women, and 512 completed measures in English (66%) (Table 1). For the total sample, the mean score on the ULS-6 was 7.00 (SD 4.76; range 0–18), with higher scores representing greater loneliness. English speakers (mean ± SD 7.29 ± 4.67) and French speakers

Table 1. Demographic characteristics of participants*

Characteristic	Overall (n = 775)	English (n = 512)	French (n = 263)
Age, mean ± SD years	55.6 ± 12.6 (n = 771)	56.4 ± 11.9 (n = 508)	54.0 ± 13.6 (n = 263)
Gender			
Women	697/771 (90.4)	461/508 (90.7)	236/263 (89.7)
Men	74/771 (9.6)	47/508 (9.3)	27/263 (10.3)
Marital status			
Not married	237/768 (30.9)	145/505 (28.7)	92/263 (35.0)
Married	531/768 (69.1)	360/505 (71.3)	171/263 (65.0)
Employment			
Not employed	449/769 (58.4)	300/507 (59.2)	149/262 (56.9)
Employed	320/769 (41.6)	207/507 (40.8)	113/262 (43.1)
Ethnicity			
White	638/765 (83.4)	426/506 (84.2)	212/259 (81.9)
Black	50/765 (6.5)	19/506 (3.8)	31/259 (12.0)
Other	77/765 (10.1)	61/506 (12.1)	16/259 (6.2)
Language			
English	512/775 (66.1)	506/506 (100)	0/263
French	263/775 (33.9)	0/506	263/263 (100)
Country			
US	244/773 (31.6)	244/510 (47.8)	0/263
Canada	192/773 (24.8)	129/510 (25.3)	63/263 (24.0)
France	198/773 (25.6)	4/510 (0.8)	194/263 (73.8)
UK	68/773 (8.8)	68/510 (13.3)	0/263
Australia	43/773 (5.6)	43/510 (8.4)	0/263
Other	28/773 (3.6)	22/510 (4.3)	6/263 (2.3)
Years since SSc diagnosis, mean ± SD years	11.6 ± 8.0 (n = 746)	12.1 ± 8.2 (n = 486)	10.7 ± 7.6 (n = 260)
Duration of education, mean ± SD years	15.8 ± 3.4 (n = 762)	15.9 ± 3.2 (n = 502)	15.6 ± 3.9 (n = 260)
SSc subtype			
Limited SSc	407/759 (52.5)	253/498 (50.8)	154/261 (59.0)
Diffuse SSc	315/759 (41.5)	219/498 (44.0)	96/261 (36.8)
Unknown per self-report	37/759 (4.9)	26/498 (5.2)	11/261 (4.2)

* Values are the no./total no. (%) unless indicated otherwise. SSc = systemic sclerosis.

(mean ± SD 6.45 ± 4.93) had a mean difference of 0.84 points (d = 0.18 [95% CI 0.03, 0.33]).

The CFA supported the expected single-factor structure (χ^2[9] = 85.56, $P < 0.001$; CFI 0.96, SRMR 0.03, RMSEA 0.11). All standardized factor loadings for items were large and statistically significant (0.60–0.86; all $P < 0.001$) (Table 2).

The overall MIMIC model fit well with language as a covariate (χ^2[14] = 147.36, $P < 0.001$; CFI 0.94, SRMR 0.04, RMSEA 0.11). Statistically significant DIF was found for 3 items across language, although standardized differences were small (β[item 2: "I feel part of a group of friends"] = 0.14, $P < 0.001$; β[item 4: "I feel isolated from others"] = –0.07, $P = 0.01$; β[item 6: "People are around me but not with me"] = 0.13, $P < 0.001$). The difference between English and French respondents in the latent factor score did not differ meaningfully when adjusting (SD –0.28 [95% CI –0.43, –0.12] or not adjusting for DIF (SD –0.29 [95% CI –0.46, –0.12]).

For the total sample, omega and alpha were both 0.87. For all study participants (Table 3), the ULS-6 total score correlated significantly and with expected directions and magnitudes with the total score for the PHQ-8 and the total score on the

Table 2. Item means and confirmatory factor analysis standardized factor loading results for the University of California, Los Angeles, Loneliness Scale-6 (ULS-6)*

Item†	Overall mean ± SD	English mean ± SD	French mean ± SD	Confirmatory factor analysis standardized factor loading
1. I lack companionship	1.04 ± 1.05	1.09 ± 1.05	0.92 ± 1.04	0.63
2. I feel part of a group of friends‡	0.84 ± 0.96	0.80 ± 0.90	0.91 ± 1.07	0.60
3. I feel left out	1.19 ± 1.02	1.29 ± 0.99	0.99 ± 1.05	0.82
4. I feel isolated from others	1.44 ± 1.07	1.58 ± 1.04	1.16 ± 1.08	0.86
5. I am unhappy being so withdrawn	1.15 ± 1.01	1.21 ± 0.99	1.05 ± 1.03	0.78
6. People are around me but not with me	1.35 ± 1.06	1.32 ± 1.04	1.41 ± 1.09	0.62
Total ULS-6 score mean	7.00 ± 4.78	7.29 ± 4.67	6.45 ± 4.93	NA

* NA = not applicable.
† On a 4-point scale, where 1 = never and 4 = often.
‡ Item 2 was reverse coded due to positive valence.

Table 3. Convergent validity of the University of California, Los Angeles, Loneliness Scale-6, per language through Pearson correlations*

Variable	Hypothesized range (direction)	Overall correlation (95% CI)	English correlation (95% CI)	French correlation (95% CI)	Hypothesis supported?	Correlation difference (95% CI)
PHQ-8 (n = 775)	Large + (|r| ≥ 0.50)	0.56 (0.58, 0.61)†	0.58 (0.52, 0.63)†	0.52 (0.43, 0.60)†	Yes	0.06 (–0.06, 0.15)
OSSS-3 (n = 775)	Large – (|r| ≥ 0.50)	–0.53 (–0.58, –0.47)†	–0.51 (–0.58, –0.45)†	–0.56 (–0.64, –0.47)†	Yes	0.05 (–0.08, 0.17)
No. of people in household (n = 774)	Moderate – (0.30 < |r| < 0.50)	–0.06 (–0.13, 0.01)	–0.04 (–0.13, 0.04)	–0.12 (–0.24, 0.001)	No; small and nonsignificant instead of moderate	0.08 (–0.07, 0.23)
No. of social interactions with 1 person (n = 769)	Moderate – (0.30 < |r| < 0.50)	–0.16 (–0.23, –0.09)†	–0.18 (0.26, –0.09)†	–0.13 (–0.25, –0.01)‡	No; small instead of moderate	0.05 (–0.19, 0.10)
No. of social interactions with multiple people (n = 769)	Moderate – (0.30 < |r| < 0.50)	–0.24 (–0.31, –0.17)†	–0.24 (–0.32, –0.15)†	–0.25 (–0.36, –0.13)†	No; small instead of moderate	0.01 (–0.14, 0.16)
Gender (n = 771)	Small, no direction predicted (|r| ≤ 0.30)	–0.06 (–0.13, 0.02)	–0.02 (–0.11, 0.06)	–0.11 (–0.23, 0.01)	Yes	0.09 (–0.06, 0.24)
Relationship status (n = 768)	Moderate – (0.30 < |r| < 0.50)	–0.17 (–0.24, –0.11)†	–0.21 (–0.29, –0.13)†	–0.13 (–0.25, –0.01)‡	No; small instead of moderate	0.08 (–0.23, 0.06)
Age (n = 771)	Moderate + (0.30 < |r| < 0.50)	–0.06 (–0.15, –0.01)‡	–0.12 (–0.20, –0.03)†	–0.03 (–0.15, 0.09)	No; small and negative instead of moderate and positive	0.09 (–0.24, 0.06)

* Values are the r value unless indicated otherwise. OSSS-3 = Oslo Social Support Scale 3; PHQ-8 = Patient Health Questionnaire 8.
† $P < 0.001$.
‡ $P < 0.05$.

OSSS-3. As expected, the ULS-6 total score had a small nonsignificant correlation with gender. Correlations with the frequency of social interactions with 1 person and multiple people were significant and negative, as expected, but small. The correlation with age was significant but small and negative, and the correlation with marital status was significant and small, with nonmarried individuals indicating higher loneliness than married individuals. Unexpectedly, the ULS-6 had a nonsignificant small negative correlation with number of people in the household. When testing the differences between correlations between English and French (Table 3), there were no significant language differences in the correlations between the ULS-6 and all hypothesized variables.

DISCUSSION

Analyses provided evidence for acceptable reliability and validity of the ULS-6 scores in English- and French-speaking adults with SSc. The CFA indicated the appropriateness of the single-factor structure, supporting use of a total score. Internal consistency calculations indicated high reliability. Although the overall MIMIC model fit well, MIMIC analyses also showed that 3 of the 6 items showed statistically significant DIF across linguistic groups. Despite these findings, differences between groups were not affected by adjusting or not adjusting for DIF. This evidence suggests that loneliness scores can be compared across languages.

Convergent validity findings did not significantly differ between the 2 groups, as expected. For both English and French speakers, the total score on the ULS-6 correlated significantly and in expected directions with total scores for depression and social support. Further, the ULS-6 was not significantly correlated with gender, as expected. For both languages, the ULS-6 had small correlations with frequencies of virtual social interactions and with nonmarried status (versus married status). Surprisingly, the ULS-6 had a small negative correlation with age. Previous literature using the ULS indicates that older age is consistently significantly associated with higher levels of loneliness in older adults and in patients with other chronic illnesses such as cancer (10,26,27). Our study finding, which differed from previous literature, could be attributed to a variety of justifications, including different patterns in patients with SSc than in older adults in the general population or patients with other chronic illnesses during the COVID-19 pandemic.

Additionally, the ULS-6 was not significantly inversely related to the number of people in the household (9). It is possible that COVID-19-related factors, such as needing to quarantine while ill with COVID, complicate this relationship and findings that would be expected to be significant. It is also possible that loneliness was more strongly rooted in the meaningfulness of interactions rather than the quantity of interactions. This interpretation aligns with the initially stated definition of loneliness as a pervasive and distressing experience involving a person's perception that their social relationships do not fulfill their social needs (1). Both English- and French-speaking patients with SSc might require more emotionally significant social interactions to reduce feelings of loneliness.

Loneliness as a latent construct has become especially relevant during the COVID-19 pandemic and may disproportionately impact chronically ill groups, especially those experiencing rare chronic illnesses such as SSc. A recent meta-analysis demonstrated that loneliness has increased since the start of the pandemic (4). Research has also demonstrated that sustained loneliness can have serious implications for mental and physical health outcomes (1,28). It is important to note, however, that in a recent SPIN study, depression levels in patients with SSc did not change from before the COVID-19 pandemic to during the COVID-19 pandemic; given our study findings that depression had a large correlation with loneliness, it is possible that loneliness levels may not have worsened for patients with SSc during the COVID-19 pandemic (29). SPIN researchers are in the process of analyzing findings regarding changes in loneliness levels during the COVID-19 pandemic, which will provide valuable information regarding the nature of loneliness in patients in the context of the COVID-19 pandemic, and whether their experiences compare to those of individuals in the general population (4).

Notably, the SPIN research team recently developed and tested an intervention, via a randomized controlled trial, targeting psychosocial outcomes including loneliness in patients with SSc during the COVID-19 pandemic (30). The COVID-19 Home-Isolation Activities Together (SPIN-CHAT) program was a 4-week telehealth group intervention providing education and mental health coping strategies, as well as social support, to reduce patient anxiety, depression, and loneliness. While developing this intervention, the SPIN patient advisory board emphasized the importance of prioritizing anxiety but believed that depression and loneliness should be less of a priority (31). They attributed this to the fact that patients with SSc already managed feelings of isolation before the COVID-19 pandemic and therefore demonstrated resiliency against depression and loneliness. Loneliness was still incorporated as an intervention target, given that patients with SSc are at increased risk of serious complications from COVID-19 and had been advised to self-isolate (31). While the intervention had small effects on anxiety, there were no intervention effects on loneliness (30). Beyond the SPIN-CHAT program, only 1 other study has specifically targeted loneliness since the start of the COVID-19 pandemic. The study's intervention was conducted for older adult clients of a Meals on Wheels program and involved 3–5 conversational phone calls per week for 4 weeks (32). The intervention successfully decreased loneliness levels on the R-ULS. Future studies should evaluate and continue to target the comparative experiences of loneliness in SSc and other chronically ill groups given the limited literature on this psychological construct.

This study had several strengths, including its large sample size, diverse group of participants, and rigorous psychometric

methods. However, the study also had notable limitations. The study sample was a convenience sample that had opted in to participating, posing a risk of selection bias. The context of the COVID-19 pandemic also created a unique environment for studying loneliness that may not be easily extrapolated to other circumstances. Further, this study did not investigate the discriminant validity of the ULS-6 or the extent to which the ULS-6 selectively captures loneliness, with the exclusion of associated yet distinct constructs such as depression. Additionally, given that the study was not designed to explicitly study loneliness, it did not incorporate specific variables of interest that would have further established convergent validity, such as strength of relationships. Additional collection of evidence is warranted to demonstrate further validity of the scale and to substantiate proposed theories for why certain findings may have differed from expectations.

In conclusion, the present study findings offer evidence of reliability and validity of the ULS-6 for use with and across English- and French-speaking patients with SSc, as demonstrated by CFA, MIMIC, and Pearson correlation findings. The limited literature on loneliness in patients with autoimmune rheumatic diseases shows that they are at higher risk of self-isolation generally and during the COVID-19 pandemic, demonstrating a need for further research (33). The ULS-6 can be used as a helpful tool in future studies evaluating and targeting loneliness through interventions for patients with SSc.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Rapoport, Choi, Mouthon, Malcarne.
Acquisition of data. Kwakkenbos, Carrier, Henry, Thombs.
Analysis and interpretation of data. Rapoport, Roesch, Thombs, Malcarne.

REFERENCES

1. Hawkley LC, Cacioppo JT. Loneliness matters: a theoretical and empirical review of consequences and mechanisms. Ann Behav Med 2010;40:218–27.
2. Petitte T, Mallow J, Barnes E, et al. A systematic review of loneliness and common chronic physical conditions in adults. Open Psychol J 2015;8 Suppl 2:113–32.
3. Mancuso CA, Duculan R, Jannat-Khah D, et al. Rheumatic disease-related symptoms during the height of the COVID-19 pandemic. HSS J 2020;16 Suppl 1:36–44.
4. Ernst M, Niederer D, Werner AM, et al. Loneliness before and during the COVID-19 pandemic: a systematic review with meta-analysis. Am Psychol 2022;77:660–77.
5. Allanore Y, Simms R, Distler O, et al. Systemic sclerosis. Nat Rev Dis Primers 2015;1:15002.
6. Atzeni F, Gerardi MC, Barilaro G, et al. Interstitial lung disease in systemic autoimmune rheumatic diseases: a comprehensive review. Expert Rev Clin Immunol 2018;14:69–82.
7. Hoffmann-Vold AM, Distler O, Bruni C, et al. Systemic sclerosis in the time of COVID-19. Lancet Rheumatol 2022;4:e566–75.
8. Russell D, Peplau LA, Cutrona CE. The revised UCLA Loneliness Scale: concurrent and discriminant validity evidence. J Pers Soc Psychol 1980;39:472–80.
9. Neto F. Loneliness among Portuguese adolescents. Soc Behav Pers 1992;20:15–21.
10. Neto F. Psychometric analysis of the short-form UCLA Loneliness Scale (ULS-6) in older adults. Eur J Ageing 2014;11:313–9.
11. Gagnier JJ, Lai J, Mokkink LB, et al. COSMIN reporting guideline for studies on measurement properties of patient-reported outcome measures. Qual Life Res 2021;30:2197–218.
12. Kwakkenbos L, Jewett LR, Baron M, et al. The Scleroderma Patient-Centered Intervention Network (SPIN) Cohort: protocol for a cohort multiple randomised controlled trial (cmRCT) design to support trials of psychosocial and rehabilitation interventions in a rare disease context. BMJ Open 2013;3:e003563.
13. Van Den Hoogen F, Khanna D, Fransen J, et al. 2013 classification criteria for systemic sclerosis: an American College of Rheumatology/-European League against Rheumatism collaborative initiative. Arthritis Rheum 2013;65:2737–47.
14. De Grace GR, Joshi P, Pelletier R. L'Échelle de solitude de l'Université Laval (ÉSUL): validation canadienne-française du UCLA Loneliness Scale. Can J Behav Sci 1993;25:12–27.
15. Kroenke K, Strine TW, Spitzer RL, et al. The PHQ-8 as a measure of current depression in the general population. J Affect Disord 2009;114:163–73.
16. Arthurs E, Steele RJ, Hudson M, et al. Are scores on English and French versions of the PHQ-9 comparable? An assessment of differential item functioning. PLoS One 2012;7:e52028.
17. Wu Y, Levis B, Riehm KE, et al. Equivalency of the diagnostic accuracy of the PHQ-8 and PHQ-9: a systematic review and individual participant data meta-analysis. Psychol Med 2020;50:1368–80.
18. Kocalevent RD, Berg L, Beutel ME, et al. Social support in the general population: standardization of the Oslo Social Support Scale (OSSS-3). BMC Psychol 2018;6:31.
19. Andrade C. Mean difference, standardized mean difference (SMD), and their use in meta-analysis: as simple as it gets. J Clin Psychiatry 2020;81:20f13681.
20. Bentler PM. Comparative fit indexes in structural models. Psychol Bull 1990;107:238–46.
21. Cheng Y, Shao C, Lathrop QN. The mediated MIMIC model for understanding the underlying mechanism of DIF. Educ Psychol Meas 2016;76:43–63.
22. Cohen J. Statistical power analysis for the behavioral sciences. Hoboken (NJ): Taylor and Francis; 2013.
23. Lee Y, Ko Y. Feeling lonely when not socially isolated: social isolation moderates the association between loneliness and daily social interaction. J Soc Pers Relat 2018;35:1340–55.
24. Maes M, Qualter P, Vanhalst J, et al. Gender differences in loneliness across the lifespan: a meta-analysis. Eur J Pers 2019;33:642–54.
25. Wolf EJ, Harrington KM, Clark SL, et al. Sample size requirements for structural equation models: an evaluation of power, bias, and solution propriety. Educ Psychol Meas 2013;73:913–34.
26. Cohen-Mansfield J, Hazan H, Lerman Y, et al. Correlates and predictors of loneliness in older-adults: a review of quantitative results informed by qualitative insights. Int Psychogeriatr 2016;28:557–76.
27. Deckx L, van den Akker M, Buntinx F. Risk factors for loneliness in patients with cancer: a systematic literature review and meta-analysis. Eur J Oncol Nurs 2014;18:466–77.

28. Jeste DV, Lee EE, Cacioppo S. Battling the modern behavioral epidemic of loneliness: suggestions for research and interventions. JAMA Psychiatry 2020;77:553–4.
29. Henry RS, Kwakkenbos L, Carrier ME, et al. Mental health before and during the pandemic in people with systemic sclerosis. Lancet Rheumatol 2022;4:e82–5.
30. Thombs BD, Kwakkenbos L, Levis B, et al. Effects of a multi-faceted education and support programme on anxiety symptoms among people with systemic sclerosis and anxiety during COVID-19 (SPINCHAT): a two-arm parallel, partially nested, randomised, controlled trial. Lancet Rheumatol 2021;3:e427–37.
31. Thombs BD, Kwakkenbos L, Carrier ME, et al. Protocol for a partially nested randomised controlled trial to evaluate the effectiveness of the Scleroderma Patient-Centered Intervention Network COVID-19 Home-Isolation Activities Together (SPIN-CHAT) program to reduce anxiety among at-risk scleroderma patients. J Psychosom Res 2020;135:110132.
32. Kahlon MK, Aksan N, Aubrey R, et al. Effect of layperson-delivered, empathy-focused program of telephone calls on loneliness, depression, and anxiety among adults during the COVID-19 pandemic: a randomized clinical trial. JAMA Psychiatry 2021;78:616–22.
33. Conway R, Grimshaw AA, Konig MF, et al. SARS–CoV-2 infection and COVID-19 outcomes in rheumatic diseases: a systematic literature review and meta-analysis. Arthritis Rheumatol 2022;74:766–75.

APPENDIX A: THE SPIN COVID-19 PATIENT ADVISORY TEAM AND SPIN INVESTIGATORS

In addition to the authors, the following primary investigators participated in the SPIN COVID-19 Patient Advisory Team: Catherine Fortuné (Ottawa Scleroderma Support Group, Ottawa, Ontario, Canada); Amy Gietzen (National Scleroderma Foundation, Tri-State Chapter, Binghamton, New York); Geneviève Guillot (Sclérodermie Québec, Longueuil, Quebec, Canada); Nancy Lewis (Toronto, Ontario, Canada); Karen Nielsen, Maureen Sauvé (Scleroderma Society of Ontario, Hamilton, Ontario, Canada); Michelle Richard (Scleroderma Atlantic, Halifax, Nova Scotia, Canada); Joep Welling (NVLE Dutch Patient Organization for Systemic Autoimmune Diseases, Utrecht, The Netherlands); and John Varga (University of Michigan, Ann Arbor).

In addition to the authors, SPIN investigators include the following: Claire E. Adams, Mara Cañedo Ayala, Vanessa Cook, Sophie Hu, Elsa-Lynn Nassar, Marieke Alexandra Neyer, Julia Nordlund, Sabrina Provencher (Jewish General Hospital, Montreal, Quebec, Canada); Susan J. Bartlett, Marie Hudson, Andrea Benedetti (McGill University, Montreal, Quebec, Canada); Karen Gottesman (National Scleroderma Foundation, Los Angeles, California); Laura K. Hummers (Johns Hopkins University School of Medicine, Baltimore, Maryland); Amanda Lawrie-Jones (Scleroderma Australia and Scleroderma Victoria, Melbourne, Victoria, Australia); Maureen D. Mayes, Shervin Assassi (University of Texas McGovern School of Medicine, Houston); Warren R. Nielson (St. Joseph's Health Care, London, Ontario, Canada); Ghassan El-Baalbaki (Université du Québec à Montréal, Montreal, Quebec, Canada); Cornelia van den Ende (Sint Maartenskliniek, Nijmegen, The Netherlands); Kim Fligelstone (Scleroderma & Raynaud's UK, London, UK); Tracy Frech (University of Utah, Salt Lake City); Daphna Harel (New York University, New York, New York); Monique Hinchcliff (Yale School of Medicine, New Haven, Connecticut); Sindhu R. Johnson (Toronto Scleroderma Program, Mount Sinai Hospital, Toronto Western Hospital and University of Toronto, Toronto, Ontario, Canada); Maggie Larche, Nader Khalidi (McMaster University, Hamilton, Ontario, Canada); Catarina Leite (University of Minho, Braga, Portugal); Christelle Nguyen, François Rannou (Université Paris Descartes and AP-HP, Paris, France); Janet Pope (University of Western Ontario, London, Ontario, Canada); Tatiana Sofia Rodriguez Reyna (Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico City, Mexico); Anne A. Schouffoer (Leiden University Medical Center, Leiden, The Netherlands); Maria E. Suarez-Almazor (University of Texas MD Anderson Cancer Center, Houston); Christian Agard (Centre Hospitalier Universitaire - Hôtel-Dieu de Nantes, Nantes, France); Nassim Ait Abdallah, Benjamin Crichi, Dominique Farge-Bancel (Hôpital St. Louis, AP-HP, Paris, France); Marc André, Louis Olagne, Perrine Smets (Centre Hospitalier Universitaire Gabriel-Montpied, Clermont-Ferrand, France); Elana J. Bernstein (Columbia University, New York, New York); Sabine Berthier (Centre Hospitalier Universitaire Dijon Bourgogne, Dijon, France); Lyne Bissonnette, Alessandra Bruns, Ariel Masetto, Sophie Roux (Université de Sherbrooke, Sherbrooke, Quebec, Canada); Patricia Carreira, Regina Fare, Maria Martin, Sheila Melchor Díaz, Esther Rodríguez Almazar (Servicio de Reumatologia del Hospital 12 de Octubre, Madrid, Spain); Marion Casadevall, Benjamin Chaigne, Bertrand Dunogue, Alexis Régent (Hôpital Cochin, AP-HP, Paris, France); Lorinda Chung (Stanford University, Stanford, California); Christopher Denton (Royal Free London Hospital, London, UK); Robyn Domsic (University of Pittsburgh, Pittsburgh, Pennsylvania); James V. Dunne, Pearce Wilcox (St. Paul's Hospital and University of British Columbia, Vancouver, British Columbia, Canada); Paul R. Fortin, Alena Ikic (CHU de Québec - Université Laval, Quebec, Quebec, Canada); Jessica Gordon, Kimberly Lakin, Robert Spiera (Hospital for Special Surgery, New York, New York); Brigitte Granel-Rey (Aix Marseille Université, and Assistance Publique - Hôpitaux de Marseille, Hôpital Nord, Marseille, France); Aurélien Guffroy, Thierry Martin, Vincent Poindron (Les Hôpitaux Universitaires de Strasbourg, Nouvel Hôpital Civil, Strasbourg, France); Genevieve Gyger (Jewish General Hospital and McGill University, Montreal, Quebec, Canada); Eric Hachulla, Marc Lambert, David Launay, Hélène Maillard, Vincent Sobanski (Centre Hospitalier Régional Universitaire de Lille, Hôpital Claude Huriez, Lille, France); Sabrina Hoa (Centre hospitalier de l'université de Montréal – CHUM, Montreal, Quebec, Canada); Niall Jones (University of Alberta, Edmonton, Alberta, Canada); Suzanne Kafaja (University of California, Los Angeles); Yvonne C. Lee (Northwestern University, Chicago, Illinois); Nancy Maltez (University of Ottawa, Ottawa, Ontario, Canada); Joanne Manning (Salford Royal NHS Foundation Trust, Salford, UK); Isabelle Marie (CHU Rouen, Hôpital de Bois-Guillaume, Rouen, France); François Maurier (Uneos-Groupe hospitalier associatif, Metz, France); Arsene Mekinian, Sébastien Rivière (Hôpital St. Antoine, AP-HP, Paris, France); Mandana Nikpour (St. Vincent's Hospital and University of Melbourne, Melbourne, Victoria, Australia); Susanna Proudman (Royal Adelaide Hospital and University of Adelaide, Adelaide, South Australia, Australia); David Robinson (University of Manitoba, Winnipeg, Manitoba, Canada); Virginia Steen (Georgetown University, Washington, DC); Evelyn Sutton (Dalhousie University, Halifax, Nova Scotia, Canada); Carter Thorne (Southlake Regional Health Centre, Newmarket, Ontario, Canada); and John Varga (University of Michigan, Ann Arbor).

Arthritis Care & Research
Vol. 75, No. 10, October 2023, pp 2166–2173
DOI 10.1002/acr.25118

Antivinculin Antibodies in Systemic Sclerosis: Associations With Slow Gastric Transit and Extraintestinal Clinical Phenotype

María Herrán,[1] Brittany L. Adler,[2] Jamie Perin,[3] Walter Morales,[4] Mark Pimentel,[4] and Zsuzsanna H. McMahan[2]

Objective. The gastrointestinal (GI) tract is commonly affected in systemic sclerosis (SSc). A positive association between antivinculin antibody levels and GI symptom severity is reported in SSc. We sought to examine whether antivinculin antibodies associate with measures of GI dysmotility and extraintestinal clinical phenotype in SSc.

Methods. A total of 88 well-characterized patients with SSc and GI disease were assayed for antivinculin antibodies by enzyme-linked immunosorbent assay. Whole-gut scintigraphy, GI symptom scores, and clinical features of SSc were compared between patients with and without antibodies.

Results. Twenty of 88 (23%) patients had antivinculin antibodies, which were more prevalent in patients with slow gastric transit (35% versus 22%). In the univariate analyses, patients who were positive for antivinculin antibodies were more likely to have limited cutaneous disease (odds ratio [OR] 9.60 [95% confidence interval (95% CI) 1.19, 77.23]) and thyroid disease (OR 4.09 [95% CI 1.27, 13.21]). Such patients were also less likely to have lung involvement based on a Medsger Severity Score of ≥2 (OR 0.25 [95% CI 0.07, 0.92]). Higher levels of antivinculin autoantibodies were associated with less gastric emptying (β coefficient –3.41 [95% CI –6.72, –0.09]). The association between antivinculin antibodies and each of these clinical features remained significant in the multivariable model. In particular, the presence of antivinculin antibodies (β coefficient –6.20 [95% CI –12.33, –0.063]) and higher levels of antivinculin antibodies (β coefficient –3.64 [95% CI –7.05, –0.23]) were each significantly associated with slower gastric transit.

Conclusion. Antivinculin antibodies associate with slower gastric transit in SSc and may provide insight into GI complications of SSc.

INTRODUCTION

Systemic sclerosis (SSc) is a rare immune-mediated connective tissue disease characterized by progressive vasculopathy, autoimmunity, and fibrosis. Over time, it can lead to multiorgan dysfunction with a highly heterogeneous clinical presentation (1). The gastrointestinal (GI) tract is the most commonly affected internal organ system in SSc (2); any segment of the GI tract can be involved, and patients may present with a variety of GI symptoms. Furthermore, patients can develop severe GI disease at any time during their disease course regardless of disease duration. It therefore remains a major challenge to identify patients at high risk for progressive GI disease and predict which areas of the GI tract are likely to be affected (3).

Serum antibodies in SSc, such as anticentromere antibody, anti–topoisomerase I antibody, and anti–RNA polymerase III antibody, are commonly detected early in the disease course and are associated with different patterns of skin and internal organ involvement. As a result, they serve as useful diagnostic and prognostic biomarkers in SSc (4,5). Previous studies have identified functional autoantibodies targeting the antimuscarinic type 3 receptor antibody, which contribute to severe lower GI dysmotility in SSc, although screening for these autoantibodies is not yet clinically available (6). Autoantibodies to ganglionic neuronal nicotinic acetylcholine receptor autoantibodies are also present in a small subset of patients and may have functional implications (7). Nevertheless, the pathophysiology of GI disease in SSc remains

Supported by the NIH (National Institute of Arthritis and Musculoskeletal and Skin Diseases grant K23-AR-071473 to Dr. McMahan), the Rheumatology Research Foundation, the Scleroderma Research Foundation, and the Jerome L. Greene Foundation.

[1]María Herrán, MD: Universidad del Rosario, Bogota, Colombia; [2]Brittany L. Adler, MD, Zsuzsanna H. McMahan, MD, MHS: Johns Hopkins University, Baltimore, Maryland; [3]Jamie Perin, PhD, MS: Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland; [4]Walter Morales, BSc, Mark Pimentel, MD, FRCP (C): Cedars-Sinai Medical Center, Los Angeles, California.

Drs. Herrán and Adler contributed equally to this work.

Author disclosures and a graphical abstract are available online at https://onlinelibrary.wiley.com/doi/10.1002/acr.25118.

Address correspondence via email to Zsuzsanna H. McMahan, MD, MHS, at zmcmaha1@jhmi.edu.

Submitted for publication July 29, 2022; accepted in revised form March 21, 2023.

SIGNIFICANCE & INNOVATIONS

- Antivinculin antibody positivity and antivinculin antibody levels are associated with slow gastric transit in systemic sclerosis.
- Extraintestinal features associated with antivinculin antibodies include limited scleroderma, thyroid disease, a higher right ventricular systolic pressure, and less severe lung disease.
- Slower gastric transit remains significantly associated with antivinculin antibody positivity and higher antivinculin antibody levels even after adjusting for age and sex.

unclear in many SSc GI subgroups, and there are likely other biomarkers of GI disease that have yet to be discovered.

The protein vinculin has been previously identified as an autoantigen in patients with different GI diseases (8,9), including irritable bowel syndrome (IBS) (10). Vinculin is a cytoskeletal protein that binds to actin and promotes cell adhesion, gut motility, and angiogenesis (11,12). Many of the functional GI diseases are temporally associated with infectious gastroenteritis, and a model has been proposed that could explain this observation: cytolethal distending toxin (CdtB) produced by bacteria during infectious gastroenteritis may result in the development of anti-CdtB antibodies, which then cross-react with vinculin protein found in the gut and lead to an immune response against vinculin (13,14). Levels of anti-CdtB and antivinculin antibodies can distinguish patients with diarrhea-predominant IBS from inflammatory bowel disease (10). More recently, a study in SSc demonstrated an association between antivinculin autoantibodies and GI symptom severity measured by the GI visual analog scale (VAS) (15). A limitation of this study was that self-reported GI symptom scores were used to measure GI severity. Furthermore, these surveys are not specific to SSc and are unable to identify specific regions of dysfunction across the GI tract. In the present study, we therefore sought to determine whether the presence and titer of antivinculin antibodies in SSc associate with 1) abnormal GI transit measured by whole-gut transit (WGT) scintigraphy; and 2) GI symptom severity measured by the validated University of California, Los Angeles, Scleroderma Clinical Trials Consortium Gastrointestinal Tract 2.0 (UCLA GIT 2.0) instrument (16).

PATIENTS AND METHODS

Patients. All participants met either the 2013 American College of Rheumatology (ACR)/EULAR classification criteria for SSc (17), the 1980 ACR criteria (18), or had at least 3 of 5 features of CREST syndrome (calcinosis, Raynaud's phenomenon, esophageal dysmotility, sclerodactyly, telangiectasias) (19) and were recruited from the Johns Hopkins Scleroderma Center during routine clinical visits. Participants with GI symptoms (including early satiety, nausea, vomiting, unintentional weight loss, distension, bloating, diarrhea, and/or constipation as determined by the treating physician) underwent the WGT study. Patients who were minimally symptomatic or asymptomatic were also recruited to capture the full spectrum of GI disease. All participants signed written, informed consent. Ethical approval was provided by the Johns Hopkins Institutional Review Board.

Clinical phenotyping of SSc. Demographic characteristics included age, sex, and race. Race was determined with a self-questionnaire that included the following categories based on the standard data collection in our Scleroderma Center database: American Indian/Alaska Native, Asian, Black or African American, Indian subcontinent, Mid-East/Arabian, Native Hawaiian/Pacific Islander, White, Other/Unknown, Not Applicable. Other clinical data such as disease duration, smoking status, SSc subtype, presence (yes/no) of telangiectasia, calcinosis, arthralgia, synovitis, tendon friction rubs, thyroid disease (any), diabetes mellitus, gastric antral vascular ectasia (GAVE), and cancer history were documented by the physician at the patient's first clinical encounter and at 6-month intervals during follow-up visits. Disease duration was defined as the interval of time between the first non-Raynaud's symptom and the WGT study. SSc subtype (diffuse cutaneous SSc or limited cutaneous SSc) was defined based on the extent of skin involvement. GI, cardiac, muscle, and lung involvement and severity were also captured at baseline and longitudinally using the Medsger Severity Score (20). The maximum modified Medsger GI Severity Score was used to characterize SSc GI severity. Five categories were evaluated including: 1) score 0 = normal (no GI symptoms); 2) score 1 = requiring gastroesophageal reflux disease (GERD) medications (including H_2 blocker or proton-pump inhibitor); 3) score 2 = high-dose GERD medications or antibiotics for bacterial overgrowth; 4) score 3 = episodes of pseudo-obstruction or malabsorption syndrome; and 5) score 4 = severe GI dysmotility requiring enteral or total parenteral nutrition. Severe GI involvement was defined as a Medsger Severity Score of ≥2. The presence of sicca symptoms was defined as the presence of any of the following: dry eyes for >3 months; the sensation of sand or gravel in the eyes; the use of artificial tears 3 times daily; dry mouth for >3 months; swollen salivary glands; and/or the requirement of liquids to swallow due to dry mouth. Other clinical variables are defined in Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25118.

Other instruments. *UCLA GIT 2.0 survey.* The UCLA GIT 2.0 questionnaire is a validated patient-reported outcome measure designed to quantify GI symptoms in SSc. The questionnaire contains 34 items organized into 7 subscales (reflux, distension/bloating, soilage, diarrhea, social functioning, emotional well-being, and constipation), and a total score is

calculated (21). The maximum UCLA GIT scores were utilized in the analysis to capture phenotype.

WGT scintigraphy. All participants underwent WGT scintigraphy. Patients fasted the night before the study and were instructed to avoid promotility agents, antibiotics, opiates, benzodiazepines, and stool softeners 3 days prior to the study. Per protocol, patients consumed a standard amount of radiolabeled In-111 water to assess esophageal and liquid gastric emptying and a radiolabeled Tc-99m egg meal to evaluate solid gastric emptying. Subsequently, a gamma camera was used to capture anterior and posterior images of the GI tract at predetermined periods of time (one-half hour, 1 hour, 2 hours, 4 hours, 6 hours, 24 hours, 48 hours, and 72 hours) to evaluate transit throughout the entire GI tract. Slow gastric transit was defined as delayed solid gastric emptying at 4 hours. Slow colonic transit was defined as <14% emptying at 24 hours, <41% emptying at 48 hours, and <67% emptying at 72 hours (22).

Antibody profiles. A second-generation enzyme-linked immunosorbent assay (ELISA) was used to measure antivinculin antibodies in the sera of patients using a validated second-generation assay established at Cedars-Sinai Medical Center (23). After undergoing epitope optimization, the antigen was mobilized onto high-binding plates and blocked with 3% bovine serum albumin in phosphate buffered saline to counter nonspecific binding. Antivinculin antibody levels were evaluated after 70 minutes using optical densities obtained after measuring the absorbance at 370 nm. The antivinculin antibody assay was considered positive when the optical density was ≥1.68, per previously published studies (10,15).

Patient sera were also screened for other autoantibodies using the commercially available Euroline Scleroderma Nucleoli Profile Uroline (IgG) immunoblot assay (Euroimmun). This assay provides the ability to screen for autoantibodies to Scl-70, centromere (CENP-A or CENP-B), RNA polymerase III (RPC11 or RPC155), fibrillarin (U3 RNP), Th/To, Ku, PM/Scl-75, and PM/Scl-100. Medium-to-strong bands (>11) on signal intensity were considered positive and were determined based on the manufacturer's cutoffs. Positivity to either subunit of centromere, RNA polymerase III, and PM/Scl were considered positive.

Statistical methods. We examined associations between dichotomous variables using chi-square or Fisher's exact tests where appropriate. Student's *t*-test was used to evaluate differences in the mean values of 2 continuous variables, and the Wilcoxon-Mann-Whitney test was used to compare the medians of highly skewed continuous variables. Spearman's rank correlation coefficient was used to assess the linear correlation between 2 continuous variables. Univariable linear and logistic regression analyses were used to examine the strength of bivariate associations. Multivariable regression analyses were constructed to determine whether associations remained after adjusting for potential confounders. All statistical analyses were conducted using Stata, version 14. Statistical tests were 2-sided, and statistical significance was defined as *P* values less than 0.05 for all analyses.

RESULTS

Antivinculin antibodies and SSc clinical characteristics. Based on the ideal cutoff levels, 20 of 88 patients (23%) with SSc had levels over the threshold of optical density ≥1.68. The detailed clinical and demographic features of SSc patients who were positive and negative for antivinculin antibodies are shown in Table 1. There was no difference in age, sex, race, or disease duration between patients positive and patients negative for antivinculin antibodies. Other clinical features of SSc, including modified Rodnan skin thickness score, cardiac involvement, myopathy, cancer, severe GI involvement, GAVE, sicca symptoms, calcinosis, tendon friction rubs, and telangiectasias, were not significantly different between the patients positive and patients negative for antivinculin antibodies.

Patients who were positive for antivinculin antibodies were significantly more likely to have limited cutaneous disease (94.1% versus 62.5%; $P = 0.02$) and thyroid disease (47.1% versus 17.9%; $P = 0.01$). Among the antivinculin antibody–positive patients with thyroid disease, all of them had hypothyroidism. Antivinculin antibody–positive patients were also less likely to have pulmonary involvement as defined by a Medsger Severity Score of ≥2 (26.7% versus 59.1%; $P = 0.04$) and had a higher estimated right ventricular systolic pressure (RVSP) by echocardiogram (37.0 mm Hg [interquartile range (IQR) 35.0–42.0] versus 31.0 mm Hg [IQR 25.0–35.0], $P = 0.03$) compared to antivinculin antibody–negative patients. Other cardiopulmonary parameters, including the mean diffusing capacity for carbon monoxide percent predicted and the mean forced vital capacity percent predicted on pulmonary function testing, were similar between the 2 groups. Of the antivinculin-positive SSc patients, 60.0% (9 of 15) were also positive for anticentromere antibodies, 6.7% (1 of 15) were also positive for anti–topoisomerase I (Scl-70), and 6.7% (1 of 15) were positive for anti–RNA polymerase III.

Univariate models. Univariate linear and logistic regression analyses were performed to examine the strength of the association between SSc-specific clinical features and positivity for antivinculin antibodies (Table 2). Antivinculin antibody–positive patients were 9.6 times as likely to have limited cutaneous disease as those who were negative for antivinculin antibodies (95% confidence interval [95% CI] 1.19, 77.23; $P = 0.03$), and they had more than triple the risk of thyroid disease (odds ratio [OR] 4.1 [95% CI 1.27, 13.21], $P = 0.02$). Antivinculin antibody–positive patients also had 75% lower risk of lung disease (OR 0.25 [95% CI 0.07, 0.92], $P = 0.04$) and had a higher estimated RVSP (OR 1.22 [95% CI 1.00, 1.49], $P = 0.05$).

Table 1. Evaluation of associations between systemic sclerosis (SSc) clinical characteristics and antivinculin antibodies*

Variable	Antivinculin antibodies		*P*
	Positive	Negative	
Age at first symptom, mean ± SD years	62 ± 10.7	57 ± 12.3	0.084
Female	15/17 (88.2)	50/56 (89.3)	1.000
Race			
White	16/17 (94.1)	40/55 (72.7)	0.095
Disease duration from first non-Raynaud's symptom to date of WGT, median (IQR)	14.3 (3.7–15.3)	7.8 (4.9–16.4)	0.520
Ever smoker	5/17 (29.4)	20/56 (35.7)	0.631
SSc type			
Limited cutaneous disease	16/17 (94.1)	35/56 (62.5)	0.016†
Maximum MRSS, median (IQR)	4.0 (2.0–6.0)	4.0 (2.0–9.0)	0.520
Significant Raynaud's phenomenon‡	4/17 (23.5)	22/56 (39.3)	0.266
Severe GI involvement, %§	12/17 (70.6)	41/55 (74.6)	0.746
GAVE	2/17 (11.8)	3/55 (5.5)	0.586
Cardiac involvement¶	2/15 (13.3)	9/50 (18.0)	1.000
Myopathy	2/17 (11.8)	7/56 (13.0)	1.000
Lung involvement#	4/15 (26.7)	26/44 (59.1)	0.039†
Cancer	7/17 (41.2)	11/56 (19.6)	0.071
Telangiectasia	15/17 (88.2)	42/56 (75.0)	0.329
Calcinosis	3/17 (17.6)	15/56 (26.8)	0.536
Arthralgia	13/17 (76.5)	40/55 (72.7)	1.000
Synovitis	4/17 (23.5)	6/55 (10.9)	0.232
Tendon friction rub	1/17 (5.9)	4/55 (7.3)	1.000
Sicca symptoms	12/17 (70.6)	43/56 (76.8)	0.604
Pulmonary function parameters, mean ± SD			
FVC%**	78 ± 25.4	79 ± 25.4	0.897
DL_{CO}%††	72 ± 24.4	61 ± 27.1	0.202
Pulmonary fibrosis	4/17 (23.5)	24/55 (43.6)	0.165
Estimated RVSP by echocardiogram, median (IQR) mm Hg	37.0 (35.0–42.0)	31.0 (25.0–35.0)	0.031†
Thyroid	8/17 (47.1)	10/56 (17.9)	0.014†
Diabetes mellitus	1/17 (5.9)	1/56 (1.8)	0.414
Dead	0/17 (0)	2/56 (3.6)	1.000
Autoantibody positive			
Anti–topoisomerase I	1/15 (6.7)	9/54 (16.7)	0.442
Anti–RNA polymerase III	1/15 (6.7)	0/54 (0.0)	0.217
Anticentromere	9/15 (60.0)	20/54 (37.0)	0.144
Anti-Ku	1/15 (6.7)	1/54 (1.9)	0.390
Anti-Th/To	0/15 (0)	0/54 (0)	NA
Anti–U3 RNP	1/15 (6.7)	1/54 (1.9)	0.390
Anti-PM/Scl	0/15 (0.0)	4/54 (7.4)	0.570
Medications			
Proton pump inhibitor	17/20 (85.0)	53/57 (93.0)	0.367
Prokinetics	5/20 (25.0)	14/61 (23.0)	0.605
Immunosuppressant	9/15 (60.0)	28/45 (62.0)	0.878

* Values are the no./total no. (%) unless indicated otherwise. DL_{CO}% = diffusing capacity for carbon monoxide percent predicted; FVC% = forced vital capacity percent predicted; GAVE = gastric antral vascular ectasia; GI = gastrointestinal; IQR = interquartile range; MRSS = modified Rodnan skin thickness score; NA = not available; RVSP = right ventricular systolic pressure; WGT = whole-gut transit scintigraphy study.
† Statistically significant.
‡ Maximum Medsger Raynaud's Severity Score of ≥2.
§ Maximum Medsger GI Severity Score of ≥2.
¶ Maximum Medsger Cardiac Severity Score of ≥1.
# Maximum Medsger Lung Severity Score of ≥2.
** Normal FVC is >70% (percent predicted).
†† Normal DL_{CO}% is >60% (percent predicted).

GI symptom scores in SSc patients with antivinculin antibodies. We explored whether antivinculin antibodies are associated with GI symptom severity as measured by the UCLA GIT 2.0 score. Patients across the cohort had median scores consistent with moderate severity for reflux (0.8 [IQR 0.3–1.3]; reference model IQR 0.5–1.0) (24), bowel distension (1.5 [IQR 0.8–2.1]; reference model IQR 1.01–1.6), diarrhea (0.5 [IQR 0.0–0.5]; reference model IQR 0.5–1.0), constipation (0.9 [IQR 0.5–1.3];

Table 2. Measuring the association between antibody positivity and systemic sclerosis (SSc) clinical characteristics*

Clinical and demographic features	OR (95% CI)	*P*
Age at first symptom, years	1.05 (0.99, 1.10)	0.089
Female	0.90 (0.16, 4.93)	0.903
Race		
White	6.00 (0.73, 49.3)	0.095
Disease duration from first non-Raynaud's symptom to date of WGT study	1.04 (0.97, 1.12)	0.237
Ever smoker	0.75 (0.23, 2.44)	0.632
SSc type		
Limited cutaneous disease	9.60 (1.19, 77.23)	0.034†
Significant Raynaud's phenomenon‡	0.48 (0.14, 1.65)	0.241
Severe GI involvement§	0.82 (0.25, 2.74)	0.747
GAVE	2.31 (0.35, 15.1)	0.382
Cardiac involvement¶	0.70 (0.13, 3.67)	0.674
Myopathy	0.93 (0.18, 4.98)	0.936
Lung involvement#	0.25 (0.07, 0.92)	0.036†
Cancer	2.86 (0.89, 9.22)	0.078
Telangiectasia	2.50 (0.51, 12.32)	0.260
Calcinosis	0.59 (0.15, 2.33)	0.447
Arthralgia	1.22 (0.34, 4.33)	0.760
Synovitis	2.51 (0.62, 10.24)	0.199
Tendon friction rub	0.80 (0.08, 7.65)	0.844
Sicca symptoms	0.73 (0.22, 2.44)	0.604
Pulmonary function parameters	1.00 (0.98, 1.02)	0.895
FVC%**		
DLco%††	1.01 (1.00, 1.04)	0.202
Estimated RVSP by echocardiogram	1.22 (1.00, 1.49)	0.051
Thyroid	4.09 (1.27, 13.21)	0.019†
Diabetes mellitus	3.44 (0.20, 58.09)	0.392
Autoantibody positive		
Anti–topoisomerase I	0.35 (0.04, 3.07)	0.348
Anticentromere	2.55 (0.79, 8.23)	0.117
Anti-Ku	3.79 (0.22, 64.39)	0.357
Anti–U3 RNP	3.79 (0.22, 64.39)	0.357

* 95% CI = 95% confidence interval; DLco% = diffusing capacity for carbon monoxide percent predicted; FVC% = forced vital capacity percent predicted; GAVE = gastric antral vascular ectasia; GI = gastrointestinal; OR = odds ratio; RVSP = right ventricular systolic pressure; WGT = whole-gut transit scintigraphy study.
† Significant.
‡ Maximum Medsger Severity Score of ≥2.
§ Maximum Medsger GI Severity Score of ≥2.
¶ Maximum Medsger cardiac Severity Score of ≥1.
# Maximum Medsger Severity Score of ≥2.
** Normal FVC is >70%.
†† Normal DLco% is >60%.

reference model IQR 0.5–1.0), and impact on emotional well-being (0.7 [IQR 0.2–1.8]; reference model IQR 0.5–1.0) and a median score consistent with mild impairment in social functioning (0.3 [IQR 0.2–1.1]; reference model IQR 0.0–0.49) based on the previously reported definitions of GI symptom severity (24). We found no significant association between the UCLA GIT 2.0 total score and positivity for antivinculin antibodies.

Association between whole-gut transit and antivinculin antibodies. To determine whether antivinculin antibodies are associated with delayed GI transit and severity, we examined the associations between delayed transit in the esophagus, stomach, small bowel, and colon as measured by the WGT study and antivinculin antibody positivity (Table 3). Antivinculin antibodies were more prevalent among patients with delayed solid gastric emptying at 4 hours than in those without (35% versus 22%), although this was not statistically significant. Antivinculin antibodies were not enriched among patients with delayed transit in other areas of the gut. Interestingly, a significant correlation was noted between higher antivinculin antibody levels and lower (worse) percent gastric emptying at 4 hours (β coefficient –3.41 [95% CI –6.72, –0.09], $P = 0.04$). No significant correlations were identified for liquid and solid gastric emptying at 1 or 2 hours, esophageal transit time, small bowel emptying (6 hours), or colonic emptying (24, 48, and 72 hours).

Multivariable models. We then sought to determine whether the associations between antivinculin antibody positivity and clinical characteristics of SSc remained after adjusting for age and sex (Table 4). In the multivariable model, antivinculin antibody positivity significantly associated with a higher odds of limited SSc (OR 8.99 [95% CI 1.05, 76.83], $P = 0.05$), 78% lower risk of lung involvement (OR 0.22 [95% CI 0.06, 0.86], $P = 0.03$), an increased odds of thyroid disease (OR 3.87 [95% CI 1.16, 12.93], $P = 0.03$), and a higher estimated RVSP (β coefficient 6.41 [95% CI 0.37, 12.45], $P = 0.04$). In the multivariable model, a significant association between the positivity for antivinculin antibodies and percent

Table 3. Linear regression analyses evaluating the association between antivinculin antibody levels and gastrointestinal transit*

Region of the gut	β coefficient (95% CI)	*P*
Esophagus		
Esophageal transit time†	–1.18 (–4.17, 1.82)	0.437
Esophageal % emptying at 10 seconds‡	1.65 (–4.94, 8.24)	0.619
Stomach		
Liquid§		
% emptying at 1/2 hour	7.03 (9.59, 23.6)	0.399
% emptying at 1 hour	–0.83 (–5.70, 4.04)	0.735
% emptying at 2 hours	1.48 (–5.73, 2.78)	0.492
Solid¶		
% emptying at 1 hour	–2.44 (–8.03, 3.13)	0.384
% emptying at 2 hours	–1.54 (–7.94, 4.86)	0.632
% emptying at 4 hours	–3.41 (–6.72, –0.09)	0.044#
Small bowel**		
% emptying at 6 hours	–2.47 (–8.40, 3.46)	0.409
Colon††		
% emptying at 24 hours	–2.90 (–9.58, 3.78)	0.389
% emptying at 48 hours	–5.68 (–25.12, 13.76)	0.562
% emptying at 72 hours	–8.38 (–28.92, 12.14)	0.419

* 95% CI = 95% confidence interval.
† Normal esophageal transit time: >15 seconds.
‡ Normal esophageal emptying at 10 seconds: ≥83%.
§ Normal liquid half-life: ≤74 minutes.
¶ Normal solid emptying at 2 hours: ≥40%; normal solid emptying at 4 hours: ≥90%.
# Significant.
** Normal small bowel transit time at 6 hours: ≥49%.
†† Normal % colonic emptying at 72 hours: ≥67%.

Table 4. Multivariable modeling with clinical and demographic features*

	% solid gastric emptying, β (95% CI)†	*P*	Thyroid, OR (95% CI)	*P*	Limited cutaneous disease, OR (95% CI)	*P*	Lung involvement, OR (95% CI)‡	*P*	RVSP, mm Hg, β (95% CI)	*P*
Antivinculin antibody (+/–)	–6.20 (–12.33, –0.063)	0.048§	3.87 (1.16, 12.93)	0.028§	8.99 (1.05, 76.83)	0.045§	0.22 (0.06, 0.86)	0.029§	6.41 (0.37, 12.45)	0.038§
Antivinculin antibody levels	–3.64 (–7.05, –0.23)	0.037§	1.78 (0.87, 3.64)	0.117	2.44 (1.05, 5.70)	0.038§	0.48 (0.22, 1.02)	0.057	3.38 (0.04, 6.72)	0.047§

* All analyses were adjusted for age and sex. Positive antivinculin antibodies had a threshold of optical density ≥1.68. 95% CI = 95% confidence interval; OR = odds ratio; RVSP = right ventricular systolic pressure (measured by echocardiogram).
† At 4 hours.
‡ Maximum Medsger Severity Score of ≥2.
§ Significant.

gastric emptying remained (β coefficient –6.20 [95% CI –12.33, –0.063], *P* = 0.05). Because thyroid disease can negatively impact GI motility, we performed an additional analysis re-evaluating the association between antivinculin antibody positivity and percent gastric emptying after also adjusting for thyroid disease. Importantly, the inverse association between antivinculin antibody levels and slow gastric transit persisted even after adjustment (β coefficient –6.94 [95% CI –13.1, –0.76], *P* = 0.03) (Table 4 and Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25118).

When examining the relationships between antivinculin antibody levels and clinical characteristics in the adjusted model, an inverse association between antivinculin antibody level and percent gastric emptying at 4 hours remained (β coefficient –3.64 [95% CI –7.05, –0.23], *P* = 0.04). The association between higher antivinculin antibody levels and a higher odds of having limited cutaneous disease (OR 2.44 [95% CI 1.05, 5.70], *P* = 0.04) and a higher RVSP (β coefficient 3.38 [95% CI 0.04, 6.72], *P* = 0.05) also remained. Trends toward an association between higher antivinculin antibody levels and less lung involvement (OR 0.48 [95% CI 0.22, 1.02], *P* = 0.06) and thyroid disease (OR 1.78 [95% CI 0.87, 3.64], *P* = 0.12) were identified, although these associations no longer reached statistical significance.

DISCUSSION

Antivinculin antibodies are a novel autoantibody specificity identified in several GI diseases including IBS, functional dyspepsia, and SSc. This is the first study to describe the extraintestinal clinical features of SSc patients who were positive for antivinculin antibodies and to examine the association between antivinculin antibodies and objective measures of GI transit by whole-gut scintigraphy. Using a previously established and validated ELISA, we found that antivinculin antibodies are common in SSc and were present in 23% of our patient cohort. This prevalence is similar to the findings of a study by Suliman et al in which antivinculin antibodies were identified in 37% of a SSc cohort enriched for GI disease and in 32% of an SSc cohort enriched for vascular disease (15). While our cohort had slightly more severe GI disease compared to patients who did not complete the WGT study (25), we also included patients who were minimally symptomatic, which could explain why the prevalence of antivinculin antibodies in our cohort was slightly lower.

We report for the first time that antivinculin antibody positivity is associated with slow gastric transit in SSc, and that antivinculin antibody levels inversely correlate with percent gastric emptying. This finding is consistent with those of a recent study that demonstrated an inverse correlation between higher levels of circulating antivinculin antibody levels and the number of interstitial cells of Cajal (ICC) in the myenteric plexus of the human stomach (26). Vinculin is located in the ICC, which function as the pacemaker cells of the GI tract. It is not yet clear whether antivinculin antibodies have a direct pathogenic effect on the ICC resulting in GI dysmotility or whether they are a marker of GI dysfunction and ICC attrition. We did not find an association between antivinculin antibodies and percent emptying in the esophagus, small bowel, or colon, despite the fact that ICC and vinculin are present throughout the GI tract. It is possible that ICC play a different role in the stomach than in other parts of the GI tract, although further studies are needed to understand why the association between antivinculin antibodies and GI dysmotility is specific for the stomach.

Antivinculin antibodies in IBS are thought to result from antibodies raised by the host against cytolethal distending toxin B (CdtB), which is a toxin produced by gram-negative bacteria that cause gastroenteritis and cross-react with vinculin. This has been demonstrated in postinfectious models of IBS (14). It is not clear if this model can be extrapolated to SSc, as GI symptoms often occur later in the disease course and are not necessarily preceded by an infectious GI illness. It is possible that small intestinal bacterial overgrowth (SIBO) in SSc leads to the build-up of bacteria and the release of toxin (27), which then generates an immune response against CdtB and vinculin. It would be instructive to screen SSc sera for antibodies against both CdtB and vinculin.

We did not find an association between positivity for antivinculin antibodies and GI symptom scores measured by the UCLA GIT 2.0 questionnaire. This is consistent with a study by Suliman et al that did not find an association between antivinculin antibodies and UCLA GIT 2.0 (15) scores. However, they did find a

positive association between higher antivinculin antibody levels and GI VAS scores. This could be due to differences in patient-reported outcome measures used to assess GI symptoms in SSc. We previously found a poor correlation between GI symptom scores measured by the GIT 2.0 and objective transit measures on the WGT study (28). Moreover, the GIT 2.0 was not administered on the same day as the blood draw and WGT study, which may have impaired our ability to determine associations with symptom scores.

We found that thyroid disease was enriched among SSc patients who were positive for antivinculin antibodies. Hypothyroidism is known to alter GI dysmotility and lead to an increased prevalence of SIBO (29). However, when we also adjusted for hypothyroidism in the multivariable model, the association remained between antivinculin antibodies and slow gastric transit, suggesting that thyroid disease is not the primary driver behind this relationship. We also found that patients with limited cutaneous disease and/or patients with less severe lung disease were more likely to be antivinculin antibody positive. Our findings suggest that patients with this clinical phenotype should be monitored for upper GI symptoms more vigilantly. Interestingly, Suliman et al previously found a trend toward more pulmonary arterial hypertension among patients with antivinculin antibodies in an SSc group enriched for vascular disease (15). The positive association that we identified between antivinculin antibodies and a higher estimated RVSP further supports this finding. Of note, we did not identify an association between vinculin antibodies and sex, which is particularly notable given that slow gastric transit, functional dyspepsia, and IBS are significantly more common among women (30,31). This suggests that the mechanisms causing GI disease in SSc are (partly) distinct from those that cause functional GI disorders in the general population.

Strengths of this study are the use of a well-characterized, prospective SSc cohort with objective transit data measured by WGT scintigraphy. There are several limitations of our study. Our cohort was enriched for GI disease, so we were unable to determine the true prevalence of antivinculin antibodies in a general SSc cohort. Also, the GIT 2.0 questionnaire was not administered concurrently with the WGT study and the blood draw. Although prior studies have shown that GI symptoms reflect damage from the disease and are relatively stable over time (32), this still could have introduced some variability into our study. Last, mechanistic investigations should be conducted to understand the association between gastric emptying and antivinculin antibodies.

In conclusion, we demonstrated that antivinculin antibodies are common in SSc and are associated with delayed gastric emptying on the WGT study, limited cutaneous disease, a high RVSP on echocardiogram, thyroid disease, and less severe lung disease. If validated in another cohort using objective transit data, antivinculin antibodies could potentially be a useful marker for slow gastric transit and extraintestinal phenotype in SSc.

ACKNOWLEDGMENTS

The authors thank Dr. Jay Pasricha for his contributions.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Herrán, Adler, McMahan.

Acquisition of data. Morales, Pimentel, McMahan.

Analysis and interpretation of data. Herrán, Adler, Perin, Morales, Pimentel, McMahan.

REFERENCES

1. Anaya JM, Shoenfeld Y, Rojas-Villarrage, et al, editors. Autoimmunity: from bench to bedside. Bogota: El Rosario University Press; 2013. p. 611.
2. Shreiner AB, Murray C, Denton C, et al. Gastrointestinal manifestations of systemic sclerosis. J Scleroderma Relat Disord 2016;1:247–56.
3. McMahan ZH. Gastrointestinal involvement in systemic sclerosis: an update. Curr Opin Rheumatol 2019;31:561–8.
4. Gourh P, Safran SA, Alexander T, et al. HLA and autoantibodies define scleroderma subtypes and risk in African and European Americans and suggest a role for molecular mimicry. Proc Natl Acad Sci U S A 2020;117:552–62.
5. Srivastava N, Hudson M, Tatibouet S, et al. Thinking outside the box: the associations with cutaneous involvement and autoantibody status in systemic sclerosis are not always what we expect. Semin Arthritis Rheum 2015;45:184–9.
6. Moroncini G, Svegliati Baroni S, et al. Agonistic antibodies in systemic sclerosis. Immunol Lett 2018;195:83–7.
7. Nakane S, Umeda M, Kawashiri SY, et al. Detecting gastrointestinal manifestations in patients with systemic sclerosis using anti-gAChR antibodies. Arthritis Res Ther 2020;22:32.
8. Vasapolli R, Schulz C, Schweden M, et al. Gut microbiota profiles and the role of anti-CdtB and anti-vinculin antibodies in patients with functional gastrointestinal disorders (FGID). Eur J Clin Invest 2021;51:e13666.
9. Talley NJ, Holtmann G, Walker MM, et al. Circulating anti-cytolethal distending toxin B and anti-vinculin antibodies as biomarkers in community and healthcare populations with functional dyspepsia and irritable bowel syndrome. Clin Transl Gastroenterol 2019;10:e00064.
10. Pimentel M, Morales W, Rezaie A, et al. Development and validation of a biomarker for diarrhea-predominant irritable bowel syndrome in human subjects. PLoS One 2015;10:e0126438.
11. Deroanne CF, Colige AC, Nusgens BV, et al. Modulation of expression and assembly of vinculin during in vitro fibrillar collagen-induced angiogenesis and its reversal. Exp Cell Res 1996;224:215–23.
12. Bays JL, DeMali KA. Vinculin in cell-cell and cell-matrix adhesions. Cell Mol Life Sci 2017;74:2999–3009.
13. Rezaie A, Park SC, Morales W, et al. Assessment of anti-vinculin and anti-cytolethal distending toxin B antibodies in subtypes of irritable bowel syndrome. Dig Dis Sci 2017;62:1480–5.
14. Pimentel M, Morales W, Pokkunuri V, et al. Autoimmunity links vinculin to the pathophysiology of chronic functional bowel changes following campylobacter jejuni infection in a rat model. Dig Dis Sci 2015;60:1195–205.
15. Suliman Y, Kafaja S, Oh SJ, et al. Anti-vinculin antibodies in scleroderma (SSc): a potential link between autoimmunity and

gastrointestinal system involvement in two SSc cohorts. Clin Rheumatol 2021;40:2277–84.
16. Khanna D, Hays RD, Maranian P, et al. Reliability and validity of the University of California, Los Angeles Scleroderma Clinical Trial Consortium Gastrointestinal Tract Instrument. Arthritis Rheum 2009;61: 1257–63.
17. Van den Hoogen F, Khanna D, Fransen J, et al. 2013 classification criteria for systemic sclerosis: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 2013;65:2737–47.
18. Subcommittee for Scleroderma Criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. Preliminary criteria for the classification of systemic sclerosis (scleroderma). Arthritis Rheum 1980;23:581–90.
19. Hachulla E, Launay D. Diagnosis and classification of systemic sclerosis. Clin Rev Allergy Immunol 2011;40:78–83.
20. Medsger TA Jr, Silman AJ, Steen VD, et al. A disease severity scale for systemic sclerosis: development and testing. J Rheumatol 1999;26: 2159–67.
21. Abignano G, Mennillo GA, Lettieri G, et al. UCLA Scleroderma Clinical Trials Consortium Gastrointestinal Tract (GIT) 2.0 reflux scale correlates with impaired esophageal scintigraphy findings in systemic sclerosis. J Rheumatol 2021;48:1422–6.
22. Antoniou AJ, Raja S, El-Khouli R, et al. Comprehensive radionuclide esophagogastrointestinal transit study: methodology, reference values, and initial clinical experience. J Nucl Med 2015;56:721–7.
23. Morales W, Rezaie A, Barlow G, et al. Second-generation biomarker testing for irritable bowel syndrome using plasma anti-CdtB and antivinculin levels. Dig Dis Sci 2019;64:3115–21.
24. Khanna D, Nagaraja V, Gladue H, et al. Measuring response in the gastrointestinal tract in systemic sclerosis. Curr Opin Rheumatol 2013;25:700–6.
25. Cheah JX, Perin J, Volkmann ER, et al. Slow colonic transit in systemic sclerosis: an objective assessment of risk factors and clinical phenotype. Arthritis Care Res (Hoboken) 2023;75: 289–98.
26. Kim JH, Nam SJ, Park SC, et al. Association between interstitial cells of Cajal and anti-vinculin antibody in human stomach. Korean J Physiol Pharmacol 2020;24:185–91.
27. Ganji L, Shirazi MH, Ebrahimi-Daryani N, et al. Carriage of CdtB encoding Campylobacter spp., Salmonella enterica, and Yersinia entercolitica in patients with gastroenteritis and irritable bowel syndrome. Dig Dis Sci 2022;67:5522–8.
28. Adler B, Hummers LK, Pasricha PJ, et al. Gastroparesis in systemic sclerosis: a detailed analysis using whole-gut scintigraphy. Rheumatology (Oxford) 2022;61:4503–8.
29. Yaylali O, Kirac S, Yilmaz M, et al. Does hypothyroidism affect gastrointestinal motility? Gastroenterol Res Pract 2009;2009:529802.
30. Houghton LA, Heitkemper M, Crowell M, et al. Age, gender and women's health and the patient. Gastroenterology 2016;S0016–5085(16)00183–9.
31. Kim YS, Kim N. Sex-gender differences in irritable bowel syndrome. J Neurogastroenterol Motil 2018;24:544–58.
32. Volkmann ER, Tashkin DP, LeClair H, et al. Treatment with mycophenolate and cyclophosphamide leads to clinically meaningful improvements in patient-reported outcomes in scleroderma lung disease: results of Scleroderma Lung Study II. ACR Open Rheumatol 2020;2: 362–70.

Arthritis Care & Research
Vol. 75, No. 10, October 2023, pp 2174–2181
DOI 10.1002/acr.25112

American College of Rheumatology
Empowering Rheumatology Professionals

Persistence of Biologics in the Treatment of Psoriatic Arthritis: Data From a Large Hospital-Based Longitudinal Cohort

Mohamad-Ali Rida,[1] Ker-Ai Lee,[2] Vinod Chandran,[3] Richard J. Cook,[2] and Dafna D. Gladman[3]

Objective. To analyze the trends in biologics use at a specialized center over a period of 20 years.

Methods. We performed a retrospective analysis of 571 patients diagnosed with psoriatic arthritis enrolled in the Toronto cohort who initiated biologic therapy between January 1, 2000, and July 7, 2020. The probability of drug persistence over time was estimated nonparametrically. The time to discontinuation of first and second treatment was analyzed using Cox regression models, whereas a semiparametric failure time model with a gamma frailty was used to analyze the discontinuation of treatment over successive administrations of biologic therapy.

Results. The highest 3-year persistence probability was observed with certolizumab when used as first biologic treatment, while interleukin-17 inhibitors had the lowest probability. However, when used as second medication, certolizumab had the lowest drug survival even when accounting for selection bias. Depression and/or anxiety were associated with a higher rate of drug discontinuation due to all causes (relative risk [RR] 1.68, $P = 0.01$), while having higher education was associated with lower rates (RR 0.65, $P = 0.03$). In the analysis accommodating multiple courses of biologics, a higher tender joint count was associated with a higher rate of discontinuation due to all causes (RR 1.02, $P = 0.01$). Older age at the start of first treatment was associated with a higher rate of discontinuation due to side effects (RR 1.03, $P = 0.01$), while obesity had a protective role (RR 0.56, $P = 0.05$).

Conclusion. Persistence in taking biologics depends on whether the biologic was used as first or second treatment. Depression and anxiety, higher tender joint count, and older age lead to drug discontinuation.

INTRODUCTION

Psoriatic arthritis (PsA) is the most common extracutaneous manifestation of psoriasis and develops in ∼24% of patients with psoriasis (1). The development of new biologic disease-modifying antirheumatic drugs (bDMARDs) for PsA has led to better disease control as well as slowing of radiographic progression (2). Current advanced therapies include tumor necrosis factor inhibitors (TNFi), interleukin-17 inhibitors (IL-17i), IL-12/IL-23 inhibitors (IL-12/IL-23i), T cell activation blockers, phosphodiesterase 2 inhibitors (PDE4i), and JAK inhibitors (3). Multiple treatment recommendations have been developed to guide treatment of PsA (4–6). However, there are limited data and no consensus on switching between different DMARDs; therefore, the practice differs from one region to another (7). This is also the result of the heterogeneity of the disease and its wide spectrum of manifestations affecting skin and musculoskeletal structures (8). Once an advanced therapy fails, due to either lack of efficacy or side effects, switching to another one from the same or different class is the norm.

Several registries have conducted studies on drug survival of biologic agents in PsA, including those from Denmark (9), Norway (10), Sweden (11), Finland (12), the US (13,14), Japan (15), Israel (16), and the UK (17). All these registries, however, included only TNFi. The choice of first biologic and subsequent switching is based on multiple factors, including drug availability, physician's assessment, patient's preferences concerning dosing frequency, safety profiles, and immunogenicity (7). In the present study, we aimed to investigate the use of advanced therapy for PsA in patients followed prospectively at a specialized center over a

The Psoriatic Arthritis Program is supported by a grant from the Krembil Foundation.

[1]Mohamad-Ali Rida, MD: Schroeder Arthritis Institute, Krembil Research Institute, and Toronto Western Hospital, University Health Network, Toronto, Ontario, Canada; [2]Ker-Ai Lee, MMath, Richard J. Cook, PhD: University of Waterloo, Waterloo, Ontario, Canada; [3]Vinod Chandran, MBBS, MD, DM, PhD, Dafna D. Gladman, MD, FRCPC: Schroeder Arthritis Institute, Krembil Research Institute, and Toronto Western Hospital, University Health Network and University of Toronto, Toronto, Ontario, Canada.

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.25112&file=acr25112-sup-0001-Disclosureform.pdf.

Address correspondence via email to Dafna D. Gladman, MD, FRCPC, at dafna.gladman@utoronto.ca.

Submitted for publication May 19, 2022; accepted in revised form March 9, 2023.

SIGNIFICANCE & INNOVATIONS

- Studies on the persistence of biologics use in psoriatic arthritis (PsA) have concentrated on anti–tumor necrosis factor (anti-TNF) agents.
- This study, performed on a longitudinal cohort of patients with PsA, investigated the persistence of several anti-TNF agents and interleukin-17 inhibitors, as well as reasons for discontinuation.
- Persistence in taking a biologic depends on the order of use.
- Depression, older age, and tender joint count lead to discontinuation of biologic therapy.

period of 20 years, with a focus on persistence and switching medications.

PATIENTS AND METHODS

Patients. Patients with PsA have been followed prospectively according to a standard protocol at the Toronto Western Hospital PsA clinic since 1978. For this analysis on advanced therapies, only patients who initiated advanced therapy between January 1, 2000 and July 7, 2020 were included. Data recorded at the initial visit and subsequent visits at intervals of 6–12 months (more often if necessary) include demographic characteristics, current medications, disease activity scores, and new prescriptions, including advanced therapy. For each medication, the treatment start and stop dates as well as reasons for stopping are recorded.

Advanced therapy. Five TNFi and 2 IL-17i have been available at the PsA clinic in the past 20 years. While other therapies have become available in the past few years, including IL-12/IL-23i, IL-23i, and PDE4i, the number of patients using each of these medications was too small, and they were not included in this analysis.

For patients who received multiple courses of the same biologic, we allowed a suspension of use up to 180 days between 2 consecutive courses for the same medication while still treating these consecutive prescriptions as 1 course. If the stop duration between 2 consecutive prescriptions was >180 days, the next prescription (even if it was the same advanced treatment) was treated as a second course. As not all the advanced therapies were available between 2000 and 2010, we considered treatments received since January 1, 2000 for graphical summary only. For regression analyses, we considered treatments received since January 1, 2010 since more advanced therapies could be offered to patients after 2010.

Ethics. All database records are fully compliant with the Canadian confidentiality requirements, including the Health Insurance Portability and Accountability Act (HIPAA) of 1996. All patients included in the cohort have signed consents, and the study was approved by the Research Ethics Board of the University Health Network.

Statistical analysis. The probabilities of discontinuation of the 5 TNFi (adalimumab, certolizumab pegol, etanercept, golimumab, and infliximab) and 3 IL-17i (combined ixekizumab, secukinumab, and brodalumab) were estimated using the Kaplan-Meier method (18). A 20-year period of data was used and then examined by 10-year periods of calendar time to account for the changing availability of various biologics over the years, primarily of the IL-17i. The probability of drug persistence was plotted to

Table 1. Characteristics of the patient population*

Variable	At the first visit following January 1, 2000 (n = 571 patients)	At the first visit following January 1, 2010 (n = 522 patients)
Age at start of treatment, years	47.7 ± 13.2	48.2 ± 13.0
Age at onset of PsA, years	36.7 ± 13.5	36.7 ± 13.5
Sex, no. (%)		
Male	325 (56.9)	301 (57.7)
Female	246 (43.1)	221 (42.3)
Education, no. (%)		
High school or lower	137 (24.4)	118 (23.0)
College or higher	424 (75.6)	394 (77.0)
Married, no. (%)	351 (62.1)	317 (61.4)
Alcohol, no. (%)		
None	256 (45.1)	233 (45.0)
Social	275 (48.5)	256 (49.4)
Daily	36 (6.3)	29 (5.6)
Actively inflamed joints†	10.0 ± 11.1	9.1 ± 10.4
Tender joints	5.7 ± 8.6	5.0 ± 7.6
Swollen joints	1.8 ± 3.5	1.3 ± 2.8
Damaged joints	4.8 ± 9.1	5.0 ± 9.4
Patients with enthesitis, no. (%)	126 (22.1)	116 (22.2)
Patients with dactylitis, no. (%)	118 (20.7)	98 (18.8)
Among patients with a radiograph, those with axial involvement, no. (%)	190 (35.1)‡	186 (37.6)§
Fibromyalgia, no. (%)	70 (12.5)	52 (10.2)
Obesity, no. (%)	177 (39.2)	169 (38.1)
Depression/anxiety, no. (%)¶	100 (18.9)	91 (18.9)

* Values are the mean ± SD unless indicated otherwise. PsA = psoriatic arthritis.
† Actively inflamed joints include those that are either tender or swollen or both.
‡ A total of 29 of 571 patients did not have a radiographic assessment.
§ A total of 27 of 522 patients did not have a radiographic assessment.
¶ Depression/anxiety includes patients who were documented to have depression or anxiety in the protocol or were taking antidepressants.

characterize the overall risk of discontinuation due to any cause as a function of time since prescription of the first biologic treatment.

To estimate the probability of drug persistence for the second biologic treatment, it is necessary to deal with possible dependence between the durations of successive courses of biologic therapies and the resultant induced dependent censoring and selection bias for the second biologic therapy (18; section 4.4.1). To address this, a semiparametric mixed failure time model was fitted with a gamma frailty (18) to examine the duration of the second course of TNFi or IL-17i treatments. An estimate of the baseline cumulative hazard for the distribution of the second course was then extracted from this joint model and used to plot the estimate of the survival distribution for each biologic treatment. The semiparametric mixed model with a gamma frailty (19) was also used to examine the duration of the first and subsequent TNFi or IL-17i treatments. Models were stratified on a 5-year window of calendar time, current type of treatment (5 TNFi and IL-17i), whether or not the current treatment was the second or later treatment, and treatment history (including the nature of the first-line treatment [TNFi versus non-TNFi] and

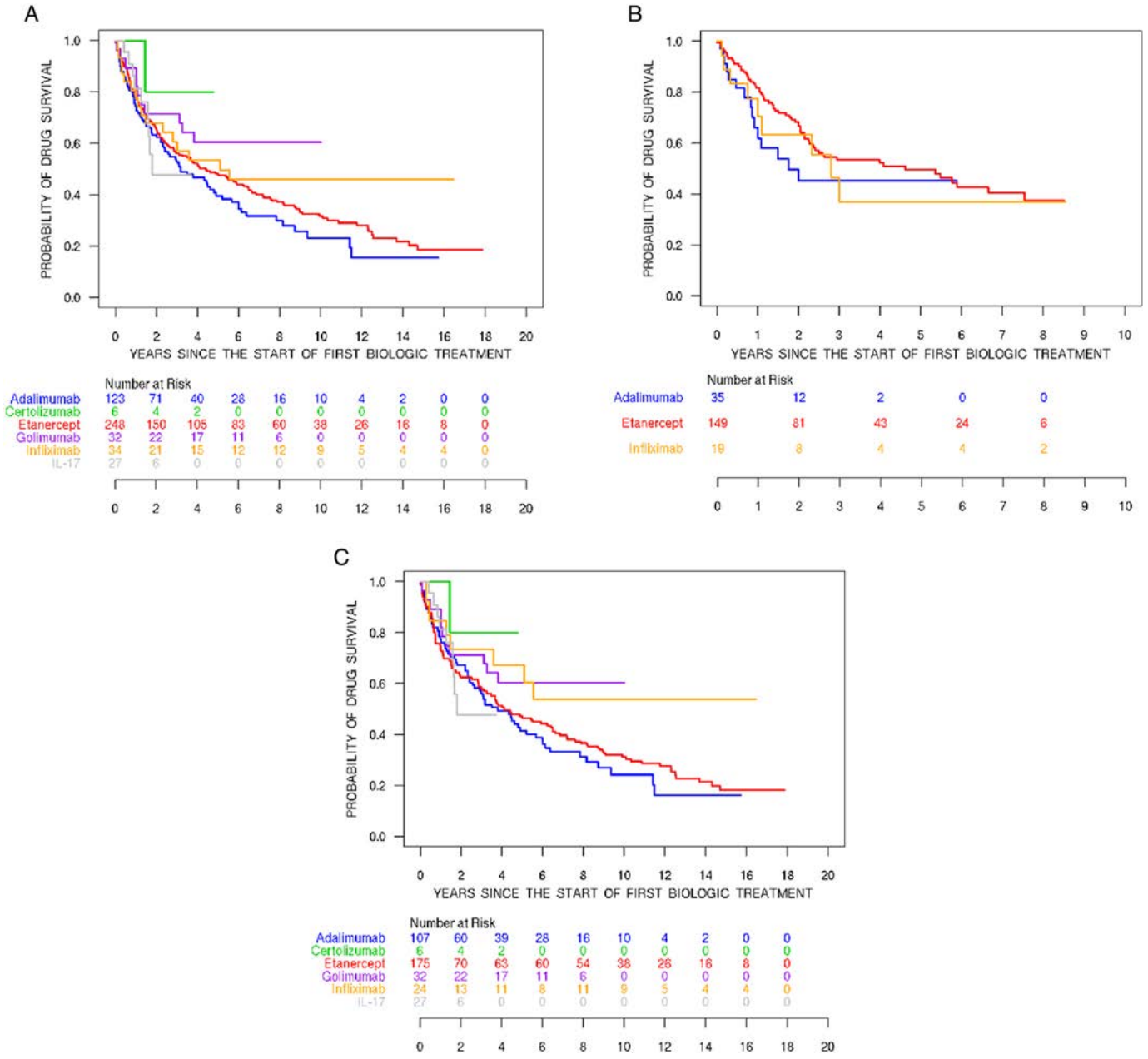

Figure 1. Probability of drug survival for the first biologic treatment over the total periods 2000–2020 (**A**), 2000–2009 (**B**), and 2010–2020 (**C**) for adalimumab (blue), certolizumab (green), etanercept (red), golimumab (purple), infliximab (orange), and IL-17 inhibitors (gray); see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25112.

the duration of the first course of biologic therapy). Similar models were fitted to examine each type of treatment alone (adalimumab, certolizumab pegol, etanercept, golimumab, infliximab, or IL-17i) and the discontinuation of treatment due to all causes.

To assess the determinants of discontinuation of the first treatment for all causes and by each specific cause, we fitted both overall and cause-specific Cox regression models (19), controlling for a 3-year period of calendar time and current type of treatment. Next, we fitted an analogous model to examine the time to discontinuation of the second treatment. In these models, we controlled for a 3-year period of calendar time, the duration between the first and second treatment, and previous and current type of treatment.

In the full multivariate models, the covariates included age at the start of treatment, sex, education level, marital status, alcohol consumption, the number of tender and/or swollen joints, fibromyalgia status, obesity, and depression/anxiety. Reduced multivariate models were obtained using backward elimination with covariates retained if *P* values were less than 0.05. Covariates retained in the reduced models are shown in the Results section. All analyses were performed using R, version 3.6.2 (20).

RESULTS

A total of 571 patients were included in the study. Among these patients, information on 584 prescriptions since January 1, 2000 was retrieved (Table 1). The most common biologic prescribed was etanercept (39%), followed by adalimumab (27%), IL-17i (13%), infliximab (8%), golimumab (7%), and then certolizumab pegol (5%). Figure 1 shows the probability of drug survival for first treatment received since January 1, 2000. Estimated 3-year drug survival probability along with the 95% confidence interval (95% CI) is provided for each biologic treatment (Figure 1). At the first biologic treatment, the highest 3-year survival probability was observed with certolizumab (0.80 [95% CI 0.52–1.00]), while IL-17i had the lowest probability (0.48 [95% CI 0.28–0.81]) (Figure 1A).

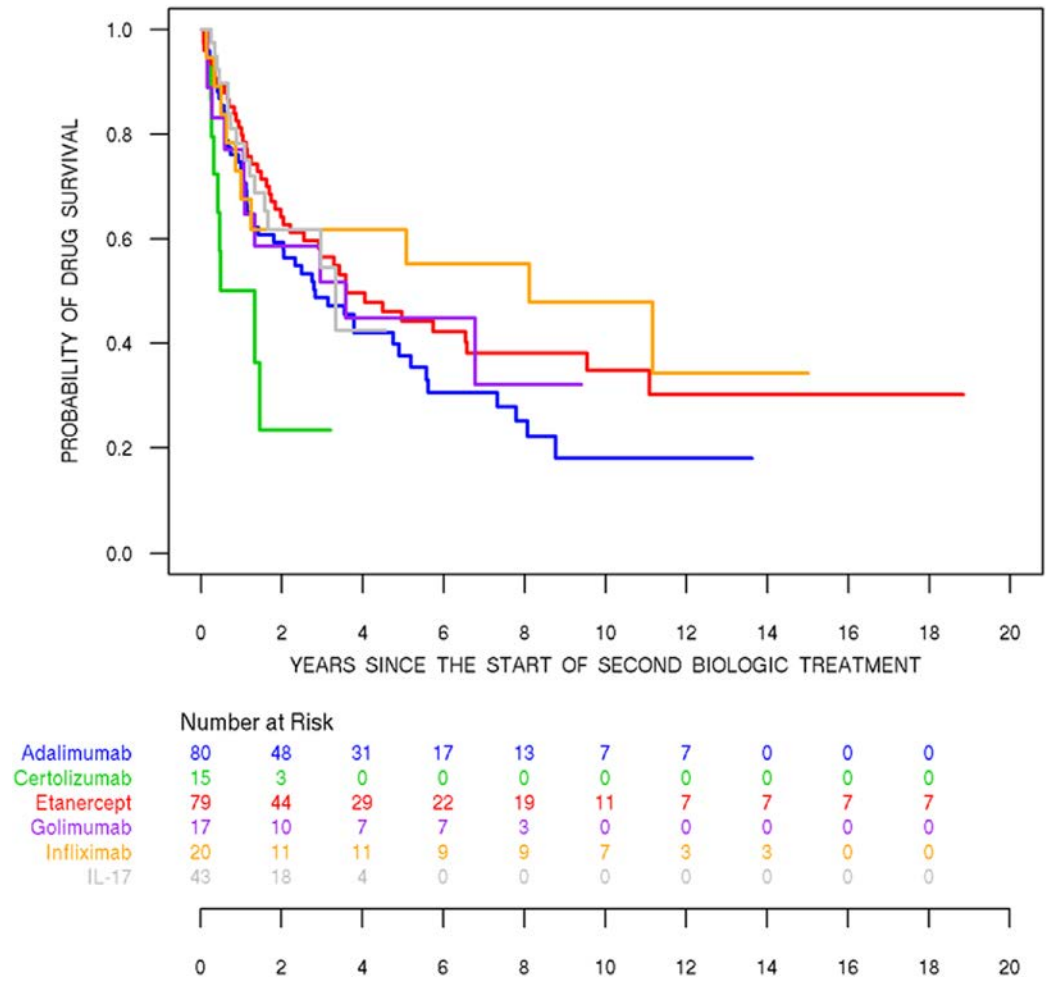

Figure 2. Probability of drug survival for the second biologic treatment received between 2000 and 2020 for adalimumab (blue), certolizumab (green), etanercept (red), golimumab (purple), infliximab (orange), and IL-17 inhibitors (gray); see Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25112. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25112/abstract.

Table 2. Analysis of the duration of all courses of biologic therapy*

	Cox model for overall discontinuation and by cause						Cox model for all course discontinuation and by biologic					
	All causes	*P*	Side effects	*P*	Other reasons	*P*	Adalimumab	*P*	Etanercept	*P*	Il-17i	*P*
Age at start of biologic, years	–		1.03 (1.01–1.05)	0.013	–		–		–		–	
College or higher	–		–		0.56 (0.33–0.96)	0.034	–		–		–	
No. of tender joints	1.02 (1.00–1.03)	0.013	1.03 (1.00–1.06)	0.025	–		–		–		1.07 (1.01–1.14)	0.029
No. of swollen joints	–		–		1.09 (1.00–1.18)	0.042	0.86 (0.75–0.98)	0.023	–		–	
Obese versus not obese	–		0.56 (0.31–1.00)	0.049	–		–		–		0.28 (0.12–0.66)	0.004
No. of both tender and swollen joints	–		–		–		–		–		1.10 (1.00–1.19)	0.038
Depression/anxiety	–		–		–		–		1.87 (1.00–3.48)	0.048	–	

* Values are the relative risk (95% confidence interval) unless indicated otherwise. These relative risks are based on mixed-effects multiplicative Cox-type models for the duration of the first and subsequent courses of biologics with a patient-specific frailty to accommodate the association between the durations of each biologic course. For analyses of discontinuation due to side effects and other causes, cause-specific models were used to address the competing risk issue. No significant covariates were found for certolizumab, golimumab, and infliximab. Il-17i = interleukin-17 inhibitors.

Table 3. Cox regression analysis for the time to discontinuation of the first course of biologic therapy for all causes and by each cause*

	All causes	P	Lack of efficacy	P	Side effects	P	Other reasons	P
College or higher	0.65 (0.44–0.96)	0.029	0.57 (0.33–0.96)	0.035	–		0.41 (0.20–0.86)	0.016
No. of tender joints	–		–		1.06 (1.02–1.1)	0.002	–	
No. of swollen joints	–		–		–		1.14 (1.03–1.25)	0.008
Depression/anxiety	1.68 (1.11–2.53)	0.013	–		–		2.32 (1.05–5.12)	0.037

* Values are the relative risk (95% confidence interval) unless indicated otherwise. Relative risks were estimated based on Cox regression models or cause-specific Cox regression models for side effects and discontinuation for other reasons.

Etanercept had the highest probability of drug survival (0.56 [95% CI 0.50–0.63]) during the period 2000–2009 (Figure 1B). Similar results were obtained for the period 2010–2020 (Figure 1C). For second biologic treatment, certolizumab had the lowest 3-year probability of drug survival (0.23 [95% CI 0.09–0.62]), while infliximab had the highest probability (0.62 [95% CI 0.42–0.90]) (Figure 2). No significant difference was found when comparing all TNFi to IL-17i (data not shown).

In the regression analyses that follow, we consider treatment received since January 1, 2010. The patient characteristics at the first visit following January 1, 2010, are also summarized in Table 1 for 522 patients who were included in the following multivariate regression analyses. The mean age was 48 years, and 58% were male. Ten percent had concomitant fibromyalgia.

The results of the analysis involving all courses of biologic therapy (Table 2) showed that a higher number of tender joints was associated with a higher risk of discontinuation due to all causes (relative risk [RR] 1.02 [95% CI 1.00–1.03], $P = 0.01$) or side effects (RR 1.03 [95% CI 1.00–1.06], $P = 0.03$). Older age (in years) at the start of a treatment was associated with an increased risk of discontinuation due to side effects (RR 1.03 [95% CI 1.01–1.05], $P = 0.01$), while obesity had a protective role (RR 0.56 [95% CI 0.31–1.00], $P = 0.05$). Having a college degree or higher lowered the risk of discontinuation due to any other reasons (RR 0.56 [95% CI 0.33–0.96], $P = 0.03$) but not due to side effects or lack of efficacy. Sex did not affect drug discontinuation. When examining the risk of drug discontinuation by each type of treatment (Table 2), a higher number of swollen joints reduced the risk of discontinuation of adalimumab, whereas having depression and/or anxiety increases the risk of discontinuation of etanercept. The risk of discontinuation of IL-17i was associated with the number of actively inflamed joints and whether or not the patient was obese. No significant covariates were found when analyses were repeated for certolizumab, golimumab, and infliximab separately.

The analysis of time to discontinuation of the first treatment showed that depression and/or anxiety had a higher risk of drug discontinuation due to all causes (RR 1.68 [95% CI 1.11–2.53], $P = 0.01$), while having a college degree or higher had a lower risk (RR 0.65 [95% CI 0.44–0.96], $P = 0.03$) (Table 3). In the analysis of time to discontinuation of the second treatment, male patients were less likely to discontinue a second treatment due to side effects (RR 0.12 [95% CI 0.02–0.69], $P = 0.02$), while obese patients were more likely to stay on the second treatment due to all causes (RR 0.64 [95% CI 0.42–1.00], $P = 0.05$) (Table 4).

DISCUSSION

In this retrospective analysis of prospectively collected data, we describe the pattern of biologic use among PsA patients in a large cohort over 20 years. TNFi, mainly etanercept and adalimumab, were the most prescribed biologics, which is in line with the availability of various biologics and their approval date in Canada. Recommendations concerning which class of biologic to initiate vary and are routinely updated. While the 2015 EULAR recommendations preferred the initiation of a TNFi (21), the updated 2019 version gives the choice according to the most symptomatic domain (4). On the other hand, the American College of Rheumatology/National Psoriasis Foundation 2018 guidelines recommended switching from a first TNFi to a second TNFi prior to switching to a different class (5). The Group for Research and Assessment of Psoriasis and Psoriatic Arthritis suggests that the choice among all potential biologics approved for PsA patients depends on the domain involved (6). These recommendations, which aim to guide the physicians in their decisions, have limitations, especially when dealing with a heterogeneous disease such as PsA (22). Recent head-to-head studies of IL-17i versus adalimumab suggest that the effect on musculoskeletal manifestations is similar, while the effect on the skin favors IL-17i (23,24).

Table 4. Overall and cause-specific Cox regression analysis for the time to discontinuation of the second biologic treatment*

	All causes	P	Side effects	P	Other reasons	P
Age at start of treatment, years	–		1.08 (1.02–1.14)	0.005	–	
Sex, male versus female	–		0.12 (0.02–0.69)	0.018	–	
Married, yes versus no	–		–		0.38 (0.16–0.91)	0.030
No. of tender joints	1.04 (1.01–1.07)	0.019	1.08 (1.01–1.15)	0.029	–	
Obese versus not obese	0.64 (0.42–1.00)	0.048	–		–	

* Values are the relative risk (95% confidence interval) unless indicated otherwise. Relative risks are based on mixed-effects Cox-type models to account for the dependence between the duration of the first and second course of biologics.

The most commonly prescribed biologics are TNFi, which is consistent with their introduction to the Canadian market, followed by IL-17i. Etanercept was the most prescribed biologic in our cohort, unlike in other cohorts, such as the nationwide Danish rheumatology database DANBIO, where it was uncommonly used, reflecting the perception that it might not be effective in treating skin disease (9,25). As expected, if a patient fails a TNFi, they may have difficulty continuing to take another drug from the same class when used as a second-line treatment. This was demonstrated in studies by Costa et al (26) and Reddy et al (27). When used as a second-line treatment, IL-17i had better persistence compared to when it was used as a first-line treatment, but persistence taking IL-17i was still inferior to the performance of TNFi. However, in a recent study conducted on a large cohort of PsA patients by Haddad et al (16), a higher persistency rate was observed for secukinumab when indicated as second-line therapy compared to adalimumab, infliximab, and ustekinumab.

Factors associated with better drug persistence in our study include having higher education and obesity, mainly for IL-17i. The latter is inconsistent with the literature, where obesity was often found to negatively affect response to TNFi (28,29). On the other hand, older age, higher number of affected joints, and the presence of anxiety and depression were found to decrease drug survival in the analysis involving all courses of biologic therapy, while sex was not statistically significant. These findings are expected, as these comorbidities indicate a more severe disease that is unlikely to be treated easily. We have previously demonstrated that obesity reduced the probability of achieving minimal disease activity in patients with PsA (30). In a recent study by Ogdie et al (31), achieving remission was positively associated with college education, which is consistent with our results. However, this study also identified that obesity, hypertension, female sex, and previous biologic use were negative predictors (31). While it is not clear why our study demonstrated that obesity was associated with persistence in taking a drug, it is possible that these patients were afraid to stop taking the drug out of fear that other drugs may not work.

Our analysis did not account for the concomitant use of methotrexate (MTX) and other conventional DMARDs. However, the combination of MTX and TNFi was not found to improve efficacy in PsA patients in the Study of Etanercept and Methotrexate (SEAM) study, which compared the effect of etanercept monotherapy to MTX monotherapy or etanercept combined with MTX (32). Similarly, combining IL-17i secukinumab and ixekizumab with MTX was not found to have better efficacy than IL-17i monotherapy (23,33). Other studies, on the other hand, found better persistence of biologics when combined with MTX (16).

This study has a major strength, as it included biologics other than TNFi and examined predictors of drug persistence. However, we acknowledge a few limitations, as we did not account for concomitant use of other conventional or targeted synthetic DMARDs. Also, ustekinumab, an IL-12/IL-23i, and guselkumab (IL-23i) were included in the data collection but not in the analysis due to the limited number of patients treated with these drugs. As this is an observational study, drug prescriptions were initiated by the treating physicians and were not dependent on specific guidelines.

In conclusion, in this retrospective analysis of this large prospective cohort, persistence of biologics in the treatment of PsA was examined. Certolizumab seemed to have the highest drug survival when used as a first-line treatment, while IL-17i had the lowest probability, although data were very sparse for IL-17i biologics. Our results were mostly consistent with the previously published literature, with few exceptions where obesity was found to be protective against drug discontinuation. While IL-17i were found to perform better as a second-line treatment in previous studies, our study did not show a difference in their performance whether prescribed as first or second biologic. Further studies and standardized protocols are needed to compare various cohorts, especially in a heterogeneous disease like PsA.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Rida, Lee, Chandran, Cook, Gladman.
Acquisition of data. Rida, Chandran, Gladman.
Analysis and interpretation of data. Rida, Lee, Chandran, Cook, Gladman.

REFERENCES

1. Rida MA, Chandran V. Challenges in the clinical diagnosis of psoriatic arthritis. Clin Immunol 2020;214:108390.
2. Landewe R, Ritchlin CT, Aletaha D, et al. Inhibition of radiographic progression in psoriatic arthritis by adalimumab independent of the control of clinical disease activity. Rheumatology (Oxford) 2019;58:1025–33.
3. So A, Inman RD. An overview of biologic disease-modifying antirheumatic drugs in axial spondyloarthritis and psoriatic arthritis. Best Pract Res Clin Rheumatol 2018;32:453–71.
4. Gossec L, Baraliakos X, Kerschbaumer A, et al. EULAR recommendations for the management of psoriatic arthritis with pharmacological therapies: 2019 update. Ann Rheum Dis 2020;79:700–12.
5. Singh JA, Guyatt G, Ogdie A, et al. 2018 American College of Rheumatology/National Psoriasis Foundation guideline for the treatment of psoriatic arthritis. Arthritis Care Res (Hoboken) 2019;71:2–29.
6. Coates LC, Kavanaugh A, Mease PJ, et al. Group for Research and Assessment of Psoriasis and Psoriatic Arthritis 2015 treatment recommendations for psoriatic arthritis. Arthritis Rheumatol 2016;68:1060–71.
7. Merola JF, Lockshin B, Mody EA. Switching biologics in the treatment of psoriatic arthritis. Semin Arthritis Rheum 2017;47:29–37.
8. Gladman DD, Antoni C, Mease P, et al. Psoriatic arthritis: epidemiology, clinical features, course, and outcome. Ann Rheum Dis 2005;64 Suppl 2:ii14–7.

9. Glintborg B, Ostergaard M, Dreyer L, et al. Treatment response, drug survival, and predictors thereof in 764 patients with psoriatic arthritis treated with anti-tumor necrosis factor alpha therapy: results from the nationwide Danish DANBIO registry. Arthritis Rheum 2011;63: 382–90.
10. Fagerli KM, Lie E, van der Heijde D, et al. The role of methotrexate comedication in TNF-inhibitor treatment in patients with psoriatic arthritis: results from 440 patients included in the NOR-DMARD study. Ann Rheum Dis 2014;73:132–7.
11. Kristensen LE, Gulfe A, Saxne T, et al. Efficacy and tolerability of antitumour necrosis factor therapy in psoriatic arthritis patients: results from the South Swedish Arthritis Treatment Group register. Ann Rheum Dis 2008;67:364–9.
12. Virkki LM, Sumathikutty BC, Aarnio M, et al. Biological therapy for psoriatic arthritis in clinical practice: outcomes up to 2 years. J Rheumatol 2010;37:2362–8.
13. Mease PJ, Karki C, Liu M,et al. Discontinuation and switching patterns of tumour necrosis factor inhibitors (TNFis) in TNFi-naive and TNFi-experienced patients with psoriatic arthritis: an observational study from the US-based Corrona registry. RMD Open 2019;5:e000880.
14. Oelke KR, Chambenoit O, Majjhoo AQ, et al. Persistence and adherence of biologics in US patients with psoriatic arthritis: analyses from a claims database. J Comp Eff Res 2019;8:607–21.
15. Yamamoto T, Ohtsuki M, Sano S, et al. Switching biologics in the treatment of psoriatic arthritis in Japan. J Dermatol 2019;46:e113–4.
16. Haddad A, Gazitt T, Feldhamer I, et al. Treatment persistence of biologics among patients with psoriatic arthritis. Arthritis Res Ther 2021; 23:44.
17. Fagerli KM, Kearsley-Fleet L, Watson KD, et al. Long-term persistence of TNF-inhibitor treatment in patients with psoriatic arthritis: data from the British Society for Rheumatology Biologics register. RMD Open 2018;4:e000596.
18. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Amer Statistical Assoc 1958:457–81.
19. Cook RJ, Lawless JF. The statistical analysis of recurrent events. New York (NY): Springer; 2007. p. 403.
20. R Core Team R: a language and environment for statistical computing. Vienna Austria: R Foundation for Statistical Computing; 2019. URL: https://www.R-project.org/.
21. Gossec L, Smolen JS, Ramiro S, et al. European League Against Rheumatism (EULAR) recommendations for the management of psoriatic arthritis with pharmacological therapies: 2015 update. Ann Rheum Dis 2016;75:499–510.
22. Ogdie A, Coates LC, Gladman DD. Treatment guidelines in psoriatic arthritis. Rheumatology (Oxford) 2020;59 Suppl 1:i37–46.
23. Mease PJ, Smolen JS, Behrens F, et al. A head-to-head comparison of the efficacy and safety of ixekizumab and adalimumab in biological-naive patients with active psoriatic arthritis: 24-week results of a randomised, open-label, blinded-assessor trial. Ann Rheum Dis 2020;79:123–31.
24. McInnes IB, Behrens F, Mease PJ, et al. Secukinumab versus adalimumab for treatment of active psoriatic arthritis (EXCEED): a double-blind, parallel-group, randomised, active-controlled, phase 3b trial. Lancet 2020;395:1496–505.
25. Ash Z, Gaujoux-Viala C, Gossec L, et al. A systematic literature review of drug therapies for the treatment of psoriatic arthritis: current evidence and meta-analysis informing the EULAR recommendations for the management of psoriatic arthritis. Ann Rheum Dis 2012;71: 319–26.
26. Costa L, Perricone C, Chimenti MS, et al. Switching between biological treatments in psoriatic arthritis: a review of the evidence. Drugs R D 2017;17:509–22.
27. Reddy SM, Crean S, Martin AL, et al. Real-world effectiveness of anti-TNF switching in psoriatic arthritis: a systematic review of the literature. Clin Rheumatol 2016;35:2955–66.
28. Viguier M, Livideanu C, Beylot-Barry M, et al. Observational case series on a group of psoriasis patients who failed to respond to any TNF blockers. J Dermatolog Treat 2014;25:75–7.
29. Di Lernia V, Ricci C, Lallas A, et al. Clinical predictors of non-response to any tumor necrosis factor (TNF) blockers: a retrospective study. J Dermatolog Treat 2014;25:73–4.
30. Eder L, Thavaneswaran A, Chandran V, et al. Obesity is associated with a lower probability of achieving sustained minimal disease activity state among patients with psoriatic arthritis. Ann Rheum Dis 2015;74: 813–7.
31. Ogdie A, Palmer JL, Greenberg J, et al. Predictors of achieving remission among patients with psoriatic arthritis initiating a tumor necrosis factor inhibitor. J Rheumatol 2019;46:475–82.
32. Mease PJ, Gladman DD, Collier DH, et al. Etanercept and methotrexate as monotherapy or in combination for psoriatic arthritis: primary results from a randomized, controlled phase III trial. Arthritis Rheumatol 2019;71:1112–24.
33. McInnes IB, Mease PJ, Ritchlin CT, et al. Secukinumab sustains improvement in signs and symptoms of psoriatic arthritis: 2 year results from the phase 3 FUTURE 2 study. Rheumatology (Oxford) 2017;56:1993–2003.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 75, No. 10, October 2023, pp 2182–2189
DOI 10.1002/acr.25111

BRIEF REPORT

Responsiveness and Minimum Clinically Important Difference in Patient-Reported Outcome Measures Among Patients With Psoriatic Arthritis: A Prospective Cohort Study

Paras Karmacharya,[1] Courtney Stull,[2] Alisa Stephens-Shields,[3] M. Elaine Husni,[4] Jose U. Scher,[5] Ethan Craig,[3] Robert Fitzsimmons,[3] Soumya M. Reddy,[5] Marina N. Magrey,[6] Alexis Ogdie,[3] and Jessica A. Walsh[7]

Objective. To determine the responsiveness to therapy and minimum clinically important improvement (MCII) for patient-reported outcome measures in psoriatic arthritis (PsA) and to examine the impact of baseline disease activity on the ability to demonstrate change.

Methods. A longitudinal cohort study was performed within the PsA Research Consortium. Patients completed several patient-reported outcomes, including the Routine Assessment of Patient Index Data, the Bath Ankylosing Spondylitis Disease Activity Index (BASDAI), the Psoriatic Arthritis Impact of Disease 12-item (PsAID12) questionnaire, and others. The mean change in the scores between visits and standardized response means (SRMs) were calculated. The MCII was calculated as the mean change in score among patients who reported minimal improvement. SRMs and MCIIs were compared among subgroups with moderate to highly active PsA and those with lower disease activity.

Results. Among 171 patients, 266 therapy courses were included. The mean ± SD age was 51 ± 13.8 years, 53% were female, and the mean swollen and tender joint counts were 3 and 6, respectively, at baseline. SRMs and MCII for all measures were small to moderate, although greater among those with higher baseline disease activity. BASDAI had the best SRM overall and for less active PsA, and the clinical Disease Activity of PsA (cDAPSA) and PsAID12 were best for those with higher disease activity.

Conclusion. SRMs and MCII were relatively small in this real-world population, particularly among those with lower disease activity at baseline. BASDAI, cDAPSA, and PsAID12 had good sensitivity to change, but selection for use in trials should consider the baseline disease activity of patients to be enrolled.

INTRODUCTION

Psoriatic arthritis (PsA) is a highly heterogeneous inflammatory disease affecting the skin and musculoskeletal system, posing a challenge for treatment selection and measurement of disease activity (1). Despite this disease heterogeneity, randomized controlled trials (RCTs) that test new medications generally enroll a homogenous subgroup of patients with PsA (polyarticular similar to rheumatoid arthritis). Findings from the RCTs are not necessarily generalizable, since the majority of patients in real-world clinical practice have oligoarticular disease. To improve our understanding of how therapies work in the real world, trials based on real-world patients are needed (2).

Presented in part at the Annual European Congress of Rheumatology of the European Alliance of Associations for Rheumatology, 2020, and at the American College of Rheumatology Annual Scientific Meeting, Philadelphia, Pennsylvania, November 2022.

Supported by the NIH/National Institute of Arthritis and Musculoskeletal and Skin Diseases (grant R01-AR-072363). Dr. Karmacharya's work was supported by the Rheumatology Research Foundation (Scientist Development award), the Spondyloarthritis Research and Treatment Network (Fellowship Pilot grant), and the Group for Research, Assessment of Psoriasis and Psoriatic Arthritis (Pilot Research grant).

[1]Paras Karmacharya, MD, MS: Vanderbilt University Medical Center, Nashville, Tennessee; [2]Courtney Stull, MD: University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania; [3]Alisa Stephens-Shields, PhD, Ethan Craig, MD, MHS, Robert Fitzsimmons, MS, Alexis Ogdie, MD, MSCE: University of Pennsylvania, Philadelphia; [4]M. Elaine Husni, MD, MPH: Cleveland Clinic, Cleveland, Ohio; [5]Jose U. Scher, MD, Soumya M. Reddy, MD: New York University, New York; [6]Marina N. Magrey, MD: University Hospitals Cleveland Medical Center, Cleveland, Ohio; [7]Jessica A. Walsh, MD: University of Utah, Salt Lake City.

Drs. Karmacharya and Stull are both first authors and Drs. Ogdie and Walsh are both last authors.

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.25111&file=acr25111-sup-0001-Disclosureform.pdf.

Address correspondence via email to Alexis Ogdie, MD, MSCE, at ogdiea@pennmedicine.upenn.edu.

Submitted for publication November 10, 2022; accepted in revised form March 9, 2023.

SIGNIFICANCE & INNOVATIONS

- With a wide variety of available patient-reported outcomes and physician-based assessments for psoriatic arthritis (PsA), there are few data comparing outcome measures to assist in identifying which to follow in the clinic or pragmatic trials.
- The Bath Ankylosing Spondylitis Disease Activity Index and the Patient-Reported Outcomes Measurement Information System depression and fatigue indices demonstrated the highest responsiveness to change. The clinical Disease Activity of PsA and PsA Impact of Disease performed better in the moderate to highly active PsA subgroup.
- The selection of measures should take into account the thresholds for meaningful improvement for the measure (minimum clinically important improvement) and projected baseline disease activity of PsA patients enrolled.

Before conducting real-world, pragmatic studies in PsA, defining appropriate outcome measures is necessary. Patient-reported outcomes are critical for understanding a patient's response to therapy in clinical practice, and depending on the trial objectives, may serve as primary outcomes for pragmatic trials (3). However, selecting patient-reported outcome measures for use in clinical practice or pragmatic trials is challenging, as limited data exist to understand how each patient-reported outcome is expected to change in routine care. Most data on the responsiveness of patient-reported outcomes are derived from clinical trials where patients have significantly greater disease activity. Which patient-reported outcomes have the psychometric properties that are required to serve as a primary outcome for pragmatic trials in PsA is not yet clear. The objectives of this study were to determine the responsiveness to therapy and minimum clinically important improvement (MCII) for candidate outcome measures in PsA and to examine the impact of baseline disease activity on the ability to demonstrate change in a real-world setting.

PATIENTS AND METHODS

Study design and patient population. PsA patients were enrolled in the PsA Research Consortium, a longitudinal observational cohort in PsA-dedicated clinics at 4 US institutions (2017–2020) (4). Inclusion criteria for the study were patients who met the Classification Criteria for PsA criteria, who were starting or switching therapy, and who had a follow-up visit with global assessment of response documented after therapy initiation or switching. Patients could contribute >1 therapy course for the primary analysis. In a sensitivity analysis, only the first treatment initiation was included.

Assessments. Patient-reported outcomes were selected to include the range of domains important to patients (5). Surveys included: the Routine Assessment of Patient Index Data (RAPID3; range 0–30), which includes the Multidimensional Health Assessment Questionnaire (range 0–10); the Health Assessment Questionnaire disability index (HAQ DI; range 0–3); the Bath Ankylosing Spondylitis Disease Activity Index (BASDAI; range 0–10); the PsA Impact of Disease 12-item questionnaire (PsAID12; range 0–10); the Patient-Reported Outcomes Measurement Information System (PROMIS) global health 10a short form physical health and mental health subscores, each with T score range of 0–100; the PROMIS fatigue 8a short form (range 0–100); the PROMIS depression 8a short form; patient pain assessment (range 0–100); and patient global assessment (range 0–10). The clinical Disease Activity of PsA (cDAPSA; range 0–154) and minimum disease activity criteria were calculated. The index date was the visit at which therapy was initiated or changed.

Follow-up visits occurred at approximately 16 weeks after therapy initiation (although we included patients with a longer follow-up interval, often up to 4 weeks delay). At follow-up, patients completed a global assessment of response and rated their status as improved, stayed the same, or worsened and rated their level of improvement or worsening as described by Ward et al (6). If patients improved, they were further asked to rate the importance of their change on a scale of 0 to 6 (0 = almost none or hardly at all, 1 = a little important, 2 = somewhat important, 3 = moderately important, 4 = a good deal important, 5 = very important, and 6 = extremely important). MCII was defined as having a score of ≥1. This global assessment of response was used as an external anchor in these analyses (reference standard). The scoring of the instruments has been previously described in detail (4,5). For PROMIS scores, T scores have a range of 0–100, and 50 represents the average person in the general population. Higher scores (i.e., above 50) represent more of the construct (i.e., depression, fatigue), and lower scores represent less of the construct.

Statistical analysis. Baseline patient characteristics and mean change in each measure were descriptively reported. Standardized response means (SRMs) were calculated among all patients initiating therapy (as therapy initiators are expected to improve). SRMs are interpreted as large if SRM >0.8, moderate if SRM = 0.5–0.8, and as low if SRM <0.5 (7). We hypothesized that tumor necrosis factor inhibitor (TNFi) and interleukin-17 inhibitor (IL-17i) would have higher SRMs than oral small molecule initiators because biologics in general have shown higher response rates compared to oral small molecules such as methotrexate or apremilast (8). We calculated the overall SRM as well as the SRM for the subgroup initiating TNFi or IL-17i for each measure. The mean change in each measure was reported separately for patients answering improved, stayed the same, or worsened on

their global assessment of response. To test discrimination, *t*-tests were used to determine whether the change in each measure statistically differed among patients reporting improvement and worsening compared to patients reporting stayed the same. We calculated the MCII for the measures using a receiver operating characteristic (ROC) curve with minimally important change on the patient global assessment of response as the reference standard. The criterion was the change in the measure that corresponded to a point on the ROC curve with minimal distance to 0,1 maximizing specificity and sensitivity simultaneously (nearest to [0,1] method) (9). Confidence intervals (CIs) for the estimates were calculated based on 2,000 bootstrapped samples. The area under the ROC curve (AUC) or accuracy was used to assess the discrimination of the measures. The misclassification rate was defined as the percentage of times a particular measure incorrectly classified the patients. Alternative methods for determining the optimal cutoff, maximizing the sum of sensitivity and specificity (the Youden method), and maximizing the product of sensitivity and specificity (the Liu method), showed similar results. Sensitivity analysis examined MCII for the subgroups with moderate to highly active PsA (≥3 swollen and ≥3 tender joint counts based on 66- and 68-joint counts, respectively) and those with less active PsA (<3 swollen and <3 tender joint counts). This cutoff is used as an inclusion criterion in most PsA RCTs (10). STATA version 16.1 was used for the analyses.

Table 1. Baseline characteristics of treatment instances (n = 266)*

Variable	No.	Value
Age, years	266	51.0 ± 13.8
Female, no. (%)	262	140 (53.4)
Race, no. (%)	266	–
White	–	242 (91.0)
African American	–	1 (0.4)
Alaska Native/American Indian	–	4 (1.5)
Asian	–	1 (0.4)
Native Hawaiian/Pacific Islander	–	0 (0)
Other	–	11 (4.1)
Ethnicity, no. (%)	266	–
Not Hispanic/Latino	–	229 (86.1)
Hispanic/Latino	–	20 (7.5)
Other	–	23 (8.7)
Body mass index, kg/m^2	232	29.9 ± 6.3
Tender joint count (0–68)	264	5.8 ± 7.4
Swollen joint count (0–66)	264	2.9 ± 4.9
Body surface area %	258	1.6 ± 3.9
Enthesitis count	260	0.6 ± 1.1
Dactylitis count	259	0.3 ± 1.2
Axial involvement, no. (%)	255	41 (16.1)
CRP elevated, no. (%)	186	52 (28.0)
Minimum disease activity, no. (%)	175	63 (36.0)
Therapies initiated, no. (%)†	–	–
TNFi	–	145 (54.5)
IL-17i	–	55 (20.7)
OSM	–	96 (36.1)
Other biologics	–	14 (5.3)

* Values are the mean ± SD unless indicated otherwise. Body surface area = body surface area affected by psoriasis; CRP = C-reactive protein; IL-17i = interleukin-17 inhibitor; OSM = oral small molecule; TNFi = tumor necrosis factor inhibitor.
† Patients could initiate >1 therapy (e.g., methotrexate and a TNFi).

RESULTS

Among 171 unique patients, 266 treatment instances met the inclusion criteria. The mean age of the participants was 51 years, 53% were female, and the majority were White (91%). The mean body mass index was 29.9, and 16% had axial disease (fulfilling the Assessment of SpondyloArthritis international Society [ASAS] axial spondyloarthritis [SpA] classification criteria) (11). Among the 266 visits, TNFi, IL-17i, oral small molecules, and another biologic or JAK inhibitor were initiated in 145, 55, 96, and 14 visits, respectively. Mean swollen and tender joint counts (66 swollen and 68 tender joints) at therapy initiation were 2.9 and 5.8, respectively (Table 1). At the time of therapy initiation/change, 83 patients had active PsA (≥3 swollen and ≥3 tender joint counts) and 183 patients had less active PsA.

Responsiveness for all measures was small to moderate (Figure 1). Overall, BASDAI was the most responsive (the only measure with moderate SRM), followed by PROMIS fatigue, PROMIS depression, physician global joint assessment, PsAID12, and patient pain assessment (low SRMs). BASDAI specifically performed well in the less active PsA subgroup, where other measures such as cDAPSA and PsAID12 performed poorly. For the PsA subgroup with moderate to highly active PsA, however, cDAPSA showed the highest SRM, followed by BASDAI and PsAID12, with similar SRMs. In general, SRMs were larger among those initiating a biologic therapy (specifically a TNFi or IL-17i) compared to any therapy.

At follow-up, 77 patients (29%) rated themselves as improved, 115 (43%) reported that they had stayed the same, and 74 (28%) reported that they had worsened. The mean changes in each measure by patient-reported response (improved, stayed the same, or worsened) are shown in Table 2. In general, the mean score for the measures increased (worsened) in patients who improved to patients who worsened, as expected (with the exception of PROMIS physical health, which decreased or worsened given the different scoring) (see Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25111). The MCII for the measures were as follows: RAPID3 –1.42 (95% CI 0.46, –3.30), BASDAI –0.70 (95% CI 0.04, –1.44), PsAID12 –0.65 (95% CI –0.41, –0.99), and cDAPSA –3.25 (95% CI –1.61, –4.89). The best accuracy (AUC) for MCII was noted for BASDAI (0.83 [95% CI 0.74, 0.91]), followed by PsAID12 (0.81 [95% CI 0.74, 0.87]), patient pain (0.79 [95% CI 0.73, 0.86]), RAPID3 (0.78 [95% CI 0.72, 0.85]), and cDAPSA (0.77 [95% CI 0.70, 0.84]). The proportion misclassified was higher for the PROMIS scores in general. Among the whole cohort of PsA patients who initiated/changed therapy, the optimal cutoff for MCII was low for most measures, and the lowest

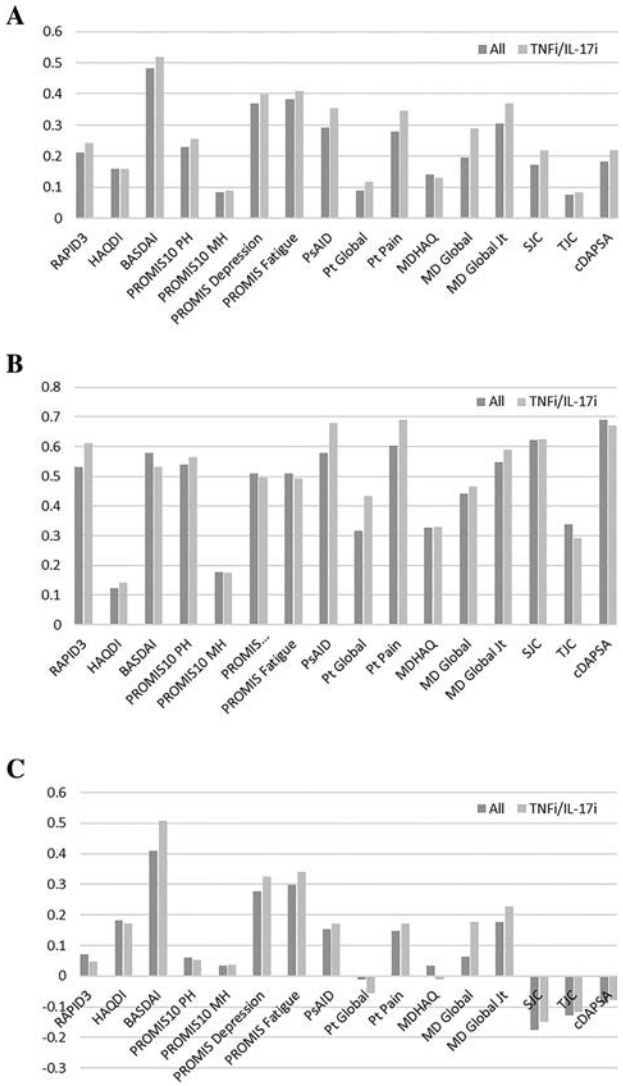

Figure 1. Standard response means of instruments tested in the Psoriatic Arthritis Research Consortium: **A**, Overall psoriatic arthritis (PsA); **B**, Moderate to highly active PsA; **C**, Less active PsA. BASDAI = Bath Ankylosing Spondylitis Disease Activity Index; cDAPSA = clinical Disease Activity of Psoriatic Arthritis; HAQDI = Health Assessment Questionnaire disability index; IL-17i = interleukin-17 inhibitor; Jt = joint; MD global = physician's global assessment; MDHAQ = Multidimensional Health Assessment Questionnaire; MH = mental health; PH = physical health; PROMIS = Patient-Reported Outcomes Measurement Information System global short form; PsAID = Psoriatic Arthritis Impact of Disease; Pt = patient; RAPID3 = Routine Assessment of Patient Index Data; SJC = swollen joint count; TNFi = tumor necrosis factor inhibitor; TJC = tender joint count.

misclassification rates were noted for PsAID12, BASDAI, patient pain, cDAPSA, and RAPID3 (Table 3). A subgroup analysis of PsA patients with moderate to highly active disease at therapy initiation/change showed higher MCIIs compared to those with less active disease (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25111): RAPID3 score of –3.25 (95% CI –1.44, –5.06) versus –1.42 (95% CI –0.24, –2.59), BASDAI –1.55 (95% CI –0.55, –2.55) versus –0.80 (95% CI 0.11, –1.71), PsAID12 –1.18 (95% CI –0.88, –1.47) versus –0.70 (95% CI –0.50, –0.91), and cDAPSA –9.25 (95% CI –5.65, –12.85) versus –2.25 (95% CI 0.41, –4.09). Similarly, sensitivity and specificity were much higher for the active PsA subgroup compared to the less active subgroup with a lower proportion of patients misclassified. MCII calculated by alternative methods showed similar results in most cases.

DISCUSSION

In this real-world PsA population, the most responsive measures were BASDAI, PROMIS fatigue, PROMIS depression, physician global joint assessment, PsAID12, and patient pain assessment. Tender and swollen joint counts had low responsiveness in this patient population. Overall, SRMs and MCII for physician-assessed and patient-reported outcome measure instruments were relatively low compared to what has been reported in RCTs, even in the subset of PsA patients initiating

Table 2. Baseline patient-reported outcomes, mean change, and discrimination among patients who completed global assessment of response*

		Baseline		All		Improved		Stayed the same		Worsened		*P* for discrimination‡	
Measure	Type	No.†	Mean ± SD	No.	Change ± SD	No.	Change ± SD	No.	Change ± SD	No.	Change ± SD	Improved vs. same	Worsened vs. same
RAPID3	PRO	229	10.93 ± 5.88	225	−1.11 ± 5.26	65	−4.77 ± 5.05	99	−0.95 ± 3.54	61	2.53 ± 5.24	<0.001	<0.001
HAQ DI	PRO	141	0.68 ± 0.56	85	−0.07 ± 0.42	35	−0.26 ± 0.33	30	−0.017 ± 0.35	20	0.20 ± 0.50	0.005	0.075
BASDAI	PRO	152	4.70 ± 2.24	95	−0.92 ± 1.91	38	−2.19 ± 1.82	32	−0.54 ± 1.15	25	0.53 ± 1.60	<0.001	0.005
PROMIS10 PH	PRO	187	42.01 ± 8.20	183	1.43 ± 6.20	58	4.96 ± 7.14	80	0.68 ± 4.19	45	−1.78 ± 5.78	<0.001	0.007
PROMIS10 MH	PRO	187	46.28 ± 10.76	184	0.54 ± 6.43	58	2.98 ± 6.97	80	−0.11 ± 5.86	46	−1.41 ± 5.82	0.006	0.232
PROMIS depression	PRO	140	60.78 ± 9.59	85	−2.68 ± 7.27	35	−6.14 ± 6.67	30	0.20 ± 6.67	20	−0.97 ± 6.95	<0.001	0.553
PROMIS fatigue	PRO	140	56.45 ± 9.93	85	−2.86 ± 7.67	35	−6.40 ± 6.87	30	0.04 ± 6.99	20	−1.01 ± 6.86	<0.001	0.604
PsAID12	PRO	231	2.00 ± 1.34	225	−0.30 ± 1.05	66	−0.97 ± 1.05	99	−0.43 ± 1.28	60	0.30 ± 1.05	<0.001	<0.001
Patient global	PRO	234	4.11 ± 2.48	234	−0.24 ± 2.66	68	−1.52 ± 2.74	103	−0.16 ± 2.32	63	1.03 ± 2.48	<0.001	0.002
Patient pain	PRO	234	4.81 ± 2.69	234	−0.75 ± 2.66	68	−2.63 ± 2.73	103	−0.62 ± 1.75	63	1.09 ± 2.50	<0.001	<0.001
MDHAQ	PRO	229	2.05 ± 1.56	225	−0.17 ± 1.24	65	−0.74 ± 1.22	99	−0.20 ± 0.87	61	0.46 ± 1.47	0.001	<0.001
Physician global	MD	257	4.54 ± 2.06	249	−0.30 ± 1.54	72	−1.24 ± 1.68	107	−0.16 ± 1.14	70	0.44 ± 1.45	<0.001	0.002
Physician global joint	MD	258	3.33 ± 2.15	250	−0.55 ± 1.80	72	−1.56 ± 1.91	107	−0.51 ± 1.35	71	0.41 ± 1.75	<0.001	<0.001
SJC	MD	264	2.90 ± 4.85	256	−0.81 ± 4.71	74	−2.92 ± 5.95	110	−0.68 ± 3.10	72	1.15 ± 4.45	0.001	0.001
TJC	MD	264	5.82 ± 7.38	256	−0.63 ± 8.35	75	−2.84 ± 9.82	110	−1.25 ± 7.25	71	2.66 ± 7.33	0.206	<0.001
cDAPSA	Both	234	17.29 ± 12.87	225	−2.26 ± 12.36	65	−9.45 ± 11.90	99	−2.47 ± 9.62	61	5.72 ± 12.16	<0.001	<0.001

* BASDAI = Bath Ankylosing Spondylitis Disease Activity Index; cDAPSA = clinical Disease Activity of Psoriatic Arthritis; HAQ DI = Health Assessment Questionnaire disability index; MDHAQ = physician's global assessment; MH = mental health; PH = physical health; PRO = patient-reported outcome; PROMIS10 = Patient-Reported Outcomes Measurement Information System global short form; PsAID12 = Psoriatic Arthritis Impact of Disease 12-item questionnaire; RAPID3 = Routine Assessment of Patient Index Data; SJC = swollen joint count; TJC = tender joint count.

† The number varies by measure because certain measures were added at different time points in the protocol (i.e., HAQ DI was added in year 2 of the study).

‡ Discrimination refers to whether there was a statistically significant difference in the mean change in each measure among the 2 groups compared.

Table 3. Minimum clinically important improvement (MCII) estimates for different measures using minimally important change on the patient global assessment of response as the standard, and the point on the ROC curve closest to (0,1) as the criterion*

Measure	Range	No.†	no.‡	AUC of the measure for MCII	MCII	Sensitivity	Specificity	% misclassified
RAPID3	0–30	225	62	0.78 (0.72, 0.85)	−1.42 (0.46, −3.30)	0.79 (0.66, 0.93)	0.68 (0 .54, 0.82)	27
HAQ DI	0–3	85	35	0.74 (0.64, 0.85)	−0.06 (0.09, −0.21)	0.69 (0.47, 0.90)	0.66 (0.47, 0.85)	33
BASDAI	0–10	95	38	0.83 (0.74, 0.91)	−0.70 (0.04, −1.44)	0.82 (0.67, 0.96)	0.68 (0.53, 0.84)	25
PROMIS10 PH	0–100	183	56	0.73 (0.64, 0.81)	2.45 (1.53, 3.37)	0.68 (0.55, 0.81)	0.69 (0.57, 0.80)	32
PROMIS10 MH	0–100	184	56	0.64 (0.56, 0.73)	1.15 (−1.87, 4.17)	0.57 (0.37, 0.77)	0.61 (0.39, 0.83)	41
PROMIS depression	0–100	85	35	0.71 (0.61, 0.82)	−1.05 (1.77, −3.87)	0.74 (0.56, 0.93)	0.58 (0.41, 0.76)	34
PROMIS fatigue	0–100	85	35	0.72 (0.61, 0.83)	−1.10 (2.36, −4.56)	0.74 (0.55, 0.93)	0.58 (0.40, 0.76)	34
PsAID12	0–10	225	63	0.81 (0.74, 0.87)	−0.70 (−0.41, −0.99)	0.79 (0.70, 0.89)	0.74 (0.65, 0.82)	23
Patient global	0–10	234	65	0.73 (0.65, 0.81)	−1.25 (−0.53, −1.97)	0.63 (0.51, 0.75)	0.82 (0.68, 0.96)	28
Patient pain	0–100	234	65	0.79 (0.73, 0.86)	−1.25 (−0.91, −1.59)	0.72 (0.62, 0.83)	0.76 (0.68, 0.85)	26
MDHAQ	0–10	225	62	0.67 (0.59, 0.75)	−0.33 (0.05, −0.72)	0.58 (0.45, 0.71)	0.68 (0.57, 0.78)	37
Physician global	0–10	249	69	0.73 (0.66, 0.80)	−0.50 (−0.30, −0.70)	0.55 (0.43, 0.67)	0.82 (0.76, 0.88)	32
Physician global joint	0–10	250	69	0.72 (0.65, 0.79)	−0.50 (−0.05, −0.95)	0.61 (0.49, 0.73)	0.75 (0.67, 0.84)	32
SJC	0–66	256	71	0.73 (0.66, 0.80)	−0.50 (−0.05, −0.96)	0.69 (0.57, 0.81)	0.68 (0.59, 0.77)	31
TJC	0–68	256	72	0.68 (0.61, 0.75)	−0.50 (0.17, −1.17)	0.69 (0.58, 0.81)	0.65 (0.56, 0.73)	33
cDAPSA	0–154	225	62	0.77 (0.70, 0.84)	−3.25 (−1.61, −4.89)	0.76 (0.65, 0.87)	0.71 (0.63, 0.80)	26

* Values are the measurement specified (95% confidence interval) unless indicated otherwise. Values are based on 2,000 bootstrapped samples. AUC = area under the curve; BASDAI = Bath Ankylosing Spondylitis Disease Activity Index; cDAPSA = clinical Disease Activity of Psoriatic Arthritis; HAQ DI = Health Assessment Questionnaire disability index; MDHAQ = physician's global assessment; MH = mental health; PH = physical health; PROMIS10 = Patient-Reported Outcomes Measurement Information System global short form; PsAID12 = Psoriatic Arthritis Impact of Disease 12-item questionnaire; RAPID3 = Routine Assessment of Patient Index Data; ROC = receiver operating characteristic; SJC = swollen joint count; TJC = tender joint count.

† No. signifies the number of patients for which the change measure was available. No. varies by measure because certain measures were added at different time points in the protocol (i.e., HAQ DI was added in year 2 of the study).

‡ no. signifies the number of patients who reported minimally important change on the patient global assessment of response.

TNFi or IL-17i. SRMs and MCII were higher for patients with moderate to highly active PsA who would fulfill RCT inclusion criteria (≥3 swollen and ≥3 tender joint counts). Lower overall SRMs and MCII therefore likely reflect the relatively low disease activity at baseline of patients in the clinical setting, and the low MCII may be impacted by the use of patient report of response to therapy as the reference standard. All measures discriminated well between patients who reported improvement and worsening compared to those who reported no change. As expected, a graded change in scores was observed across the 3 patient-reported global assessment of response categories (improved, stayed the same, and worsened), indicating construct validity.

Among patients switching or starting therapy in this real-world PsA population, the mean swollen and tender joint counts were approximately 3 and 6, respectively, compared to an average of 12 and 20 in RCTs (12). More than two-thirds of treatment instances in our study would not have been eligible for a typical PsA RCT. Because the majority of patients had lower baseline disease activity, there was less room to improve, which decreased the chances of showing a response (13). Consequently, most of the measures in our study had relatively low MCII. Importantly, swollen and tender joint counts, which are the crux of the American College of Rheumatology (ACR) response criteria, showed very low SRMs and optimal MCII cutoffs (in the order of improvement in 1 swollen and 1 tender joint count, respectively) with a high misclassification rate. This finding was mostly true even in the moderate to highly active PsA subgroup. Given the low responsiveness of these measures, the ACR criteria may not work well as a primary outcome in pragmatic trials. It is important to consider this while interpreting traditional RCT results, which may not be applicable to the average patient starting treatment in clinical practice. Furthermore, pragmatic trials in PsA will require unique designs that allow the observation of improvement among those with lower disease activity at baseline.

BASDAI had the highest SRM overall and performed the best in those patients with less active PsA, where most measures did not do well. The MCII for BASDAI was similar to a previous study in ankylosing spondylitis (14). BASDAI also showed excellent discrimination (AUC of 0.83) for MCII. We and others have found that despite being developed for axial SpA, BASDAI works quite well as a patient-reported outcome in PsA, regardless of the presence of axial disease (15,16). The questions in BASDAI are primarily focused on disease activity (peripheral joint pain, back/hip/neck pain, tenderness to touch, fatigue, and morning stiffness). Notably, these constructs are all of importance to patients with PsA (5). Moreover, BASDAI addresses only current symptoms (in the past week). This questionnaire contrasts with other instruments that also evaluate function/impact that can be affected by long-term damage/disease consequences, even in those with currently inactive disease. Thus, BASDAI may be desirable to isolate current disease activity for measuring responsiveness. Additional studies are needed to identify whether BASDAI may be a reasonable outcome measure to use across the spectrum of SpA.

PROMIS fatigue and depression were also found to be relatively responsive in this study. However, their discrimination for MCII was relatively lower, with poor sensitivity and specificity and a higher proportion misclassified at their optimal cutoff values. These are not the traditionally used measures such as the Functional Assessment of Chronic Illness Therapy–Fatigue, the Hospital Anxiety and Depression Scale, and the Short Form 36 (SF-36) health survey. The more commonly used health-related quality of life measure in RCTs and observational studies, the SF-36, has a mental component score, but this score generally does not change substantially even in this high disease activity population (17). While observational studies have shown that these 2 PROMIS questionnaires may be reasonable ways of measuring the change in fatigue and psychosocial burden (18), they still need further testing.

PsA-specific outcome measures, cDAPSA and PsAID12 (endorsed by the Group for Research and Assessment of Psoriasis and Psoriatic Arthritis Outcome Measures in Rheumatology), showed overall lower SRMs compared to the above-mentioned measures such as BASDAI, PROMIS fatigue, and depression scores in our study. This finding was specifically true in the less active PsA subgroup. For the moderate to highly active PsA subgroup, however, cDAPSA showed the highest SRM, and PsAID12 performed similar to BASDAI. Both cDAPSA and PsAID12 demonstrated good discrimination for MCII, with good sensitivity and specificity. Thus, while they perform well overall, they may not be expected to change greatly in the subset of patients with low baseline disease activity.

Strengths of this study include testing of PROMIS measures (few studies have tested PROMIS in longitudinal studies in PsA), examining MCIIs and their cutoffs across a variety of patient-reported outcomes, and testing patient-reported outcomes responsiveness in a multicenter real-world setting. However, our study has some limitations. First, there were relatively few patients who responded in the minimally improved group (thus CIs were wide). Additionally, some instruments were added to the patient surveys later than others, including the global assessment of response, so there are different numbers of patients for several of the outcome measures. Next, the standard (truth) was based on improvement judged by the patient (13). However, since patients are the ones experiencing the change, patient judgment may be the most valid standard for this purpose. Axial involvement in the cohort was defined as fulfillment of ASAS axial SpA classification criteria. Workup for axial involvement was done as deemed necessary by examining physicians. Therefore, not all patients underwent evaluation with radiographs or MRI to ascertain axial involvement, and asymptomatic or mild cases might have been missed. Lastly, while the patients included in our study are representative of PsA in the real world, there might be selection bias, as the centers that participated were academic centers with expertise in PsA (i.e., more severe or treatment-resistant PsA).

In conclusion, the responsiveness of patient-reported outcomes in clinical practice is lower compared to that in traditional PsA clinical trials. While PsAID12 and cDAPSA measures performed better in those patients with higher disease activity, BASDAI performed better with low disease activity. Therefore, the selection of measures in future pragmatic trials should account for baseline disease activity. The SRM and MCII for different patient-reported outcomes from this study can guide outcome measure selection, interpretation of the observed change, and sample size calculations for future pragmatic trials in PsA.

ACKNOWLEDGMENTS

The authors thank Tori Fischer and Kat Bush for administrative support, Dr. Michael Ward for sharing his expertise on the methods and interpretation, and all the patients participating in this study.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Karmacharya, Stull, Ogdie, Walsh.

Acquisition of data. Karmacharya, Stull, Scher, Craig, Magrey, Ogdie, Walsh.

Analysis and interpretation of data. Karmacharya, Stull, Stephens-Shields, Husni, Fitzsimmons, Reddy, Ogdie, Walsh.

REFERENCES

1. FitzGerald O, Ogdie A, Chandran V, et al. Psoriatic arthritis. Nat Rev Dis Primers 2021;7:59.
2. Ogdie A, Coates L. The changing face of clinical trials in psoriatic arthritis. Curr Rheumatol Rep 2017;19:21.
3. Ogdie A, Coates LC, Mease P. Measuring outcomes in psoriatic arthritis. Arthritis Care Res (Hoboken) 2020;72 Suppl 10:82–109.
4. Wan MT, Walsh JA, Craig ET, et al. A comparison of physical function instruments in psoriatic arthritis: HAQ-DI vs MDHAQ vs PROMIS10 global physical health. Rheumatology (Oxford) 2021;60:2307–16.
5. Orbai AM, de Wit M, Mease P, et al. International patient and physician consensus on a psoriatic arthritis core outcome set for clinical trials. Ann Rheum Dis 2017;76:673–80.
6. Ward MM, Guthrie LC, Alba M. Domain-specific transition questions demonstrated higher validity than global transition questions as anchors for clinically important improvement. J Clin Epidemiol 2015; 68:655–61.
7. Cohen J. Statistical power: analysis for the behavioural sciences. New York: Academic Press; 1977.
8. Ogdie A, Coates LC, Gladman DD. Treatment guidelines in psoriatic arthritis. Rheumatology (Oxford) 2020;59 Suppl 1:i37–46.
9. Zhou XH, Obuchowski NA, McClish DK. Statistical methods in diagnostic medicine. New York: John Wiley; 2002.
10. Mease PJ, Gladman DD, Collier DH, et al. Etanercept and methotrexate as monotherapy or in combination for psoriatic arthritis: primary results from a randomized, controlled phase III trial. Arthritis Rheumatol 2019;71:1112–24.
11. Rudwaleit M, van der Heijde D, Landewé R, et al. The assessment of SpondyloArthritis International Society classification criteria for peripheral spondyloarthritis and for spondyloarthritis in general. Ann Rheum Dis 2011;70:25–31.
12. Vandendorpe AS, de Vlam K, Lories R. Evolution of psoriatic arthritis study patient population characteristics in the era of biological treatments. RMD Open 2019;5:e000779.
13. Ward MM, Castrejon I, Bergman MJ, et al. Minimal clinically important improvement of Routine Assessment of Patient Index Data 3 in rheumatoid arthritis. J Rheumatol 2019;46:27–30.
14. Van Tubergen A, Black PM, Coteur G. Are patient-reported outcome instruments for ankylosing spondylitis fit for purpose for the axial spondyloarthritis patient? A qualitative and psychometric analysis. Rheumatology (Oxford) 2015;54:1842–51.
15. Taylor WJ, Harrison AA. Could the Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) be a valid measure of disease activity in patients with psoriatic arthritis? Arthritis Rheum 2004;51:311–5.
16. Reddy S, Husni M, Scher J, et al. Use of the BASDAI in psoriatic arthritis patients with and without axial disease [abstract]. Arthritis Rheumatol 2020;72 Suppl 10. URL: https://acrabstracts.org/abstract/use-of-the-basdai-in-psoriatic-arthritis-patients-with-and-without-axial-disease/.
17. Skougaard M, Jørgensen TS, Jensen MJ, et al. Change in psoriatic arthritis outcome measures impacts SF-36 physical and mental component scores differently: an observational cohort study. Rheumatol Adv Pract 2021;5:rkab076.
18. Walsh JA, McFadden ML, Morgan MD, et al. Work productivity loss and fatigue in psoriatic arthritis. J Rheumatol 2014;41:1670–4.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 75, No. 10, October 2023, pp 2190–2198
DOI 10.1002/acr.25100

Longitudinal Patterns of Renal Function in Antineutrophil Cytoplasmic Antibody–Associated Vasculitis

Jennifer S. Hanberg,[1] Claire Cook,[1] Xiaoqing Fu,[1] Hyon K. Choi,[2] Yuqing Zhang,[2] and Zachary S. Wallace[2]

Objective. A spectrum of chronic kidney disease (CKD) and end-stage renal disease (ESRD) may occur in antineutrophil cytoplasmic antibody–associated vasculitis (AAV). The longitudinal trajectory of renal function in AAV is poorly understood.

Methods. Patients with ≥2 creatinine measurements, including at baseline (±30 days of treatment initiation), were included from the Mass General Brigham AAV Cohort. We calculated estimated glomerular filtration rate (eGFR). We incorporated longitudinal changes in eGFR into a group-based trajectory model to identify patients with similar patterns of change in renal function. The chi-square test and the Kruskal-Wallis test were used to evaluate differences between groups in categorical variables and non-normally distributed continuous variables, respectively.

Results. In 255 AAV patients, we identified 4 renal trajectory groups: rapid decline (n = 20), impaired (n = 82), preserved (n = 129), and recovery (n = 24). The rapid decline and impaired groups had greater baseline comorbidity ($P = 0.01$) and lower prevasculitis eGFR ($P = 0.02$). Clinically significant CKD (eGFR <60 ml/minute/1.73 m^2) persisted over 5 years in >75% of the impaired group, compared to <40% of patients in the preserved group ($P < 0.001$). ESRD occurred most frequently in the rapid decline (100%), followed by the impaired and preserved groups (7% each). Baseline AAV renal involvement was present prior to 95% of ESRD. However, ESRD etiology varied, with 90% of rapid-onset ESRD attributed to vasculitis, versus 17–44% in impaired or preserved groups ($P = 0.001$).

Conclusion. We identified 4 longitudinal patterns of renal function after AAV diagnosis. Our findings highlight the burden of CKD in AAV and provide a framework for future research into personalized care in this vulnerable population.

INTRODUCTION

Renal involvement is common in antineutrophil cytoplasmic antibody (ANCA)–associated vasculitis (AAV), affecting more than half of patients with AAV (1). Renal manifestations of AAV span a range of severity, including microscopic hematuria and proteinuria, to transient acute kidney injury, to rapidly progressive glomerulonephritis and end-stage renal disease (ESRD). While ESRD is among the most consequential manifestations of AAV for the approximately 20% of patients reaching this outcome, AAV patients likely experience a spectrum of chronic kidney disease (CKD) with or without ESRD. However, the longitudinal trajectory of renal function following a diagnosis of AAV has not been well characterized in patients with diverse presenting features of AAV (2). Filling this knowledge gap will facilitate efforts to study factors driving renal function patterns after diagnosis, personalize care, and prevent progressive renal disease.

Previous analyses have indicated that not all ESRD in AAV occurs precipitously at the onset of disease. Lionaki et al found that, in a cohort of 136 AAV patients who reached the outcome of ESRD, 43% of the events occurred in patients with no clinical evidence of active vasculitic renal disease at the time of ESRD onset (3). There has been little work using longitudinal data to characterize and differentiate slow progressors from those who experience rapid renal function deterioration early in their disease course. While several studies have developed algorithms to identify AAV patients at high risk of ESRD, most require biopsy data, use only 1 renal function timepoint, do not differentiate slow and

The data reported here have been supplied by the United States Renal Data System. The interpretation and reporting of these data are the responsibility of the authors and in no way should be seen as an official policy or interpretation of the U.S. government.

Supported by the Massachusetts General Hospital Department of Medicine Scholars Research Funding and by the Rheumatology Research Foundation (Resident Research Award). Dr. Wallace's work was supported by the NIH/National Institute of Arthritis and Musculoskeletal and Skin Diseases (K23-AR-073334 and R03-AR-078938) and by the Rheumatology Research Foundation (K Supplement).

[1]Jennifer S. Hanberg, MD, Claire Cook, MPH, Xiaoqing Fu, MS: Massachusetts General Hospital, Boston; [2]Hyon K. Choi, MD, DrPH, Yuqing Zhang, DSc, Zachary S. Wallace, MD, MSc: Massachusetts General Hospital and Harvard Medical School, Boston.

Author disclosures and a graphical abstract can be found online at https://onlinelibrary.wiley.com/doi/10.1002/acr.25100.

Address correspondence via email to Zachary S. Wallace, MD, MSc, at zswallace@mgh.harvard.edu.

Submitted for publication April 1, 2022; accepted in revised form February 7, 2023.

SIGNIFICANCE & INNOVATIONS

- Chronic kidney disease is a common and potentially devastating complication of antineutrophil cytoplasmic antibody–associated vasculitis (AAV) when it culminates in end-stage renal disease (ESRD). Attention has focused on identification of risk factors for ESRD, but less is known about the spectrum of longitudinal changes in renal function in AAV following diagnosis.
- We present an innovative trajectory analysis of longitudinal renal function data that identifies 4 renal trajectory groups, including rapidly declining renal function, impaired renal function, preserved renal function, and renal recovery. The rapid decline and impaired groups had a greater burden of clinically significant kidney disease as well as overall comorbidity compared to the groups with preserved function or recovery.
- Our findings provide an approach that may be leveraged in future studies to inform how we might develop, test, and implement strategies that personalize care for patients with AAV. Such personalized approaches may help prevent or slow the progression of chronic kidney disease in this vulnerable population.

rapid ESRD progress, and are derived from cohorts defined by renal involvement, which may introduce selection biases when identifying risk factors (4,5). Given these gaps in knowledge, we aimed to use a large cohort of AAV patients with diverse disease manifestations, followed over time, to assess whether distinct patterns of renal function change could be identified using trajectory analysis, an agnostic approach (6).

PATIENTS AND METHODS

Study population. We included patients from the Mass General Brigham AAV cohort, a longitudinal inception cohort including patients treated between 2002 and 2017; this cohort has been previously described in detail and is defined by the use of both a validated algorithm and manual chart review for identification of cases (7,8). Inclusion in this study required that a patient have a baseline creatinine value recorded, i.e., a serum creatinine available within ±30 days of the onset of AAV-directed therapy, plus at least 1 subsequent creatinine measurement. The date of initiation of AAV-directed therapy was the index date. The study was approved by the Mass General Brigham Institutional Review Board, protocol number 2016P000633.

Data collection. We extracted data from both structured sources (laboratory values, demographic data, including self-identified race and ethnicity, and medication data) and unstructured sources (clinical notes and chart review) in the electronic medical record. The extraction of variables, including Birmingham Vasculitis Activity Score/granulomatosis with polyangiitis (GPA) score (9), clinical phenotypes, and dates of treatment, has been previously described (7). Follow-up for outcomes began at the index date and was truncated 10 years after treatment initiation. For development of trajectory models, the estimated glomerular filtration rate (eGFR) was evaluated up to monthly for 1 year prior to and 2 years after treatment initiation. If >1 creatinine measurement was available in any month, including the baseline month, the mean of all available measurements in that month was recorded as the monthly value. eGFR measurements after the date of ESRD were recorded as 0 to avoid false signals from fluctuations during dialysis or improvement in renal function after transplantation. See Supplementary Methods in Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25100, for further details regarding data collection for comorbidities and AAV-related treatments.

We calculated eGFR using CKD-epidemiology, without the use of a race multiplier (10). We assessed renal function at baseline, a timepoint we defined as the creatinine closest to the index date (between –30 and + 30 days). In the subset of patients who had such data available, we also assessed pretreatment creatinine prior to vasculitis treatment (between –365 and –30 days relative to the index date). Biopsy results were classified according to the schema of Berden et al based on the original interpretation available in the electronic health record (11).

Renal involvement by AAV was defined according to Birmingham Vasculitis Activity Score/GPA score classification (9). Renal treatment resistance was defined by the absence of remission within 6 months of treatment initiation. Renal remission was defined by stabilization or improvement of the creatinine level with the absence of hematuria for at least 1 month. In patients who did not have a repeat urinalysis available within 6 months of initiation of treatment, stabilization or improvement of creatinine was considered sufficient evidence of remission. This definition is similar to the one previously described by Lionaki et al (3). ESRD was defined as 1) a need for dialysis for >60 days, 2) dialysis until death if the patient died between 14 and 60 days of follow-up, or 3) renal transplant, as identified by chart review and US Renal Data System records (12).

Statistical analysis. *Derivation of trajectory groups.* We used semiparametric, group-based mixture modeling (GBTM) (PROC TRAJ in SAS), to identify groups with similar longitudinal change in renal function, defined as the percentage of baseline eGFR for 1 year prior to and 2 years after treatment initiation (the index date) (6,13). This approach sorts each patient's longitudinal set of measurements (in this case, change in renal function) into clusters and estimates distinct trajectories based on the clusters. A strength of this approach is that renal function

assessments do not have to be available at the same time or at a prescribed frequency.

When selecting a measurement type to evaluate over time as a primary input in the GBTM model, we chose to use the percent change in eGFR rather than absolute change in eGFR or eGFR itself. Percent change was calculated relative to the baseline eGFR measurement (within 30 days of AAV treatment initiation). We selected the percent change metric to avoid biasing the GBTM models overly strongly toward the baseline level of renal function. Due to non-normality of the distribution of percent change in eGFR, we applied a Yeo-Johnson transformation to normalize the variable (14).

We performed model selection according to an accepted iterative procedure; details of the models evaluated in the selection process are reported in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25100 (15,16). We expected to ultimately evaluate models with ≥3 groups, as we anticipated ≥2 trajectories with decreasing renal function over time (3), plus a subset with stable renal function, and minor or no renal involvement, and likely some patients with improvement in an initial renal function insult. When selecting the degree of polynomial to test in our models, we anticipated that there may be up to 2 inflection points in renal trajectories, reflecting the initial renal insult and subsequent stabilization at very low eGFR (i.e., ESRD) or improved renal function (patients with renal recovery); thus, we evaluated up to the third-order polynomial. We did not evaluate models with >5 groups due to decreasing size of trajectory groups (i.e., <5% of the overall cohort). Our final model was selected from all possible candidate models containing up to 5 groups and up to the third-order polynomial, using statistical validity (Bayesian information criterion and the estimate of the log Bayes factor), group size criteria (>5% of the sample) and face validity based on the authors' clinical expertise. Code for the final model and additional details of model development are included in the Supplementary Methods in Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25100. After deciding on a model through this process, we labeled each trajectory (e.g., rapid decline, impaired, preserved, recovery) based on the apparent renal function trend in each group. Similar approaches have been published previously (17–19).

Statistical methods for group comparisons. Measures of central tendency are reported as mean ± SD or median (25th–75th percentile with interquartile range [IQR]). Between-group differences were tested using the chi-square or Fisher's exact test for categorical data, Kruskal-Wallis test for continuous variables, or analysis of variance for normally distributed continuous variables, and were evaluated between all 4 trajectory groups unless otherwise specified. We used the nonparametric method of the log-rank test to evaluate differences in time to the composite outcome of ESRD or death between groups; we did not perform proportional hazards testing due to violation of the proportional hazards assumption.

Due to the small sample size of the more severe renal dysfunction trajectories (rapid decline and impaired), we collapsed these trajectories into 1 group and compared it with any other group (i.e., preserved/recovery group) for regression analyses. We performed logistic regression to evaluate the relation of hypertension and diabetes mellitus (baseline characteristics a priori known to be associated with kidney disease) to the risk of group membership (coded binarily as rapid decline/impaired versus preserved/recovery). In the logistic regression model we adjusted for age, sex, and ANCA type. Statistical significance was defined as a 2-tailed *P* value less than 0.05. SAS software, version 9.4 was used for all statistical analysis. Patients were not directly involved in the design of this research.

RESULTS

Of the 484 patients in the overall Mass General Brigham AAV cohort, 255 were included in the final analysis, as they had a baseline creatinine value within 30 days of the index date and at least 1 additional creatinine measurement. The median number of monthly renal function measurements was 9 (IQR 5–15), and the median number of measurements was numerically similar across trajectory groups (P = 0.08) (Table 1). The majority of patients were female (60%) and the mean ± SD age at treatment initiation was 61 ± 17 years (Table 1). Comorbid baseline diabetes mellitus and hypertension at baseline were present in 41 patients (16%) and 122 patients (48%), respectively.

Among those with pretreatment (–365 to –30 days of the index date) creatinine measurements available (n = 143), the median pretreatment eGFR was 65 ml/minute/1.73 m^2 (IQR 34–81), with 48% of patients having an eGFR of 60 ml/minute/1.73 m^2 or less. The median baseline (±30 days of the index date) eGFR among the entire cohort was 40 ml/minute/1.73 m^2 (IQR 17–80).

Describing trajectory groups. We identified 4 trajectory groups of renal function using group-based trajectory modeling. Based on the trajectory of renal function in each group, we refer to these as rapid decline (n = 20 [8%]), impaired (n = 82 [32%]), preserved (n = 129 [51%]), and recovery (n = 24 [9%]) (Figure 1).

Differences in renal function between groups were observable at baseline. Among the entire cohort, the rapid decline group had the lowest baseline eGFR with a median of 7 ml/minute/1.73 m^2 (IQR 6–9) compared to the impaired (median 25 ml/minute/1.73 m^2 [IQR 17–36]), preserved (median 78 ml/minute/1.73 m^2 [IQR 51–92]), and recovery groups (median 10 ml/minute/1.73 m^2 [IQR 7–16]). Because pretreatment renal function measurements were relatively sparse in this data set, there is significant fluctuation, with apparent large increases and decreases in mean eGFR in the smaller trajectory groups (rapid

Table 1. Characteristics of patients at diagnosis stratified by renal function trajectory*

Characteristic	Overall	Rapid decline	Impaired	Preserved	Recovery	*P*
No. (%) of cohort	255 (100)	20 (8)	82 (32)	129 (51)	24 (9)	–
Posterior probability of group membership, median (IQR)	–	1.0 (1.0–1.0)	0.99 (0.98–1.0)	0.99 (0.92–1.0)	0.99 (0.99–1.0)	–
No. of visits (36 months of follow-up), median (IQR)	9 (5–15)	7 (4–10)	10 (7–16)	8 (6–15)	8 (4–15)	0.08
Demographic data						
Age at diagnosis, mean ± SD	61 ± 17	62 ± 17	64 ± 16	59 ± 17	63 ± 19	0.21
Female	153 (60)	11 (55)	53 (65)	78 (60)	11 (46)	0.40
Race/ethnicity						
White	213 (87)	18 (95)	70 (86)	103 (84)	22 (92)	–
Black	6 (2)	0 (0)	3 (4)	2 (2)	1 (4)	–
Hispanic	9 (4)	0 (0)	2 (2)	7 (6)	0 (0)	–
Asian	4 (2)	1 (5)	0 (0)	3 (2)	0 (0)	–
Other	5 (2)	0 (0)	1 (1)	3 (2)	1 (4)	–
Not recorded	9 (4)	0 (0)	5 (6)	4 (3)	0 (0)	–
Clinical characteristics at baseline						
ANCA type						0.91
MPO	182 (71)	15 (75)	60 (73)	91 (71)	16 (67)	–
PR3	73 (29)	5 (25)	22 (27)	38 (29)	8 (33)	–
BVAS/GPA, median (IQR)	4 (4–6)	5 (4–7)	4 (4–6)	4 (3–6)	6 (4.5–7)	<0.001
Renal†	175 (69)	20 (100)	68 (83)	63 (49)	24 (100)	<0.001
Mucosal/ocular†	24 (9)	0 (0)	3 (4)	20 (16)	1 (4)	0.01
Pulmonary†	106 (42)	11 (55)	26 (32)	63 (49)	6 (25)	0.02
Neurologic†	25 (10)	2 (10)	6 (7)	16 (12)	1 (4)	0.49
Biopsy category						
No.‡	66	9	31	17	9	–
Crescentic	16 (24)	2 (22)	8 (26)	3 (18)	3 (19)	–
Focal	10 (15)	0 (0)	7 (23)	3 (18)	0 (0)	–
Mixed	14 (21)	2 (22)	6 (19)	3 (18)	3 (33)	–
Sclerotic	20 (30)	4 (44)	8 (26)	6 (35)	2 (22)	–
Normal or other	6 (9)	1 (11)	2 (7)	2 (12)	1 (11)	–
Comorbidities						
CCI, median (IQR)	4 (2–6)	4.5 (3.5–6)	4 (3–6)	3 (1–7)	4 (2.5–5)	0.01
Baseline diabetes mellitus	41 (16)	3 (15)	17 (21)	19 (15)	2 (8)	0.46
Baseline hypertension§	122 (48)	12 (60)	43 (52)	54 (42)	13 (54)	0.25
Baseline renal function						
Pretreatment eGFR (−365 to −30 days), median (IQR)¶	65 (34–81)	48 (37–88)	44 (25–72)	71 (50–81)	75 (41–88)	0.02
Baseline eGFR (±30 days), median (IQR)¶	40 (17–80)	7 (6–9)	25 (17–36)	78 (51–92)	10 (7–16)	<0.001
Induction treatment						
CYC-based	100 (39)	8 (40)	34 (41)	47 (36)	12 (50)	–
RTX-based	121 (47)	11 (55)	40 (49)	58 (45)	12 (50)	–
Other	34 (13)	1 (5)	8 (10)	24 (19)	1 (4)	–
Plasma exchange#	67 (26)	15 (75)	25 (30)	13 (10)	14 (58)	<0.001
Maintenance treatment						
RTX**	113 (44)	6 (30)	39 (48)	59 (46)	9 (38)	0.07
Non-RTX immunosuppression	116 (45)	7 (35)	38 (46)	62 (48)	9 (38)	0.18
None	33 (13)	6 (30)	10 (12)	15 (12)	2 (8)	0.02
Lost to follow-up or deceased before maintenance period	20 (8)	3 (15)	5 (6)	7 (5)	5 (21)	0.02

* Values are the number (%) unless indicated otherwise. *P* values shown reflect analysis of variance (for normally distributed continuous variables), Kruskal-Wallis (for other continuous variables), and chi-square or Fisher's test (for categorical variables) results across the 4 groups. Column sums are >100% in some cases for maintenance treatment because patients could be categorized as receiving both non-rituximab (RTX) maintenance immunosuppression and maintenance RTX. Data in this table are not from the United States Renal Data System. ANCA = antineutrophil cytoplasmic antibody; BVAS = Birmingham Vasculitis Activity Score; CCI = Charlson Comorbidity Index; CYC = cyclophosphamide; eGFR = estimated glomerular filtration rate (ml/minute/1.73 m^2); GPA = granulomatosis with polyangiitis; IQR = interquartile range; MPO = myeloperoxidase; PR3 = proteinase 3.
† Organ system involvement by BVAS/GPA score.
‡ Total number with biopsy in each group. Columns may add up to smaller numbers due to nonclassifiable or normal biopsies.
§ Includes n = 3 total patients with comorbid hypertension of unknown onset time.
¶ Relative to the index date (date of treatment initiation); n = 143 for pretreatment eGFR, n = 255 for baseline eGFR.
# Plasma exchange was not mutually exclusive with other treatment regimens.
** Patients who received short, "bridging" courses of cyclophosphamide with initiation of RTX were categorized as receiving primarily RTX-based induction treatment (n = 67 total, including n = 6 rapid decline, n = 24 impaired, n = 31 stable, and n = 6 recovery.) Maintenance treatment column sums may exceed 100% because patients could be classified as using both RTX and non-RTX maintenance immunosuppression.

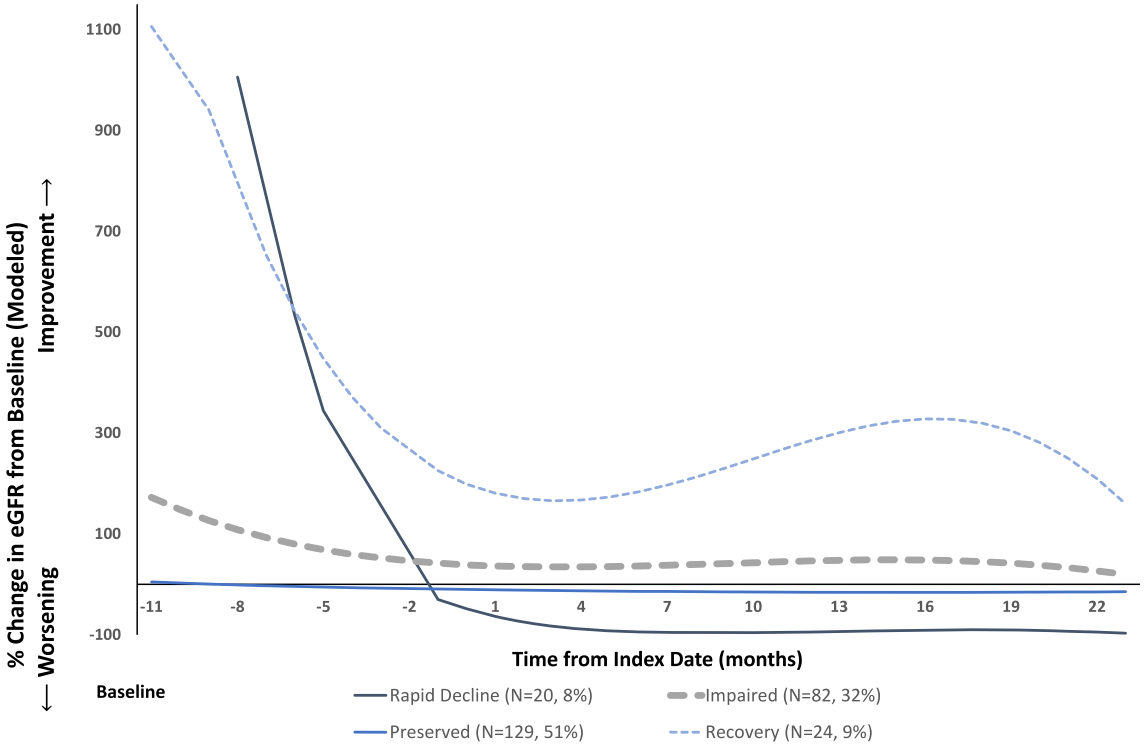

Figure 1. Renal function trajectories in 4 groups identified by group-based trajectory modeling. This chart displays the model estimated renal function measurement at each monthly timepoint within a given trajectory group. eGFR = estimated glomerular filtration rate.

decline and recovery) prior to treatment in the raw within-group averages over time as shown in Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25100. These pretreatment fluctuations are less likely to be clinically meaningful than patterns of renal function after treatment initiation; the modeled trajectories, which emphasize the trends that are best supported by our models, are shown in Figure 1.

The trajectory of renal disease after treatment initiation varied between groups (Figures 1 and 2 and Table 2). The rapid decline group was characterized by the rapid development of ESRD. In the impaired group, the eGFR decreased from a pretreatment (–30 to –365 days prior to treatment initiation) median of 44 to 25 ml/minute/1.73 m^2 at initiation of treatment (among those with pretreatment eGFR available [n = 49 of 82]). This initial insult observed at the time of diagnosis in the impaired group appeared to slightly improve over follow-up at a group level; however, renal function in this group remained substantially impaired (5-year median eGFR 48 ml/minute/1.73 m^2). The impaired group had a lower median eGFR at 2 and 5 years of follow-up compared to the preserved and recovery groups; lower median eGFR was reflected in a greater burden of CKD in the impaired group over time, with >75% of patients in the impaired group remaining with CKD stage 3 or greater at years 1, 2, and 5 (Table 2 and Figure 2). These patterns of kidney disease severity over time were also observed in a sensitivity analysis including only patients who had 5-year follow-up renal function data available (n = 160) (see Supplementary Figure 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25100).

The renal recovery group showed an initial decrement in renal function that dramatically improved; unlike the impaired group, the recovery group frequently had resolution of clinically significant CKD (stage 3+) (Figure 2). The recovery group had a baseline median eGFR of 10 ml/minute/1.73 m^2 (IQR 7–16) which improved to 24 ml/minute/1.73 m^2 (IQR 18–35) by 1 year and 58 ml/minute/1.73 m^2 (IQR 33–71) by 5 years of follow-up. The preserved group had little change in renal function during follow-up. ESRD was uncommon in the recovery and preserved groups (1 [4%] and 9 [7%], respectively).

Baseline features associated with trajectory group membership. Age, sex, and race were not statistically different between groups (Table 1). However, the baseline comorbidity burden, as measured by the Charlson Comorbidity Index, was greater in the rapid decline and impaired groups (4.5 [IQR 3.5–6] and 4 [IQR 3–6], respectively) compared to the preserved group (3 [IQR 1–7]); P = 0.01). The preserved renal function group had the lowest proportion, with baseline AAV-associated renal involvement (n = 63 [49%] in preserved compared to n = 20 [100%] in rapid decline, n = 68 [83%] in impaired, and n = 24 [100%] in recovery). The distribution of myeloperoxidase- versus proteinase 3–ANCA type was similar across groups (P = 0.91).

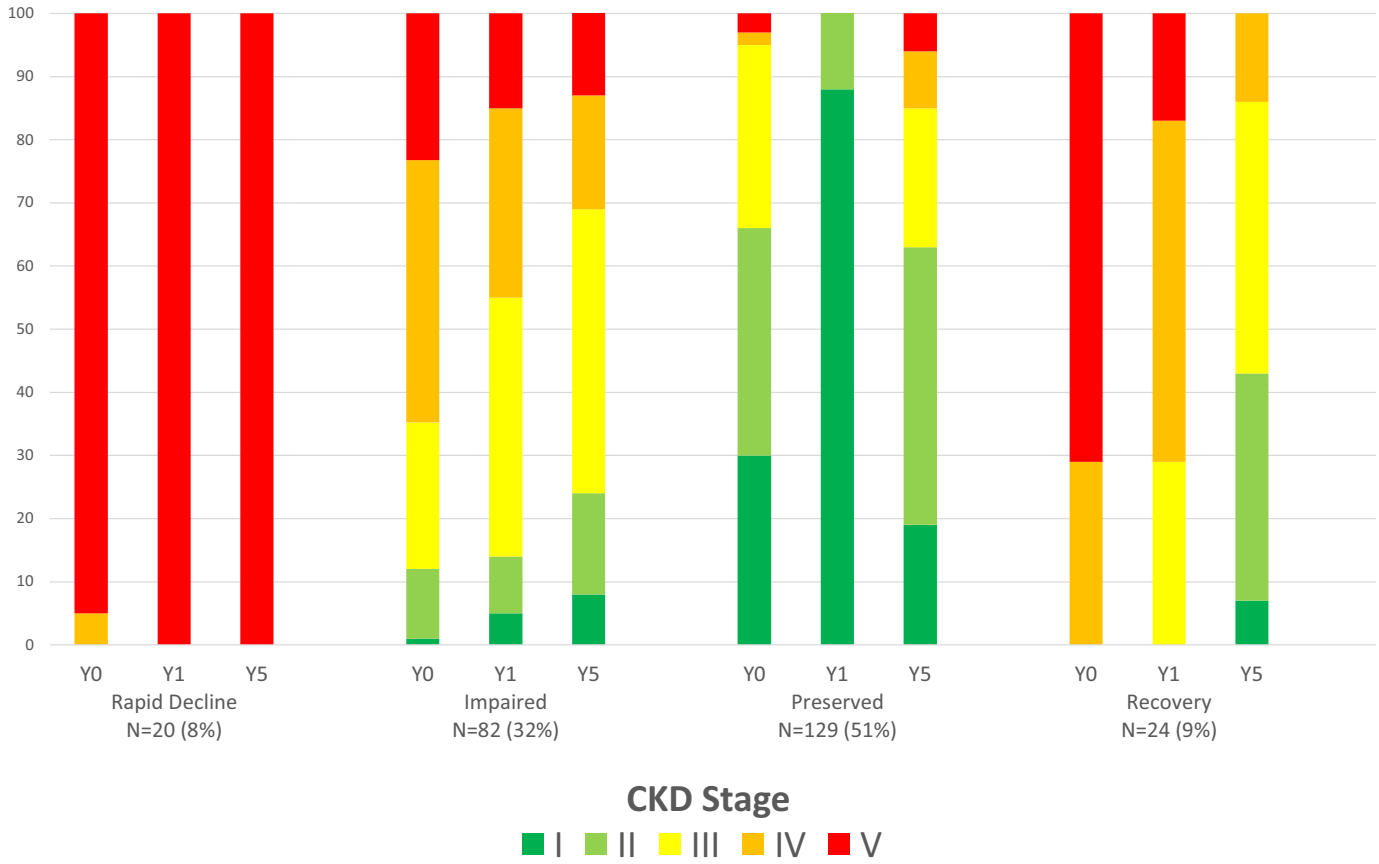

Figure 2. Progression of renal disease as represented by chronic kidney disease (CKD) stage over time among trajectory groups. The horizontal axis depicts the year of assessment of renal function on the top line and the trajectory group on the second line. The CKD stage is depicted in colors ranging from stage 1 (green, estimated glomerular filtration rate [eGFR] >90) to stage 5 (red, eGFR <15). Data in this figure are not from the United States Renal Data System. Y0 = baseline renal function within 30 days of initiation of therapy; Y1 = CKD stage at 1 year of follow-up, averaged over 1 year; Y5 = CKD stage at 5 years of follow-up, averaged over 1 year.

Renal biopsies were uncommon (n = 66 [26%]) in this cohort. Histopathologic categorization among rapid decline patients (number with biopsy = 9 of 20) was more often sclerotic (n = 4 [44%]), compared to the impaired group (number with biopsy = 31 of 82; sclerotic categorization observed in 8 patients [26%]). Despite this trend, no statistically significant differences were observed across groups ($P = 0.86$).

We observed differences in the proportion of subjects in each trajectory with hypertension, a key driver of CKD in the general population. A history of hypertension at baseline was most common among patients in the rapid decline group (n = 12 [60%]), followed by patients in the impaired and recovery groups (n = 43 [52%] and n = 13 [54%], respectively). Hypertension was less common in the preserved group, although these differences did not meet statistical significance (n = 54 [42%]; $P = 0.25$ for difference across all 4 groups) (Table 1). There was no strong association observed between a history of diabetes mellitus at baseline and trajectory group ($P = 0.46$). After adjustment for age, sex, and ANCA type, neither baseline hypertension (adjusted odds ratio [OR] 1.1 [95% confidence interval (95% CI) 0.6–2.1]) nor diabetes mellitus (adjusted OR 1.3 [95% CI 0.6–2.8]) was associated with the outcome of membership in a composite renal dysfunction group (collapsed rapid decline and impaired groups together). In the subgroup of patients with CKD stage 3 or 4 at baseline, differences in comorbidity burden were less striking between the impaired and preserved groups, suggesting that other factors may drive renal function trajectories in this subgroup (see Supplementary Table 2 and Supplementary Results in Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25100).

Use of AAV-specific treatments across trajectory groups. Cyclophosphamide- and rituximab-based induction regimens were used at similar rates between groups, with the exception that non-rituximab, non-cyclophosphamide–based induction regimens (steroids alone or conventional disease-modifying antirheumatic drugs) were used more frequently in the preserved group (n = 24 [19%] versus 4–10% in all other groups). Plasma exchange was used more frequently among patients with either rapid decline

Table 2. Clinical outcomes, longitudinal renal function, and mortality stratified by trajectory group*

Outcome	Overall	Rapid decline	Stable impaired	Stable preserved	Recovery	*P*
No. (%) of cohort	255 (100)	20 (8)	82 (32)	129 (51)	24 (9)	–
Treatment resistance, no. (%)						
Yes	51 (20)	17 (85)	15 (18)	15 (12)	4 (17)	–
No	107 (42)	1 (5)	49 (60)	39 (30)	18 (75)	–
Insufficient data	19 (7)	2 (10)	8 (10)	7 (5)	2 (8)	–
No renal involvement	78 (31)	0 (0)	10 (12)	68 (53)	0 (0)	–
Renal outcomes						
Any dialysis, no. (%)	52 (20)	20 (100)	14 (17)	13 (10)	5 (21)	<0.001
Transplant, no. (%)	8 (3)	4 (20)	1 (1)	3 (2)	0 (0)	<0.001
Permanent ESRD, no. (%)†	36 (14)	20 (100)	6 (7)	9 (7)	1 (4)	<0.001
Active vasculitis as cause of ESRD, no. (%)‡	23 (62)	18 (90)	1 (17)	4 (44)	0 (0)	0.001
Time to ESRD‡	0.34 (0.02–4.7)	0.02 (0.004–0.15)	4.2 (1.1–6.8)	4.7 (2.5–5.6)	10 (10–10)	–
eGFR at 1 year	43 (23–76)	2 (1–5)	32 (22–46)	71 (44–89)	24 (18–36)	<0.001
eGFR at 2 years	53 (31–77)	0 (0–0)	41 (26–58)	69 (45–86)	48 (29–61)	<0.001
eGFR at 5 years	54 (29–77)	0 (0–0)	48 (26–60)	69 (43–84)	58 (36–71)	<0.001
CKD 3+ at 1 year, no. (%)	162 (64)	20 (100)	71 (87)	47 (36)	24 (100)	<0.001
CKD 3+ at 2 years, no. (%)	117 (58)	12 (100)	54 (79)	39 (36)	12 (75)	<0.001
CKD 3+ at 5 years, no. (%)	89 (56)	10 (100)	39 (76)	32 (38)	8 (57)	<0.001
Mortality						
Death, no. (%)	70 (27)	8 (40)	27 (33)	29 (22)	6 (25)	0.21
Follow-up time to ESRD, death, or censorship	6.5 (3.9–9.8)	0.2 (0.004–0.15)	6.4 (4.0–9.8)	7.3 (5.0–10)	7.4 (5.1–10)	<0.001
Follow-up time to death or censorship	7.3 (4.7–10)	6.8 (1.8–10)	6.6 (4.0–10)	7.4 (5.5–10)	7.4 (5.1–10)	–

* Values are the median (interquartile range) unless indicated otherwise. End-stage renal disease (ESRD) was defined as 1) a need for dialysis for >60 days, 2) dialysis until death if the patient died between 14 and 60 days of follow-up, or 3) renal transplant, as identified by chart review and US Renal Data System records. Because temporary dialysis was included in the count of patients with dialysis, more patients received dialysis than experienced ESRD by this definition. Data in this table are not from the United States Renal Data System. CKD = chronic kidney disease; CKD 3+ = CKD stage 3 or higher, i.e., estimated glomerular filtration rate (eGFR) <60 ml/minute/1.73 m^2.
† ESRD newly occurring during follow-up, i.e., no earlier than 1 year prior to initiation of treatment for vasculitis.
‡ Among those with ESRD.

or renal recovery, which was expected given the indication for consideration of this treatment among individuals with severe renal injury (n = 15 [75%] rapid decline and n = 14 [58%] recovery, versus impaired n = 25 [30%] and preserved n = 13 [10%]; $P < 0.001$). Rituximab and non-rituximab maintenance medications were used at statistically similar rates across groups overall ($P = 0.07$ and $P = 0.18$, respectively); however, they were somewhat less frequently prescribed to patients in the rapid decline group.

Outcomes of ESRD and mortality across trajectory groups. Over a mean follow-up of 73 months, 36 patients (14%) experienced ESRD (2.3 ESRD events per 100 person-years) (Table 2). Of the 36 patients with ESRD, 2 did not have baseline renal involvement by AAV; ESRD was attributed to an unrelated immune complex glomerulonephritis in 1 case and a preexisting condition dating back to childhood in another case. The majority of ESRD occurred in the rapid decline group (n = 20 [56% of all ESRD events]), followed by the preserved group (n = 9 [25%]) and the impaired group (n = 6 [17%]). Among those patients who experienced ESRD in the rapid decline group and impaired groups, ESRD occurred at a median of 0.02 years (IQR 0.004–0.2) after the index date compared to 4.2 years (IQR 1.1–6.8) in the impaired group. The composite outcome of ESRD or death occurred earliest in the rapid decline group, followed by the impaired group and then the preserved and recovery groups (log-rank $P < 0.001$). Renal treatment resistance was more common in the rapid decline than the impaired group (n = 17 [85%] versus n = 15 [18%]; $P < 0.001$).

Differences in the etiology of ESRD were observed across trajectory groups. Active vasculitis was thought to be the cause of ESRD in 90% of patients categorized as rapid decline (n = 18 of 20), compared to 17% of patients categorized as impaired (n = 1 of 6), 44% of patients in the preserved group (n = 4 of 9) and no patients in the recovery group (n = 0 of 1; $P = 0.001$ for comparison across 4 groups).

Over a mean follow-up of 81 months, 70 patients (27%) died (4.1 deaths per 100 person-years) (Table 2). Although the small size of some trajectory groups limits conclusions, a greater proportion of deaths was observed in the groups characterized by worse renal function (rapid decline: n = 8 [40%], impaired: n = 27 [33%], preserved: n = 29 [22%], recovery: n = 6 [25%]).

DISCUSSION

We used an agnostic methodologic approach to identify 4 patterns of longitudinal renal function in an incident cohort of

AAV patients with diverse manifestations. Each trajectory group was characterized by distinct courses with respect to renal function, such that patients with persistent renal dysfunction generally follow 1 of 2 clinical courses: precipitous decline to ESRD, or chronic impaired renal function, during which many patients are left with advanced CKD and a portion develop ESRD. Nearly 70% of patients with ESRD reach that endpoint from these 2 CKD trajectories, which are distinguished from other groups by demographic and clinical features. Our findings provide an innovative approach to conceptualize the impact of a new diagnosis of AAV on renal function and highlight groups that would benefit from personalized, multidisciplinary approaches to care. These trajectories serve as a framework within which strategies can be developed, tested, and implemented to further improve AAV outcomes, especially for CKD.

Patients in the rapid decline group had quick, nearly universal onset of ESRD with very severe renal impairment at treatment initiation. Given that the eGFR prior to diagnosis in the rapid decline group was lower than that observed in the recovery group (at 48 ml/minute/1.73 m^2 compared to >70), they may have been affected by a more indolent, subclinical progression of vasculitic renal damage prior to diagnosis. Alternatively, patients in the rapid decline group may have had preexisting kidney disease due to hypertension and other comorbidities that predisposed them to rapid, irreversible renal deterioration with AAV onset. These possibilities highlight the potential impact of delays in AAV diagnosis as well as the uncertainties regarding whether ideal treatment for patients presenting with AAV renal involvement might vary based on their prediagnosis renal function, if available (20,21). A strength of our study in contrast to others was the availability, in a subset of patients, of prediagnosis measures of renal function.

Of particular interest regarding renal function trajectories are the characteristics that distinguish the impaired group from preserved and recovery groups, given the differences between groups in the etiology of ESRD occurring months to years after AAV diagnosis. Previous studies have suggested that a subset of AAV patients develop late-onset ESRD in the absence of clinically evident active AAV; these patients likely reflect the impaired trajectory phenotype, as illustrated in our study (3). The impaired group was older than the preserved and recovery groups and had higher rates of general comorbidity compared to the preserved group; renal involvement by AAV and hypertension was more common in the impaired group than the preserved group. These observations highlight high-risk patients who might benefit from a personalized approach to care based on their age, comorbidity burden, and renal function trajectory. Indeed, even by 3 months, there were striking differences in the trajectory of renal function among those classified in the impaired versus recovery groups, highlighting the implications of persistent renal dysfunction at this time point.

Our findings provide important empirical evidence in an incident AAV cohort, followed from diagnosis, that both comorbidities and the history of prior renal involvement by AAV are influential factors contributing to CKD and progression to ESRD months to years after diagnosis. Additional studies are needed to evaluate whether longitudinal assessment of renal biomarkers beyond serum creatinine and urinalysis, such as soluble CD163, CD25 and others, may identify patients who stand to benefit from modified or intensified immunosuppression as opposed to more aggressive ESRD risk factor modification (22,23). This type of personalized care for patients at risk for renal impairment may improve outcomes by helping to prevent or reduce the progression of CKD in AAV. Identifying and studying these trajectories is increasingly important as survival improves for patients with AAV and management strategies evolve (24–27). The need to identify and study is especially true in the face of the ongoing COVID-19 pandemic, as providers and patients weigh individualized decisions regarding the risks of decreasing or holding immunosuppression. Robust data to inform renal risk stratification in these scenarios would be of significant clinical utility.

Previous work has established several risk models for the outcome of ESRD in AAV, incorporating biopsy features, age, ANCA serotype, induction therapy type, and initial renal function (4,28–30). This study adds to the existing literature via an unbiased approach to describe the longitudinal arc of renal function and CKD in AAV in more nuanced terms than prior ESRD-focused work in this space. Our work quantitatively confirms the existence and frequency of a phenotype with largely stable but clinically significant renal function impairment. Additionally, we performed these analyses in an AAV cohort with diverse manifestations and have provided detailed examination of the clinical features of patients exhibiting these trajectories, which lends greater generalizability to our work.

Our study has certain limitations. First and most importantly, we believe that the current investigation provides important preliminary evidence that there are clinically distinct renal phenotypes or trajectories that exist in AAV; however, given our small sample size, the single health care system in which our study was conducted, and the need for validation in other cohorts, based on this study alone we cannot apply the concept of trajectory group membership to individual patients in clinical practice. Additional studies will be necessary to validate these observed trajectories in other cohorts. Second, attributing causality of renal outcomes or trajectory to the treatments received is not possible, due to the likely role of confounding by indication. However, most treatment strategies were distributed similarly across trajectory groups. Third, renal biopsies were not obtained in all patients, limiting our ability to associate specific renal biopsy features with trajectory membership. However, biopsies are uncommonly obtained in the context of positive ANCA tests with a consistent clinical context, as previously observed in our health care system (12). Fourth, our definition of renal treatment resistance may be biased by persistent hematuria in patients with renal damage despite lack of ongoing active vasculitis; however, the definition

of treatment resistance reflects a similar definition used in another study assessing progression to ESRD in AAV. Finally, small sample sizes in some of our trajectory groups limit the assessment of statistically significant differences between groups and may increase the likelihood of Type 2 errors, especially for heterogeneous outcomes, such as treatment or AAV-related organ involvement.

In this AAV cohort, we identified 4 distinct patterns of change in renal function. The increased baseline comorbidity and risk of clinically significant CKD in the rapid decline and impaired trajectories underscores the importance of personalized, multidisciplinary care for AAV patients; further investigation of tailored strategies to preserve renal function is warranted. Our findings provide a framework for future research into next steps to improve renal outcomes for this vulnerable population.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Hanberg, Wallace.

Acquisition of data. Hanberg, Cook, Fu, Wallace.

Analysis and interpretation of data. Hanberg, Fu, Choi, Zhang, Wallace.

REFERENCES

1. Binda V, Moroni G, Messa P. ANCA-associated vasculitis with renal involvement. J Nephrol 2018;31:197–208.
2. Sagmeister MS, Grigorescu M, Schonermarck U. Kidney transplantation in ANCA-associated vasculitis. J Nephrol 2019;32:919–26.
3. Lionaki S, Hogan SL, Jennette CE, et al. The clinical course of ANCA small-vessel vasculitis on chronic dialysis. Kidney Int 2009;76:644–51.
4. Brix SR, Noriega M, Tennstedt P, et al. Development and validation of a renal risk score in ANCA-associated glomerulonephritis. Kidney Int 2018;94:1177–88.
5. Menez S, Hruskova Z, Scott J, et al. Predictors of renal outcomes in sclerotic class anti-neutrophil cytoplasmic antibody glomerulonephritis. Am J Nephrol 2018;48:465–71.
6. Nagin DS, Jones BL, Passos VL, et al. Group-based multi-trajectory modeling. Stat Methods Med Res 2018;27:2015–23.
7. Wallace ZS, Fu X, Harkness T, et al. All-cause and cause-specific mortality in ANCA-associated vasculitis: overall and according to ANCA type. Rheumatology (Oxford) 2020;59:2308–15.
8. Watts R, Lane S, Hanslik T, et al. Development and validation of a consensus methodology for the classification of the ANCA-associated vasculitides and polyarteritis nodosa for epidemiological studies. Ann Rheum Dis 2007;66:222–7.
9. Mukhtyar C, Lee R, Brown D, et al. Modification and validation of the Birmingham Vasculitis Activity Score (version 3). Ann Rheum Dis 2009;68:1827–32.
10. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. Ann Int Med 2009;150:604–12.
11. Berden AE, Ferrario F, Hagen EC, et al. Histopathologic classification of ANCA-associated glomerulonephritis. J Am Soc Nephrol 2010;21:1628–36.
12. Cook CE, Fu X, Zhang Y, et al. Validation of antineutrophil cytoplasmic antibody-associated vasculitis as the cause of end-stage renal disease in the US Renal Data System. ACR Open Rheumatol 2022:4;8–12.
13. Jones BL, Nagin DS. Advances in group-based trajectory modeling and an SAS procedure for estimating them. Sociol Methods Res 2007;35:542–71.
14. Yeo IK, Johnson RA. A new family of power transformations to improve normality or symmetry. Biometrika 2000;87:954–9.
15. Nguena Nguefack HL, Page MG, Katz J, et al. Trajectory modelling techniques useful to epidemiological research: a comparative narrative review of approaches. Clin Epidemiol 2020;12:1205–22.
16. Andruff H, Carraro N, Thompson A, et al. Latent class growth modelling: a tutorial. Tutorials Quant Meth Psychol 2009;5:11–24.
17. Hanberg JS, Akgun KM, Hsieh E, et al. Incidence and presentation of sarcoidosis with and without HIV infection. Open Forum Infect Dis 2020;7:ofaa441.
18. Marshall BD, Tate JP, McGinnis KA, et al. Long-term alcohol use patterns and HIV disease severity. AIDS 2017;31:1313–21.
19. Boucquemont J, Loubere L, Metzger M, et al. Identifying subgroups of renal function trajectories. Nephrol Dial Transplant. 2017;32 Suppl 2:ii185–93.
20. Taimen K, Mustonen A, Pirila L. The delay and costs of diagnosing systemic vasculitis in a tertiary-level clinic. Rheumatol Ther 2021;8:233–42.
21. Sreih AG, Cronin K, Shaw DG, et al. Diagnostic delays in vasculitis and factors associated with time to diagnosis. Orphanet J Rare Dis 2021;16:184.
22. Dekkema GJ, Abdulahad WH, Bijma T, et al. Urinary and serum soluble CD25 complements urinary soluble CD163 to detect active renal anti-neutrophil cytoplasmic autoantibody-associated vasculitis: a cohort study. Nephrol Dial Transplant 2019;34:234–42.
23. Moran SM, Scott J, Clarkson MR, et al. The clinical application of urine soluble CD163 in ANCA-associated vasculitis. J Am Soc Nephrol 2021;32:2920–32.
24. Wallace ZS, Miloslavsky EM. Management of ANCA associated vasculitis. BMJ 2020;368:m421.
25. Wallace ZS, Lu N, Miloslavsky E, et al. Nationwide trends in hospitalizations and in-hospital mortality in granulomatosis with polyangiitis (Wegener's). Arthritis Care Res (Hoboken) 2017;69:915–21.
26. Wallace ZS, Lu N, Unizony S, et al. Improved survival in granulomatosis with polyangiitis: a general population-based study. Semin Arthritis Rheum 2016;45:483–9.
27. Holle JU, Gross WL, Latza U, et al. Improved outcome in 445 patients with Wegener's granulomatosis in a German vasculitis center over four decades. Arthritis Rheum 2011;63:257–66.
28. De Lind van Wijngaarden RA, Hauer HA, Wolterbeek R, et al. Clinical and histologic determinants of renal outcome in ANCA-associated vasculitis: a prospective analysis of 100 patients with severe renal involvement. J Am Soc Nephrol 2006;17:2264–74.
29. Hauer HA, Bajema IM, Van Houwelingen HC, et al. Determinants of outcome in ANCA-associated glomerulonephritis: a prospective clinico-histopathological analysis of 96 patients. Kidney Int 2002;62:1732–42.
30. Lee T, Gasim A, Derebail VK, et al. Predictors of treatment outcomes in ANCA-associated vasculitis with severe kidney failure. Clin J Am Soc Nephrol 2014;9:905–13.

Arthritis Care & Research
Vol. 75, No. 10, October 2023, pp 2199–2206
DOI 10.1002/acr.25116

American College of Rheumatology
Empowering Rheumatology Professionals

Updating and Validating the Rheumatic Disease Comorbidity Index to Incorporate ICD-10-CM Diagnostic Codes

Anthony Dolomisiewicz,[1] Hanifah Ali,[2] Punyasha Roul,[2] Yangyuna Yang,[2] Grant W. Cannon,[3] Brian Sauer,[3] Joshua F. Baker,[4] Ted R. Mikuls,[1] Kaleb Michaud,[5] and Bryant R. England[1]

Objective. To update and validate the Rheumatic Disease Comorbidity Index (RDCI) utilizing International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) codes.

Methods. We defined ICD-9-CM (n = 1,068) and ICD-10-CM (n = 1,425) era cohorts (n = 862 in both) spanning the ICD-9-CM to ICD-10-CM transition in a multicenter, prospective rheumatoid arthritis registry. Information regarding comorbidities was collected from linked administrative data over 2-year assessment periods. An ICD-10-CM code list was generated from crosswalks and clinical expertise. ICD-9– and ICD-10–derived RDCI scores were compared using intraclass correlation coefficients (ICC). The predictive ability of the RDCI for functional status and death during follow-up was assessed using multivariable regression models and goodness-of-fit statistics (Akaike's information criterion [AIC] and quasi information criterion [QIC]) in both cohorts.

Results. Mean ± SD RDCI scores were 2.93 ± 1.72 in the ICD-9-CM cohort and 2.92 ± 1.74 in the ICD-10-CM cohort. RDCI scores had substantial agreement in individuals who were in both cohorts (ICC 0.71 [95% confidence interval 0.68–0.74]). Prevalence of comorbidities was similar between cohorts with absolute differences <6%. Higher RDCI scores were associated with a greater risk of death and poorer functional status during follow-up in both cohorts. Similarly, in both cohorts, models including the RDCI score had the lowest QIC (functional status) and AIC (death) values, indicating better model performance.

Conclusion. The newly proposed ICD-10-CM codes for the RDCI-generated comparable RDCI scores to those derived from ICD-9-CM codes and are highly predictive of functional status and death. The proposed ICD-10-CM codes for the RDCI can be used in rheumatic disease outcomes research spanning the ICD-10-CM era.

INTRODUCTION

The Rheumatic Disease Comorbidity Index (RDCI) was designed as a tool to quantify the burden of comorbidities and to account for the association of this burden with long-term health outcomes in patients with rheumatic diseases. The RDCI was initially developed using patient questionnaires to collect data regarding 10 comorbid conditions, which were weighted to generate a score ranging from 0 to 9 (1). RDCI scores have been shown to be predictive of physical function, quality of life, and death (1,2). Previously, we validated RDCI scores using administrative claims data and International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) codes (2). Using claims data, the RDCI favorably compared to other existing general population comorbidity indices in predicting physical function

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

Dr. Baker's work was supported by VA Clinical Science Research & Development and Rehabilitation Research & Development (merit grants CX-001703 and RX-003644, respectively). Dr. Mikuls's work was supported by VA Biomedical Laboratory Research & Development (grant I01-BX-004660), the Department of Defense (grant PR-200793), the Rheumatology Research Foundation, and the NIH/National Institute of General Medical Sciences (grant U54-GM-115458). Dr. Michaud's work was supported by the Rheumatology Research Foundation and the NIH/National Institute of General Medical Sciences (grant U54-GM-115458). Dr. England's work was supported by VA Clinical Science Research & Development (grant IK2-CX-002203).

[1]Anthony Dolomisiewicz, MD, Ted R. Mikuls, MD, MSPH, Bryant R. England, MD, PhD: Veterans Affairs Nebraska-Western Iowa Health Care System and University of Nebraska Medical Center, Omaha; [2]Hanifah Ali, Punyasha Roul, MS, Yangyuna Yang, MBBS, PhD: University of Nebraska Medical Center, Omaha; [3]Grant W. Cannon, MD, Brian Sauer, PhD: Salt Lake City VA and University of Utah, Salt Lake City; [4]Joshua F. Baker, MD, MSCE: Corporal Michael J. Crescenz VA and University of Pennsylvania, Philadelphia; [5]Kaleb Michaud, PhD: University of Nebraska Medical Center, Omaha, and FORWARD, The National Data Bank for Rheumatic Diseases, Wichita, Kansas.

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.25116&file=acr25116-sup-0001-Disclosureform.pdf.

Address correspondence via email to Bryant R. England, MD, PhD, at bryant.england@unmc.edu.

Submitted for publication January 31, 2023; accepted in revised form March 21, 2023.

SIGNIFICANCE & INNOVATIONS

- The Rheumatic Disease Comorbidity Index (RDCI) was developed to determine the contribution of comorbidities to death, physical disability, and quality of life in patients with rheumatic diseases.
- The RDCI was previously developed from patient questionnaires and validated using International Classification of Diseases, Ninth Edition, Clinical Modification (ICD-9-CM) codes, but has not been updated and validated for the ICD-10-CM system.
- Use of a proposed set of ICD-10-CM codes to generate the RDCI produced comparable RDCI scores, comorbidity prevalence, and prediction of death and functional status as the ICD-9-CM–derived RDCI.
- With the proposed code set, the RDCI can be used in rheumatic disease outcomes research encompassing the ICD-10-CM era.

and death in individuals with rheumatoid arthritis (RA) (2). Additionally, the RDCI is more feasible to obtain and a more versatile index, performing well in both self-reported and administrative data (1,2). Because of its performance, feasibility, and versatility, it has been broadly used in RA studies (3) as well as in studies of psoriatic arthritis (4,5), spondyloarthritis (5–7), gout (8,9), lupus (10), vasculitis (11,12), and osteoporosis (13).

In the US, health care systems transitioned from the ICD-9-CM classification system for medical conditions to the ICD-10-CM classification system on October 1, 2015. The ICD-10-CM coding system contains a nearly 5-fold increase in the number of codes available in order to provide improved specificity in classifying medical conditions (14). The transition to ICD-10-CM has required updating the coding of chronic conditions and comorbidity indices that were developed when the previous classification system was in place in order to confirm the validity of these indices using ICD-10-CM codes (15–17). The RDCI, derived from ICD-10-CM diagnostic codes, has not yet been validated. Because different comorbidity indices utilize unique definitions for conditions and diagnostic code sets, the generation of an ICD-10-CM code set specific to the RDCI is necessary in order to use the code set in studies spanning both timeframes. Therefore, the objectives of this study were to translate the ICD-9-CM codes used to generate the RDCI to ICD-10-CM codes and to validate the predictive ability for death and functional status of the RDCI derived from ICD-10-CM codes.

PATIENTS AND METHODS

Study design and patient population. We performed a cohort study using the Veterans Affairs Rheumatoid Arthritis (VARA) registry. This study has been approved by the Nebraska-Western Iowa Health Care System VA Subcommittee of Human Studies institutional review board. The VARA registry is a multicenter prospective cohort of US veterans, >18 years of age who have been diagnosed with RA by a rheumatologist and fulfilled the 1987 American College of Rheumatology classification criteria for RA (18). We assembled ICD-9-CM and ICD-10-CM cohorts within VARA over 2 separate time periods immediately before and after the transition to ICD-10-CM codes to represent the different classification eras in the US (Figure 1). A subgroup of patients was included in both cohorts. Within each period, the initial 2 years were designated as comorbidity ascertainment (ICD-9-CM October 1, 2013 to September 30, 2015; ICD-10-CM January 1, 2016 to December 31, 2017). Patients were then followed up from the index date (ICD-9-CM October 1, 2015; ICD-10-CM January 1, 2018) up to December 31, 2021, considered the outcome observation period. A 3-month gap between the introduction of ICD-10-CM in the US and the start of our ICD-10-CM cohort was implemented to account for health care systems and providers acclimating to the new coding system.

In the current study, VARA participants were included if they were alive as of the index date, had enrolled in the VA ≥2 years prior to the index date, and had ≥1 VARA visit during the 2 years prior to the index date. Participants could be included in both the ICD-9-CM and ICD-10-CM cohorts.

Translating ICD-9-CM codes to ICD-10-CM codes. We translated the previously validated ICD-9-CM codes to ICD-10-CM codes using tools available from www.ICD9Data.com and www.ICD10Data.com (2). The websites are owned and operated by Alkaline Software, and the tools provide comprehensive lists of ICD-9-CM and ICD-10-CM codes, as well as suggested conversions between the 2 classification systems (19,20). Utilizing these tools, a list of potential ICD-10-CM codes and their descriptions for each comorbidity in the RDCI was generated (Table 1). The list was reviewed by a physician (BRE) for clinical relevance and accuracy.

Comorbidity data collection. We collected data regarding comorbid conditions comprising the RDCI by linking the VARA registry participants to national VA administrative databases in the Corporate Data Warehouse (CDW). A condition was considered present if ≥1 diagnostic code from a VA or non-VA inpatient or outpatient encounter was recorded during the 2-year comorbidity ascertainment period. RDCI scores (range 0–9) were calculated from the individual conditions using the original formula: 2 × lung disease + (2 × [heart attack, other cardiovascular disease, or stroke] or 1 × hypertension) + fracture + depression + diabetes + cancer + (ulcer or stomach symptom) (2).

Study outcomes for predictive modeling. Death and functional status were assessed over the outcome observation period after collection of comorbidities. Vital status was determined using linked VA mortality data, and functional status was

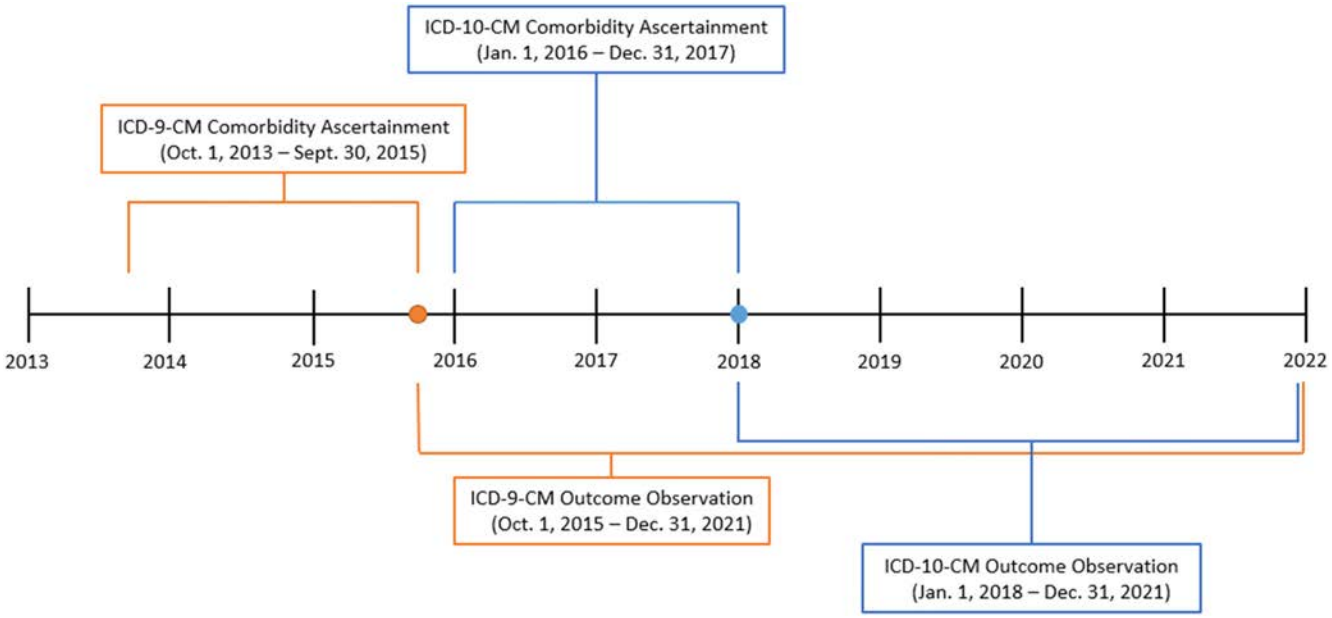

Figure 1. Study design and data collection timelines. International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) and ICD-10-CM cohorts were identified over separate time periods immediately before and after the transition to ICD-10-CM. The initial 2 years were designated as the comorbidity ascertainment period (ICD-9-CM October 1, 2013 to September 30, 2015; ICD-10-CM January 1, 2016 to December 31, 2017). The outcome observation period started on the index date (colored circles) (ICD-9-CM October 1, 2015 [orange]; ICD-10-CM January 1, 2018 [blue]) and continued until December 31, 2021.

obtained from the VARA registry. Functional status in participants in the VARA registry was measured using the Multidimensional Health Assessment Questionnaire (MDHAQ) as part of routine care (21). All MDHAQ values recorded after the index date within the outcome observation period in each cohort were utilized in analyses.

Table 1. RDCI ICD-9-CM to ICD-10-CM crosswalk*

Comorbidity/condition	ICD-9-CM	ICD-10-CM
Myocardial infarction	410.x-412.x	I20.0; I21.x-I22.x; I24.x; I25.110; 125.2; I25.700; I25.710; I25.720; I25.730; I25.750; I25.760; I25.790
Hypertension	401.x; 405.x	I10.x; I15.x; I16.x
Diabetes mellitus	249.x-250.x	E08.x-E11.x; E13.x
Depression	296.2-296.39; 300.4; 311.x	F32.x (excluding F32.81); F33.x; F34.1
Ulcer or stomach problem	531.x–535.7x; 536.3; 536.8–536.9; 578.9	K25.x-K30.x; K31.84; K31.89; K31.9; K52.81; K92.2
Stroke	430.x–431.x; 433.x–435.x; 997.02	G45.x (excluding G45.3 and G45.4); G46.0-G46.4; I60.x-I61.x; I63.x; I65.x-I66.x; I97.81x-I97.82x
Spine, hip, or leg fracture	733.13–733.16; 733.93; 733.96–733.98; 805.x–806.x; 808.x; 820.x–821.x; 823.x; 827.x	M48.4x-M48.5x; M80.05x-M80.06x; M80.08x; M80.85x-M80.86x; M80.88x; M84.35x-M84.36x; M84.45x-M84.46x; M84.55x-M84.56x; M84.65x-M84.66x; M84.75x; S12.x (excluding S12.8x); S22.0x; S32.x; S72.x; S82.1x-S82.2x; S82.311x; S82.312x; S82.319x S82.4x; S82.81x-S82.83x; S82.86x; S82.89x; S82.9x
Other CVD	394.x–396.x; 402.x; 404.x; 413.x–414.x; 424.0–424.3; 425.x–428.x	I05.x-I06.x; I08.x; I11.x; I13.x; I20.1; I20.8; 120.9; I25.10; I25.111; I25.118; I25.119; I25.3; I25.41; I25.42; I25.5; I25.6; I25.701; I25.708; I25.709; I25.711; I25.718; I25.719; I25.721; I25.728; I25.729; I25.731; I25.738; I25.739; I25.751; I25.758; I25.759; I25.761; I25.768; I25.769; I25.791; I25.798; I25.799; I25.810; I25.811; I25.812; I25.82; I25.83; I25.84; I25.89; I25.9; I34.x-I35.x; I36.x; I37.x; I42.x-I45.x; I46.x; I47.x-I50.x; R00.1
Lung disease	490.x–493.99; 494.x; 495.x; 496.x; 500. x–505.x; 515.x–517.x; 714.81	J40.x-J47.x; J60.x-J67.x; J84.x; J99; M05.10-M05.19
Cancer	140.x–209.39	C00.x-C26.x; C30.x-C41.x; C43.x-C58.x; C60.x-C75.x; C76.x-C86.x; C88.2-C88.9; C90.x-C93.x; C94.0x-C94.3x; C94.8x; C95.x-C96.x; C7A.x; C4A.x; D03.x

* CVD = cardiovascular disease; ICD-9-CM = International Classification of Diseases, Ninth Revision, Clinical Modification; RDCI = Rheumatic Disease Comorbidity Index.

Study covariates and descriptive variables. Several demographic and RA-related factors were selected as covariates for predictive models or as descriptive variables. Age, sex, race, smoking status (current, former, and never), RA disease duration, and Disease Activity Score in 28 joints (DAS28) were obtained from the VARA registry, where they were collected by treating providers. These variables were collected at registry enrollment, with the exception of the DAS28, which is routinely collected over the course of care. The closest DAS28 value preceding the index date was selected. Anti–cyclic citrullinated peptide (anti-CCP) antibody was measured in a standardized manner among VARA participants using serum collected at the time of enrollment, as previously described (22). Use of conventional synthetic disease-modifying antirheumatic drugs (csDMARDs), biologic DMARDs (bDMARDs) or targeted synthetic DMARDs (tsDMARDs), and glucocorticoids as indicated by pharmacy dispensings, was obtained from linked VA CDW pharmacy data (23). DMARDs were assessed over the year prior to the index date, while glucocorticoid use was assessed over the 90 days prior to the index date.

Statistical analysis. Patient characteristics were summarized using descriptive statistics. Comorbidities and RDCI scores were descriptively assessed in each cohort separately, as well as among the individuals who were included in both the ICD-9-CM cohort and ICD-10-CM cohort. Agreement in individual comorbidities among individuals who were included in both cohorts was assessed using an unweighted Cohen's κ coefficient. Intraclass correlation coefficients (ICC) in a 2-way mixed effects model with absolute agreement were calculated to assess agreement of the RDCI scores. Kappa and ICC values ranged from 0 to 1 and were interpreted as follows: values <0.01 indicating chance agreement, values between 0.01 and 0.20 indicating slight agreement, values between 0.21 and 0.40 indicating fair agreement, values between 0.41 and 0.60 indicating moderate agreement, values between 0.61 and 0.80 indicating substantial agreement, and values >0.81 indicating almost perfect agreement (24–26).

Cox proportional hazards regression models were used to assess the ability of the RDCI score to predict death. We used generalized estimating equations models to assess how the RDCI predicted functional status in each coding system. Two sets of covariates were used in regression models. The first model adjusted for age, sex, and race, while the fully adjusted model additionally included smoking status, csDMARDs, bDMARDs/tsDMARDs, glucocorticoid use, and anti-CCP positivity.

We subsequently determined the improvement of model performance for both death and functional status after including the RDCI score. In mortality analyses, the Akaike's Information Criterion (AIC) was calculated using Cox regression models with and without the RDCI score. For functional status, quasi information criterion (QIC) values were calculated using generalized estimating equations models with and without the RDCI. AIC and QIC differences after RDCI inclusion were calculated to compare model performance, with lower AIC and QIC values indicating better model fit (2). Analyses limiting follow-up in the ICD-9-CM cohort to the maximum follow-up duration in the ICD-10-CM cohort produced similar results compared to the primary approach (data not shown). Missing covariate data were addressed using the missing covariate indicator method (27). Analyses were conducted using Stata version 17 within the VA Informatics and Computing Infrastructure environment.

RESULTS

Patient characteristics. Both the ICD-9-CM cohort (n = 1,068) and ICD-10-CM cohort (n = 1,425) were predominantly male (ICD-9-CM 89.2%, ICD-10-CM 87.3%), White (ICD-9-CM 76.8%, ICD-10-CM 74.0%), and had a mean age of 67–68 years old (ICD-9-CM mean ± SD 67.3 ± 10.2 years, ICD-10-CM mean ± SD 68.2 ± 9.9 years) (Table 2). The majority of patients were anti-CCP antibody positive (ICD-9-CM 77.5%, ICD-10-CM 78.1%). The most frequent RA treatments were csDMARDs in both cohorts (ICD-9-CM 59.5%, ICD-10-CM 65.3%). A total of 862 patients were included in both cohorts. The median follow-up time was 6.3 years in the ICD-9-CM cohort and 4.0 years in the ICD-10-CM cohort.

Comorbidity prevalence and RDCI scores. RDCI scores were generated using the proposed ICD-9-CM to ICD-10-CM crosswalk codes shown in Table 1. The mean ± SD RDCI scores were 2.93 ± 1.72 in the ICD-9-CM cohort and 2.92 ± 1.74 in the ICD-10-CM cohort. Among individuals who were in both cohorts, RDCI scores demonstrated substantial agreement (ICC 0.71 [95% confidence interval (95% CI) 0.68–0.74]) (Table 3). The prevalence of individual comorbidities in the RDCI was also similar among the cohorts defined during the 2 coding ascertainment periods, with all absolute differences <6% (range 0.1–5.7%). Myocardial infarction, hypertension, diabetes mellitus, depression, stroke, other cardiovascular disease, lung disease, and cancer had moderate agreement or higher (range κ = 0.47–0.84) among individuals in both cohorts. Fracture (κ = 0.13) and ulcer/gastrointestinal (GI) problems (κ = 0.27) had slight and fair agreement, respectively.

Mortality prediction. In the ICD-9-CM cohort, 228 deaths occurred over 5,716 patient-years compared to 210 deaths over 5,144 patient-years in the ICD-10-CM cohort. Higher RDCI scores were associated with a greater risk of death in both cohorts (Table 4) (ICD-9-CM adjusted hazard ratio [HR] 1.17 [95% CI 1.08, 1.27]; ICD-10-CM adjusted HR 1.24 [95% CI 1.14, 1.35]). Comparing models with and without the RDCI, model performance improved with the addition of the RDCI in both ICD-9-CM and ICD-10-CM cohorts (Table 4). The reduction

Table 2. Patient characteristics according to study cohort*

Demographic and clinical characteristics	ICD-9-CM cohort (n = 1,068)†	ICD-10-CM cohort (n = 1,425)‡
Age, mean ± SD years	67.3 ± 10.2	68.2 ± 9.9
Male sex	953 (89.2)	1,244 (87.3)
White	820 (76.8)	1,054 (74.0)
Smoking status		
Current	264 (24.7)	341 (23.9)
Former	553 (51.8)	740 (51.9)
Never	231 (21.6)	293 (20.6)
RA disease duration, mean ± SD years	15.4 ± 10.9	16.6 ± 11.3
Anti-CCP antibody positive	753 (77.5)	1,051 (78.1)
DAS28-CRP, mean ± SD	3.1 ± 1.2	2.9 ± 1.2
RA treatments		
csDMARDs	635 (59.5)	931 (65.3)
bDMARDs or tsDMARDs	297 (27.8)	470 (33.0)
Glucocorticoids	292 (27.3)	330 (23.2)

* Except where indicated otherwise, values are the number (%) of patients. Percentages are among those with available data. Anti-CCP = anti–cyclic citrullinated peptide; bDMARDs = biologic disease-modifying antirheumatic drugs; csDMARDs = conventional synthetic DMARDs; Clinical Modification; RA = rheumatoid arthritis; tsDMARDs = targeted synthetic DMARDs.
† In the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) cohort, data were missing for smoking status (n = 20), anti–cyclic citrullinated peptide (anti-CCP antibody) (n = 97), and Disease Activity Score in 28 joints using the C-reactive protein level (DAS28-CRP) (n = 269).
‡ In the ICD-10-CM cohort, data were missing for smoking status (n = 51), anti-CCP antibody positivity (n = 79), and DAS28-CRP (n = 384).

in AIC in the ICD-9-CM cohort was 15.29 in the age-, sex-, and race-adjusted model and 11.53 in the fully adjusted model. AIC reductions were even larger in the ICD-10-CM cohort with the reduction of 25.93 in the age-, sex-, and race-adjusted model and 23.85 in the fully adjusted model.

Functional status prediction. The mean ± SD MDHAQ scores during follow-up were 1.0 ± 0.7 in the ICD-9-CM cohort and 0.9 ± 0.6 in the ICD-10-CM cohort. Higher RDCI scores were significantly associated with poorer functional status in both ICD-9-CM cohort and ICD-10-CM cohort (Table 5) (ICD-9-CM β = 0.06 [95% CI 0.04–0.08]; ICD-10-CM β = 0.07 [95% CI 0.05–0.09]). Functional status prediction was improved in models that included the RDCI in both the ICD-9-CM cohort and ICD-10-CM cohort as indicated by reductions in the QIC (Table 5). The reduction in QIC in the ICD-9-CM cohort was 101.5 in the age-, sex-, and race-adjusted model and 89.12 in the fully adjusted model. QIC reduction in the ICD-10-CM cohort was 104.4 in the age-, sex-, and race-adjusted model and 93.2 in the fully adjusted model.

DISCUSSION

In this study, we aimed to translate the ICD-9-CM codes defined for the RDCI to ICD-10-CM codes and validate the predictive ability of the RDCI for the key long-term rheumatic disease outcomes of death and functional status. We created a set of ICD-10-CM codes that generated comparable RDCI scores and individual comorbidity prevalence estimates compared to those derived from ICD-9-CM codes within a large, multicenter RA registry, then tested the predictive ability of the RDCI score using the ICD-10-CM code set. We found the ICD-10-CM code set was able to predict death and functional status as well as the ICD-9-CM codes. With the updated codes (Table 1), the RDCI can continue to be used in rheumatic disease outcomes research utilizing ICD-10-CM era data.

After assembling an ICD-10-CM code set using a crosswalk and clinical expertise, we calculated RDCI scores in RA cohorts spanning the ICD-9-CM and ICD-10-CM time periods. RDCI scores were comparable between cohorts (mean scores of 2.93 versus 2.92) and, among individuals observed during both ICD-9-CM and ICD-10-CM eras, RDCI scores had substantial

Table 3. Agreement between RDCI scores and prevalence of comorbidities*

	All participants		Participants in both cohorts (n = 862)		
	ICD-9-CM (n = 1,068)	ICD-10-CM (n = 1,425)	ICD-9-CM	ICD-10-CM	ICC (95% CI) or Cohen's κ†
RDCI score, mean ± SD	2.93 ± 1.72	2.92 ± 1.74	2.89 ± 1.70	3.02 ± 1.76	0.71 (0.68, 0.74)
Comorbid condition, %					
Myocardial infarction	7.0	4.9	6.5	5.6	0.47
Other CVD	34.6	36.5	32.1	37.8	0.63
Stroke	5.1	6.5	4.5	6.3	0.49
Hypertension	65.5	63.8	65.1	63.2	0.71
Diabetes mellitus	29.0	28.3	27.8	31.2	0.84
Lung disease	29.8	32.3	29.4	33.1	0.62
Cancer	20.5	19.9	20.3	21.8	0.58
Ulcer or GI problems	7.8	7.7	6.8	8.0	0.27
Fracture‡	1.7	2.7	1.3	2.0	0.13
Depression	27.6	23.6	28.3	25.8	0.61

* CVD = cardiovascular disease; GI = gastrointestinal; ICD-9-CM = International Classification of Diseases, Ninth Revision, Clinical Modification.
† Intraclass correlation coefficient (ICC) (95% confidence interval [95% CI]) was used for Rheumatic Disease Comorbidity Index (RDCI) score. For all comorbid conditions, Cohen's κ was used.
‡ Spine, hip, or leg fracture.

Table 4. Performance of the RDCI score for predicting death in Cox regression models*

Model	No. of deaths	Adjusted HR for RDCI (95% CI)	*P*	Model AIC no RDCI	Model AIC RDCI	ΔAIC
ICD-9-CM (n = 1,068)	228					
Age-, sex-, and race-adjusted		1.19 (1.10–1.29)	<0.001	2,951.75	2,936.46	−15.29
Fully adjusted†		1.17 (1.08–1.27)	<0.001	2,930.67	2,919.14	−11.53
ICD-10-CM (n = 1,425)	210					
Age-, sex-, and race-adjusted		1.25 (1.15–1.36)	<0.001	2,894.92	2,868.99	−25.93
Fully adjusted†		1.24 (1.14–1.35)	<0.001	2,892.93	2,869.08	−23.85

* 95% CI = 95% confidence interval; AIC = Akaike's information criterion; HR = hazard ratio; RDCI = Rheumatic Disease Comorbidity Index. See Table 2 for other definitions.
† Model includes age, sex, race, smoking status, csDMARDS, bDMARDs or tsDMARDs, glucocorticoid use, and anti-CCP antibody positivity.

agreement (κ = 0.71). Similarly, when we assessed the prevalence of individual comorbidities between coding systems, the absolute differences were small (all <6%). While evaluating individuals in both cohorts, we discovered that agreement between ICD-9-CM and ICD-10-CM varied across individual conditions. Individual conditions with greater chronicity had greater agreement compared to acute conditions. For example, diabetes mellitus, hypertension, other cardiovascular disease, lung disease, and depression (chronic conditions) all had at least substantial agreement (κ >0.60). In contrast, ulcer or GI problems and spine, hip, or leg fracture (acute conditions) had fair and slight agreement (κ = 0.27 and 0.13, respectively).

We postulate that the greater agreement for conditions that are more chronic is primarily the result of these conditions being continually addressed and recorded during clinical visits conducted over the course of the study. However, it is also possible that medical care for acute conditions may have solely occurred outside the VA system (28). Finally, the variability in agreement may be related to differences in the prevalence of these conditions, with those having the lowest prevalence also having poorer agreement.

Other studies have similarly developed ICD-10-CM code sets for comorbidity indices developed during the ICD-9-CM era. Quan et al constructed ICD-10-CM coding algorithms for the Charlson and Elixhauser comorbidity indices using Canadian administrative data (17). They found that the frequency of most comorbidities was similar between ICD-9-CM and ICD-10-CM coding algorithms and found the ICD-10-CM code set to either match or outperform the ICD-9-CM versions in predicting in-hospital death. As in our study, peptic ulcer disease was a condition that differed in frequency between ICD-9-CM and ICD-10-CM. Glasheen et al have since updated the ICD-10-CM coding for the Charlson comorbidity index, validating the predictive performance of ICD-10-CM codes for hospital admission and death (16). Sears and Rundell developed updated ICD-9-CM and ICD-10-CM code lists for the Functional Comorbidity Index (FCI), assessing concordance before and after the transition to ICD-10-CM (15). The frequency of individual comorbidities was consistent between coding algorithms for 13 of 18 comorbidities. While the FCI retained predictive value for length of hospital stay with ICD-10-CM codes, there was an interaction between the index and coding algorithm suggesting modest reduced performance with ICD-10-CM codes. Our findings are consistent with these studies; appropriate ICD-10-CM code sets for comorbidity indices can agree with ICD-9-CM versions and retain predictive value for health outcomes.

Comorbidity burden is a crucial determinant of long-term outcomes that include death and functional status among individuals with rheumatic diseases (3,22,29–31). Thus, we tested the validity of the ICD-10-CM codes by assessing the ability of the RDCI score to predict death and functional status. As expected, higher RDCI scores, indicating a greater comorbidity burden, were associated with a higher risk of death and poorer functional status during follow-up in both the ICD-9-CM cohort and ICD-10-CM cohort. Moreover, metrics of overall model performance (AIC, QIC) for predicting these outcomes in each cohort also substantially improved when the RDCI was included. Taken together, these findings demonstrate that the RDCI calculated with the proposed ICD-10-CM codes maintains expected predictive value for key long-term outcomes.

Table 5. Performance of the RDCI for predicting Multidimensional Health Assessment Questionnaire scores in generalized estimating equations models*

Model	No. of observations	β for RDCI (95% CI)	*P*	Model QIC (No RDCI)	Model QIC (RDCI)	ΔQIC
ICD-9-CM (n = 931)	8,769					
Age-, sex-, and race-adjusted		0.06 (0.04–0.08)	<0.001	3,485.37	3,383.86	−101.51
Fully adjusted†		0.06 (0.04–0.08)	<0.001	3,435.45	3,346.33	−89.12
ICD-10-CM (n = 1,175)	6,864					
Age-, sex-, and race-adjusted		0.07 (0.05–0.09)	<0.001	2,567.83	2,463.43	−104.4
Fully adjusted†		0.07 (0.05–0.09)	<0.001	2,510.82	2,417.62	−93.2

* 95% CI = 95% confidence interval; QIC = quasi information criterion; RDCI = Rheumatic Disease Comorbidity Index. See Table 2 for other definitions.
† Model includes age, sex, race, smoking status, csDMARDS, bDMARDs or tsDMARDs, glucocorticoid use, and anti-CCP antibody positivity.

This study has limitations. The cohorts were predominantly male, White, and of middle to older age, which may limit generalizability, since RA is a condition that predominantly occurs in women. However, the RDCI was previously found to perform similarly for predicting death and functional status in VA and non-VA cohorts (2). The ICD-9-CM cohort was defined during the end of the ICD-9-CM era, when clinicians may have been more experienced with diagnostic code selection, compared to the ICD-10-CM, cohort which was defined during the early period of the ICD-10-CM era. To reduce misclassification resulting from this, we excluded the initial 3-month time period after ICD-10-CM code implementation. Patients may have received care outside the VA, so our data may underestimate the prevalence of individual conditions and RDCI scores, though we do not expect this to differ by ICD era. Finally, we evaluated the predictive ability for death and functional status but recognize comorbidity burden is also associated with other long-term outcomes, such as quality of life, which were not available.

In conclusion, we generated a set of ICD-10-CM codes for the RDCI that produce comparable RDCI scores and chronic disease prevalence estimates compared to those derived from previously validated ICD-9-CM codes. RDCI scores calculated using these ICD-10-CM codes are highly predictive of functional status and death, comparing favorably to scores based on ICD-9-CM codes. The proposed RDCI codes can be used in rheumatic disease outcomes research spanning the ICD-10-CM era.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Ali, Roul, Yang, Mikuls, Michaud, England.
Acquisition of data. Dolomisiewicz, Ali, Roul, Yang, Mikuls, Michaud, England.
Analysis and interpretation of data. Dolomisiewicz, Roul, Yang, Cannon, Sauer, Baker, Mikuls, Michaud, England.

REFERENCES

1. Michaud K, Wolfe F. Comorbidities in rheumatoid arthritis. Best Pract Res Clin Rheumatol 2007;21:885–906.
2. England BR, Sayles H, Mikuls TR, et al. Validation of the Rheumatic Disease Comorbidity Index. Arthritis Care Res (Hoboken) 2015;67:865–72.
3. Aslam F, Khan NA. Tools for the assessment of comorbidity burden in rheumatoid arthritis. Front Med (Lausanne) 2018;5:39.
4. Bavière W, Deprez X, Houvenagel E, et al. association between comorbidities and quality of life in psoriatic arthritis: results from a multicentric cross-sectional study. J Rheumatol 2020;47:369–76.
5. Ogdie A, Maksabedian Hernandez EJ, Shaw Y, et al. Side effects of methotrexate and tumor necrosis factor inhibitors: differences in tolerability among patients with psoriatic arthritis and rheumatoid arthritis. ACR Open Rheumatol 2022;4:935–41.
6. Kieskamp SC, Paap D, Carbo MJ, et al. Central sensitization has major impact on quality of life in patients with axial spondyloarthritis. Semin Arthritis Rheum 2022;52:151933.
7. Putrik P, Ramiro S, Moltó A, et al. Individual-level and country-level socioeconomic determinants of disease outcomes in SpA: multinational, cross-sectional study (ASAS-COMOSPA). Ann Rheum Dis 2019;78:486–93.
8. Coburn BW, Michaud K, Bergman DA, et al. Allopurinol dose escalation and mortality among patients with gout. Arthritis Rheumatol 2018;70:1298–307.
9. Spaetgens B, Wijnands JM, van Durme C, et al. Content and construct validity of the Rheumatic Diseases Comorbidity Index in patients with gout. Rheumatology (Oxford) 2015;54:1659–63.
10. Katz P, Pedro S, Trupin L, et al. The impact of asthma and chronic obstructive pulmonary disease (COPD) on patient-reported outcomes in systemic lupus erythematosus (SLE). ACR Open Rheumatol 2021;3:221–30.
11. Elfishawi M, Rakholiya J, Gunderson TM, et al. Lower frequency of comorbidities prior to onset of giant cell arteritis: a population-based study. J Rheumatol doi: https://doi.org/10.3899/jrheum.220610. 2022. E-pub ahead of print.
12. Luo Y, Xu J, Jiang C, et al. Trends in the inpatient burden of coronary artery disease in granulomatosis with polyangiitis: a study of a large national dataset. J Rheumatol 2021;48:548–54.
13. Cornelissen D, Boonen A, Evers S, et al. Improvement of osteoporosis Care Organized by Nurses: ICON study - protocol of a quasi-experimental study to assess the (cost)-effectiveness of combining a decision aid with motivational interviewing for improving medication persistence in patients with a recent fracture being treated at the fracture liaison service. BMC Musculoskelet Disord 2021;22:913.
14. Centers for Disease Control and Prevention, National Center for Health Statistics. International Classification of Diseases, (ICD-10-CM/PCS) transition - background. 2015. URL: https://www.cdc.gov/nchs/icd/icd10cm_pcs_background.htm.
15. Sears JM, Rundell SD. Development and testing of compatible diagnosis code lists for the Functional Comorbidity Index: International Classification of Diseases, ninth revision, Clinical Modification and International Classification of Diseases, 10th revision, Clinical Modification. Med Care 2020;58:1044–50.
16. Glasheen WP, Cordier T, Gumpina R, et al. Charlson comorbidity index: ICD-9 update and ICD-10 translation. Am Health Drug Benefits 2019;12:188–97.
17. Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care 2005;43:1130–9.
18. Mikuls TR, Reimold A, Kerr GS, et al. Insights and implications of the VA Rheumatoid Arthritis Registry. Fed Pract 2015;32:24–9.
19. The web's free ICD-9-CM & ICD-10-CM medical coding reference. URL: www.icd9data.com.
20. The web's free 2021 ICD-10-CM/PCS medical coding reference. www.icd10data.com.
21. Pincus T, Summey JA, Soraci SA Jr, et al. Assessment of patient satisfaction in activities of daily living using a modified Stanford Health Assessment Questionnaire. Arthritis Rheum 1983;26:1346–53.
22. Mikuls TR, Fay BT, Michaud K, et al. Associations of disease activity and treatments with mortality in men with rheumatoid arthritis: results from the VARA registry. Rheumatology (Oxford) 2011;50:101–9.
23. Cannon GW, Mikuls TR, Hayden CL, et al. Merging Veterans Affairs rheumatoid arthritis registry and pharmacy data to assess methotrexate adherence and disease activity in clinical practice. Arthritis Care Res (Hoboken) 2011;63:1680–90.
24. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.

25. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016;15: 155–63.
26. Burnham KP, Anderson DR. Multimodel inference: understanding AIC and BIC in model selection. Sociol Methods Res 2004;33:261–304.
27. Groenwold RH, White IR, Donders AR, et al. Missing covariate data in clinical research: when and when not to use the missing-indicator method for analysis. CMAJ 2012;184:1265–9.
28. Kartje R, Dixon BE, Schwartzkopf AL, et al. Characteristics of veterans with non-VA encounters enrolled in a trial of standards-based, interoperable event notification and care coordination. J Am Board Fam Med 2021;34:301–8.
29. Nikiphorou E, de Lusignan S, Mallen C, et al. Prognostic value of comorbidity indices and lung diseases in early rheumatoid arthritis: a UK population-based study. Rheumatology (Oxford) 2020;59: 1296–305.
30. Stouten V, Westhovens R, De Cock D, et al. Having a co-morbidity predicts worse outcome in early rheumatoid arthritis despite intensive treatment: a post hoc evaluation of the pragmatic randomized controlled CareRA trial. Rheumatology (Oxford) 2021;60:3699–708.
31. Ben Tekaya A, Hannech E, Saidane O, et al. Association between Rheumatic Disease Comorbidity Index and factors of poor prognosis in a cohort of 280 patients with rheumatoid arthritis. BMC Rheumatol 2022;6:78.

Arthritis Care & Research
Vol. 75, No. 10, October 2023, pp 2207–2214
DOI 10.1002/acr.25128

American College of Rheumatology
Empowering Rheumatology Professionals

Beliefs, Experiences, and Openness Regarding Dietary Interventions: Data From an Urban Hispanic Population With Rheumatic Disease in the US

Sandy Lee, Jack Rodman, Vera Hsu, and Leanna Wise

Objective. To obtain descriptive data on the beliefs, behaviors, and openness regarding dietary changes for rheumatic diseases in an urban US Hispanic patient population with rheumatic disease as foundational data for future intervention design.

Methods. We distributed a voluntary survey to our primarily Hispanic population at an outpatient rheumatology clinic for 19 weeks. This survey queried individuals' behaviors as they related to dietary intake used for the treatment of rheumatic disease, perceptions of the effect of food groups on rheumatic disease activity, barriers to physician-recommended diets, and willingness to try future interventions. We used descriptive statistics and Pearson's chi-square test to evaluate associations.

Results. More than 40% of survey respondents from our primarily (88%) Hispanic population noted a link between what they ate and their underlying rheumatic disease activity. More than one-third of patients had, at some point, modified dietary intake to affect their rheumatic disease. Vegetables, fruit, and white meats were commonly reported to improve disease, while red meat and processed foods were reported to worsen disease. Barriers to following a prespecified diet included cost and lack of knowledge. More than 70% of respondents indicated willingness to attempt certain eating patterns should it help their underlying rheumatic disease.

Conclusion. In this primarily Hispanic rheumatic disease patient population, many have not only noted a correlation between dietary intake and rheumatic disease activity but are also open to future nutrition-related interventions. As this population experiences poor rheumatic disease outcomes and a high rate of lifestyle-related comorbidities, an intervention to optimize healthy eating patterns would likely be beneficial as well as acceptable.

INTRODUCTION

Autoimmune rheumatic diseases, such as rheumatoid arthritis (RA) and systemic lupus erythematosus (SLE), often require lifetime immunosuppression to treat the underlying disease and prevent irreversible damage. Individuals with chronic diseases, including those with rheumatic diseases, often explore dietary measures, such as dietary supplements or specific eating patterns, as an adjunct to their medication regimens (1). Indeed, data support the use of some naturally occurring dietary compounds, such as polyphenolic compounds and polyunsaturated fatty acids, in the treatment of some rheumatic diseases (2–5). Additionally, diet and caloric balance may play a role in the disease process, with overweight or obese status potentially affecting disease development, disease activity, response to medication, and overall prognosis (6–12). Finally, a notable proportion of select patient populations with rheumatic disease has noted that what they eat subjectively affects their underlying rheumatic disease (13,14). Consequently, both patients and their providers have shown interest in modulating nutritional intake to enhance disease control.

Ethnic and racial minorities, such as the Hispanic population in the US, have long born a variety of health disparities as they relate to rheumatic diseases. These disparities include, but are not limited to, longer time to RA diagnosis, more active RA at any given time point, a greater incidence and severity of SLE, and worse COVID-19 outcomes in the setting of SLE (15–19). Diet-related comorbidities, such as obesity and type 2 diabetes

Supported by the NIH (National Center for Advancing Translational Sciences grants UL1-TR-001855 and UL1-TR-000130).

Sandy Lee, DO, Jack Rodman, MPH, Vera Hsu, MS, Leanna Wise, MD, MPH: University of Southern California, Los Angeles.

Author disclosures are available online at https://onlinelibrary.wiley.com/doi/10.1002/acr.25128.

Address correspondence via email to Leanna Wise, MD, MPH, at Leanna.Wise@med.usc.edu.

Submitted for publication December 19, 2022; accepted in revised form April 4, 2023.

SIGNIFICANCE & INNOVATIONS

- As data suggest that certain eating patterns may be associated with rheumatic disease activity as well as with the incidence and outcome of chronic comorbidities, it is key to understand patients' perceptions of the relationship between diet and rheumatic disease activity, especially in a unique population that has experienced health disparities, such as the Hispanic population in the US.
- These descriptive data on this population's experiences and beliefs regarding diet and rheumatic disease will provide the foundation for a culturally acceptable intervention.
- This is the first study that evaluates the beliefs and behaviors of an urban, lower socioeconomic, primarily Hispanic rheumatic disease population as they relate to the patient-reported relationship between nutrition and rheumatic disease activity.
- The relatively high proportion of individuals noting a link between dietary intake and rheumatic disease activity, as well as those willing to try future nutrition-related intervention, provide preliminary data for the construction of a culturally acceptable and evidenced-based dietary intervention for this specific population.

mellitus, also occur at higher rates in this population relative to the non-Hispanic White population (20,21). Thus, the Hispanic rheumatic disease population in the US may benefit from adjunctive dietary interventions that both treat the underlying rheumatic disease and lessen the burden of nutrition-related comorbidities. For any discrete ethnic or racial group, such as the Hispanic population in the US, it is vital that any proposed interventions not only be evidence-based but also be practical and well accepted from a cultural and social perspective.

To date, there are no data regarding the beliefs, experiences with, and openness to dietary interventions in the Hispanic rheumatic disease population in the US. The patient targets of prior research into patient-reported experiences with diet and rheumatic diseases have been primarily highly educated non-Hispanic White and European populations. This research has queried patients from a sociocultural context that markedly differs from that of the Hispanic population in the US. The opinions of Hispanic rheumatic disease patients in the US are vital to inform evidence-based dietary interventions in this high-risk group, as they would lay the foundation for interventions that would be the most acceptable and sustainable.

Our rheumatology clinic is composed predominantly of a Hispanic, underserved, disadvantaged patient population with a high burden of rheumatic diseases and concomitant comorbidities. In this study, we sought to obtain preliminary cross-sectional survey-reported data regarding dietary habits and interventions for rheumatic disease patients from our Los Angeles County rheumatology outpatient rheumatology clinic. The purpose of this survey collection is to provide preliminary pilot data that will generate the foundation for more detailed hypothesis testing. We hope that this descriptive pilot data can provide us with a foundation from which we can create (with patient input) and implement an acceptable dietary intervention for this unique population.

PATIENTS AND METHODS

Study design. Institutional review board (IRB) exempt approval was obtained by the University of Southern California's IRB. An anonymous survey was created based on the authors' clinical experience and prior literature assessing dietary interventions in chronic autoimmune populations (see Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25128). This survey was created in English and pilot tested for comprehension, acceptability, and length with a bilingual registered dietician, a bilingual physician, and 6 patients from the target population (3 English-speaking; 3 Spanish-speaking, with certified medical interpretation services). After modifications were made to the survey based on the aforementioned individuals' input, the final version was translated into Spanish by a native bilingual research assistant. Feedback from the patients assisted in refining the length of survey questions and the structure of questions involving barriers to physician-recommended dietary patterns and openness to different interventions. The final Spanish version was reviewed by the same bilingual dietician, a bilingual layperson, and a different bilingual physician for accuracy. All bilingual individuals involved in the initial review were native bilingual English and Spanish speakers.

Our survey was distributed during 2 half-day rheumatology clinics for a 19-week period at Los Angeles County + University of Southern California Medical Center (LAC+USC MC) rheumatology clinic. Approximately 1,500 individuals were offered the survey over this time period. Approximately 50% of our rheumatology clinic patients have a primary diagnosis of RA; 25% a diagnosis of SLE; 8% a diagnosis of vasculitis; the remaining 17% a mix of inflammatory myositis, spondyloarthritis, systemic sclerosis, and others. Only 25% of our clinic patients receive their medical care in English; 75% require medical translation services. Of this 75% non–English-speaking population, ~90–95% speak Spanish. On average 85% of our clinic self-identifies as Hispanic and is of Mexican or other Central American heritage, and 85% are insured by Medi-Cal (California's Medicaid program). Our rheumatology clinic at LAC+USC MC is staffed by 5–6 attending physicians per half day, along with 5–6 rheumatology fellows and 4–7 internal medicine residents.

Self-reported demographic characteristics were obtained for each survey; categorical National Institutes of Health terminology was used for self-reported race and ethnicity. The questions were designed to explore patients' prior dietary/nutritional modifications (if any), patients' perceptions as to whether or not there were certain foods that affected their disease activity, barriers to physician-recommended diets, and willingness to try prespecified

Table 1. Patient-reported demographic characteristics (n = 308)*

Variable	Value
Age (n = 305)	
18–30	29 (9.5)
31–40	45 (14.8)
41–50	80 (26.2)
51–60	88 (28.9)
61–70	55 (18.0)
71+	8 (2.6)
Sex (n = 300)	
Male	48 (16.0)
Female	252 (84.0)
Race (n = 191)	
White	87 (45.6)
Black/African American	7 (3.7)
Asian American	16 (8.4)
American Indian/Alaska Native	0 (0.0)
Native Hawaiian/Pacific Islander	1 (0.5)
Mixed race	16 (8.4)
Other†	64 (33.5)
Ethnicity (n = 302)	
Hispanic	266 (88.1)
Non-Hispanic	32 (10.6)
Prefer not to state	4 (1.3)
Preferred language (n = 307)	
English	92 (30.0)
Spanish	209 (68.1)
Other	6 (2.0)
Education (n = 300)	
Less than high school	107 (35.7)
High school equivalent	149 (49.7)
Bachelor's degree or above	44 (14.7)
Marital status (n = 302)	
Single/widowed	115 (38.1)
Married or in a domestic partnership	135 (44.7)
Divorced/separated	52 (17.2)
Diagnosis (n = 294)	
Lupus	66 (22.5)
Rheumatoid arthritis	133 (45.2)
Other/>1	95 (32.3)

* Values are the number (%). A total of 308 surveys were collected; each variable assessed had missing patient-reported data.
† Respondents who indicated "other" self-identified as ethnically Hispanic and noted a national identity for race (i.e., Mexican or El Salvadorian).

dietary patterns based on patients' prior inquiries regarding vegetarian/vegan diet, a Mediterranean diet, a ketogenic diet, and prior research in fasting. The questions were all close ended and approximated to a sixth-grade reading level; the final question was open ended and allowed for free-text answers. Physical copies of the surveys were offered to all outpatient rheumatology patients at the Los Angeles County/University of Southern California Department of Health Services outpatient rheumatology clinic from April 19, 2022, through August 30, 2022. All patients, regardless of ethnicity, race, or language preference, were invited to take the survey upon check-in to the clinic; however, the survey was only distributed in English and Spanish. No compensation was offered for completing the survey.

Statistical analysis. Surveys were collected from 308 individuals after each rheumatology clinic, and data were manually entered into REDCap, a Health Insurance Portability and Accountability Act–compliant data management system. Two authors (SL and LW) double checked each survey entry in REDCap. All survey questions were categorical and were summarized using frequency and percentage. Associations between categorical variables of interest were evaluated using Pearson's chi-square or Fisher's exact test, as appropriate. To adjust for multiple comparisons and control false discovery rate, Benjamini-Hochberg correction was used. The acceptable false discovery rate was set to 0.05 (5%). All tests were 2-sided, and *P* values less than or equal to 0.05 were considered statistically significant. Post hoc power/sample size analysis is demonstrated in see Supplementary Appendix B, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25128. Analyses were performed in R, version 4.2.1, and SAS 9.4 for Windows.

RESULTS

A total of 308 surveys were collected (Table 1). Patients were predominantly female (84.0%) and of Hispanic ethnicity (88.1%). Most patients selected Spanish as their preferred language (68.1%), and 85.3% of patients had a high school diploma or less. The most represented rheumatic diseases were RA (45.2%) and SLE (22.5%).

More than 40% of respondents noted that there was a link between dietary intake and perceived rheumatic disease activity (Table 2). Many (35.6%) reported either current or past supplement use to affect their rheumatic disease activity, and 39.3% of patients reported current or past adherence to a certain eating pattern to affect their rheumatic disease activity (Table 2). Respondents who were 18–40 years of age reported the highest proportion of current or past diet use for disease management (42.8%) compared to those 41–60 years (40.0%) and ≥61 years (31.4%) (*P* = 0.004; see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25128). The top 3 barriers reported to adherence to a

Table 2. Patient-reported experiences with diet and rheumatic disease*

Variable	Value
What I eat affects my rheumatic disease	
Yes	116 (40.3)
No	47 (16.3)
I don't know/unsure	125 (43.4)
Current/past supplement use to treat disease	
Yes, current	54 (20.2)
Yes, past	41 (15.4)
Never	172 (64.4)
Current/past adherence to special diet to treat disease	
Yes, current	65 (24.8)
Yes, past	38 (14.5)
Never	159 (60.7)

* Values are the number (%).

prespecified diet were cost (25.4%), knowledge/understanding of how to shop or cook for a special diet (21.9%), and taste relative to prior ways of eating (14.1%) (Figure 1). Of note, 32% of respondents indicated the absence of perceived barriers to a recommended eating pattern. Finally, almost 71% of total respondents stated that a rheumatologist had never discussed diet in the context of their rheumatic disease. There was a statistically significant relationship between the patient's preferred language and prior dietary discussions with a rheumatologist, with more non-English speaking patients denying such a discussion ($P < 0.0001$; see Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25128). There was also a correlation between educational attainment and dietary discussions with a rheumatologist, with significantly more individuals with a bachelor's degree or greater having had such a discussion ($P = 0.018$; see Supplementary Figure 1, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25128).

Certain foods were noted to improve disease, most commonly vegetables (67.9% of all respondents), fruit (62.3%), and white meat protein (chicken, fish; 62.0%) (Figure 2). Regarding foods that worsened disease, the most frequently noted food groups were fast/processed food (dessert, fast food, soda; 41.2%), red meat protein (beef, pork; 35.1%), and alcohol (25.0%) (Figure 2).

Interestingly, significantly more female than male patients noted that fast/processed foods (44.4% versus 22.9%; $P = 0.031$), red meat protein (37.7% versus 22.9%; $P = 0.030$), and grains and dairy ($P = 0.031$ and 0.030, respectively) worsened disease activity (see Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25128). Similarly, female patients numerically noted more foods that improved their disease, but the differences were not statistically significant.

In all patients, 50.5% noted improvement in joint/muscle pain, 50% noted improvements in overall energy, and 33.3% noted improvement in sleep (Table 3). Similar rates of symptom improvement were seen in patients who noted that what they eat effects their disease (47.4%, 42.2%, and 26.7%, respectively). A significantly higher proportion of these patients reported that vegetables (79.3% versus 59.9%; $P = 0.006$), fruits (70.7% versus 57.0%; $P = 0.049$), white meat protein (72.4% versus 53.5%; $P = 0.006$), and grains (46.7% versus 32.0%; $P = 0.044$) improved their disease, compared to patients who said that diet had no effect (see Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25128).

Our survey also included several questions assessing patients' willingness to pursue particular eating patterns should they help their rheumatic disease. We assessed willingness to eat more fruits and vegetables, willingness to trial a vegetarian diet, willingness to eat more white meat protein and less red meat protein, willingness to trial a brief period of fasting for a few days per month (a "fasting-mimicking diet"), and willingness to reduce simple carbohydrate intake. The overwhelming majority (>70% for all aforementioned dietary patterns) of patients indicated a willingness to try a specific eating pattern (Table 4). Finally, over 60% noted interest in participating in a future study to evaluate the effect of nutrition on rheumatic disease activity (Table 4). More specifically, a significantly higher proportion of patients 18–40 years of age indicated interest in participating in future studies (62.7%) compared to those who were 41–60 years (60.2%) and ≥61 years (57.1%) ($P = 0.025$; see Supplementary Table 1, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25128). Patients who reported that diet does affect their disease were also more willing to participate in future studies compared to those who said what they eat had no effect (66.0% versus 57.4%; $P = 0.007$; see Supplementary Table 3, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25128).

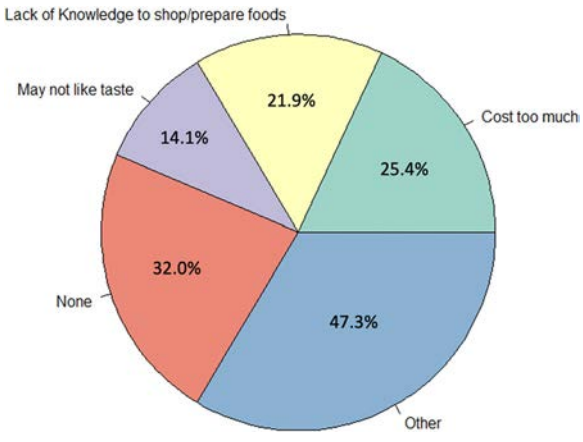

Figure 1. Patient-reported barriers to adherence to a prespecified diet. Patients could select >1 barrier.

DISCUSSION

Individuals diagnosed with a chronic disease often explore the use of diet to quell disease activity, and patients living with rheumatic disease are no exception. Indeed, nearly 50% of RA patients use some form of dietary supplementation, and data from the general population suggest that dietary supplementation is frequently used without any input from, or disclosure to, the treating providers (22–24). Our results demonstrate that more than one-third of our predominantly Hispanic rheumatic disease population has explored dietary modifications or supplementation to treat their underlying rheumatic disease.

Our data also provide insights into several key themes. Some commonly consumed food groups (fruits, vegetables, and white meats) were noted to improve rheumatic disease activity, whereas others (red meats, fast/processed foods, alcohol) were noted to worsen rheumatic disease activity. While several barriers

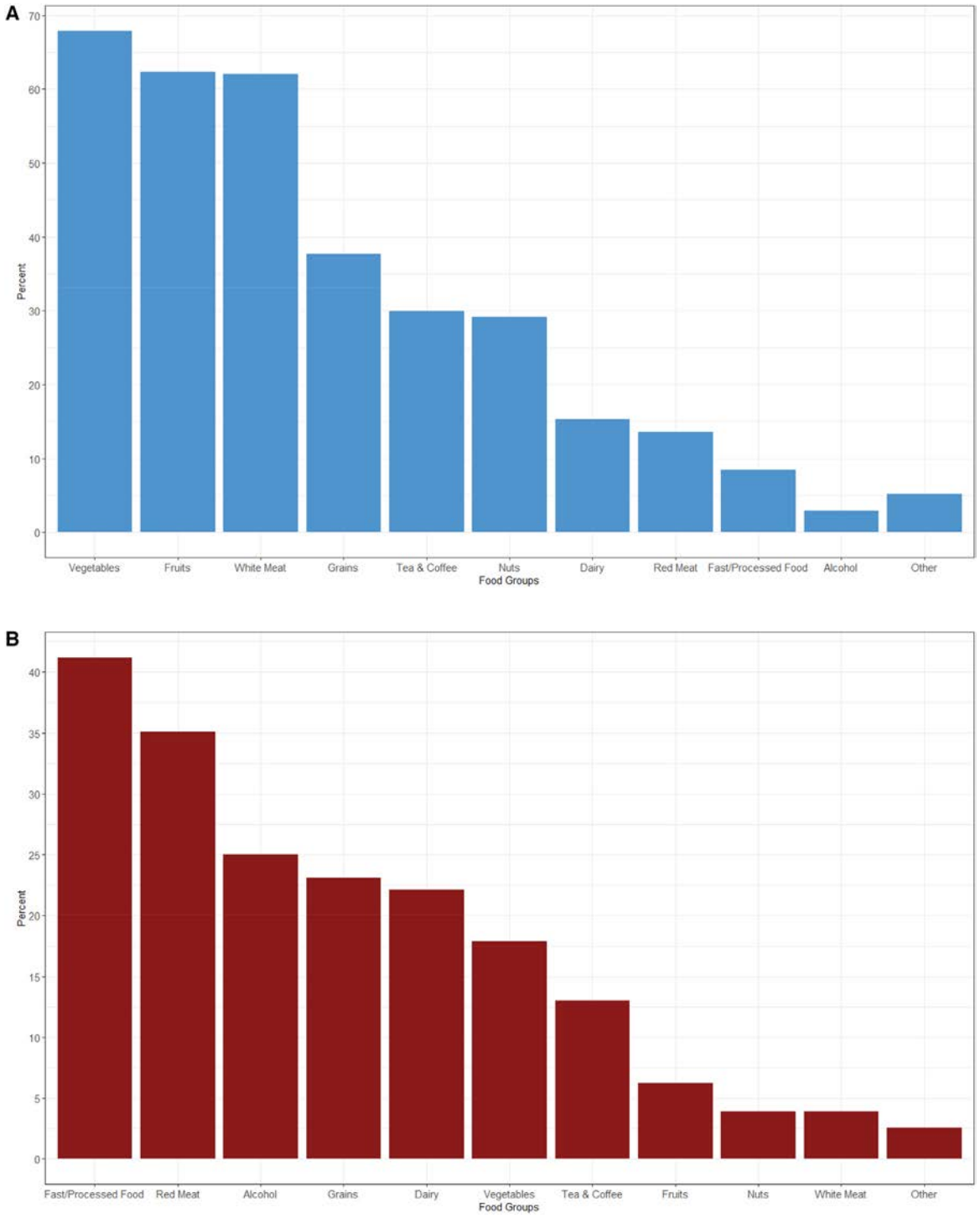

Figure 2. Top foods that improved disease for all patients (**A**) and top foods that worsened disease for all patients (**B**). Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25128/abstract.

were noted in response to following a recommended diet (should one be recommended by a rheumatologist), the vast majority of our patients were very open to dietary interventions as an adjunct to rheumatic disease treatment, with the majority also noting a willingness to participate in future interventional studies.

Prior survey-based research in autoimmune populations has also shown a patient-reported association between nutrition and underlying disease activity, as well as an interest in dietary modulation, including in classic rheumatic diseases such as psoriatic arthritis and non-rheumatic diseases such as inflammatory bowel disease and multiple sclerosis (13,14,25–29). Many of these

Table 3. Patient-reported symptoms that improved with dietary changes*

Variable	Value
Symptoms that have improved	
Overall energy	93 (50.0)
Headaches	39 (21.0)
Mood	40 (21.5)
Sleep	62 (33.3)
Joint/muscle pain	94 (50.5)
Stomach problems	52 (28.0)
Other	11 (5.9)
Missing, no.	122

* Values are the number (%) unless indicated otherwise.

participants noted improvement in disease activity from higher intake of fruits and vegetables and from decreased consumption of processed foods. A large online-based survey (n = 1,206) of primarily non-Hispanic White, highly educated patients with psoriasis (almost 44% of whom had psoriatic arthritis) noted similar trends to our study: a beneficial effect of fruit and vegetable intake on psoriasis activity, and a detrimental effect seen from sugar consumption. However, in contrast to our patients' responses, red meat intake was not mentioned as frequently as was in our population, in regard to foods that worsen patient-reported disease activity. Interestingly, non-White race was associated with a beneficial disease response to avoiding red meat and the addition of fruits. The most commonly reported barriers to dietary adherence in this study were very different than ours, as lack of willpower and time constraints were prominent in this psoriasis study, while our population's most commonly-reported barriers were cost and lack of knowledge in regard to food preparation. Of note, all prior survey-based nutrition studies in autoimmune disease patient populations were markedly different from ours in terms of geography, race/ethnicity, and socioeconomic status, with the vast majority being non-Hispanic White and of relatively high socioeconomic status.

Interestingly, many patients in our study reported that overall energy levels, pain, sleep, and mood improved with dietary manipulation, suggesting that these quality of life (QoL) domains could benefit from nutritional interventions in rheumatic disease patients. This is especially important, as rheumatic disease patients, such as those with SLE, often have unmet QoL issues that do not correlate well with measured disease activity or use of immunosuppression (30,31). Indeed, prior studies of nutrition and/or dietary supplementation on patient-reported outcomes and QoL measures in rheumatic disease patients have found that nutritional intake favorably affects a variety of QoL domains, particularly in SLE and RA (32–34). Furthermore, not only should dietary changes favorably affect patient-reported disease activity and QoL, but they also should meaningfully affect coexisting comorbidities. Hypertension, dyslipidemia, and type 2 diabetes, which can be modulated by nutrition, are all highly prevalent in the Hispanic population (35). Accordingly, a patient-centered approach to dietary optimization should be considered in rheumatic disease patients, especially those with persistent QoL deficits and/or the aforementioned comorbidities.

Prior research suggests that various dietary modifications, including a highly antiinflammatory diet (the "ITIS" diet), fasting, a plant-based diet, and a low-purine diet may quell both objective and subjective disease activity in common rheumatic disease conditions such as RA, SLE, and gout, respectively (25,36–38). The mechanism behind nutritional intake and disease activity is likely complex, but modulation of the gut microbiome via macro- and micronutrients found in food, along with downstream effects on gut wall permeability and systemic immune function, likely plays an important role (reviewed in [39]).

It has already been documented that "cultural competence" is essential in introducing a dietary intervention into the Hispanic community (40,41). Of note, the use of specific dietary patterns and/or supplements has been observed in the evaluation of other Southern California Hispanic populations. One cross-sectional survey (n = 179) in a Los Angeles Hispanic population similar to ours found high rates of herbal medicine use (57%) and dietary supplement use (21%), with many patients using complementary or alternative medicine modalities for pain control and low energy (42). Another survey of a Hispanic population in Southern California (n = 318) also found that ~90% of respondents used some kind of herbal medicine (43). Still another survey in a

Table 4. Patient-reported willingness for future interventions

Variable	Value
I would be willing to eat more fruits and vegetables if it made my disease better	
Disagree/strongly disagree	39 (15.9)
Neither agree nor disagree	6 (2.5)
Agree/strongly agree	200 (81.6)
I would be willing to try a vegetarian diet if it made my disease better	
Disagree/strongly disagree	45 (18.8)
Neither agree nor disagree	24 (10.0)
Agree/strongly agree	170 (71.1)
I would be willing to eat more lean protein and less red meat if it made my disease better	
Disagree/strongly disagree	37 (15.3)
Neither agree nor disagree	13 (5.4)
Agree/strongly agree	192 (79.3)
I would be willing to try a low-calorie diet for a few days a month if it made my disease better	
Disagree/strongly disagree	41 (17.1)
Neither agree nor disagree	9 (3.8)
Agree/strongly agree	190 (79.2)
I would be willing to cut back on bread, rice, & tortillas if it made my disease better	
Disagree/strongly disagree	44 (18.4)
Neither agree nor disagree	24 (10.0)
Agree/strongly agree	171 (71.6)
I would be willing to be part of a future study to see if certain foods, herbal supplements, or eating patterns help my disease	
Disagree/strongly disagree	50 (21.2)
Neither agree nor disagree	43 (18.2)
Agree/strongly agree	143 (60.6)

primarily Hispanic Southern California population (n = 150) demonstrated frequent use of vitamins or supplements (32%), herbal medicine (29%), and dietary/nutritional therapy (26%) (44). Indeed, the Hispanic population in the US has the fastest growing rates of use of complementary and alternative medicine modalities (45). The aforementioned studies, along with our data, suggest that evidence-based, culturally appropriate nutritional interventions will be welcome in our Hispanic rheumatic disease population.

Limitations of our research include the fact that our survey only measured patient-reported data, so selection bias may have played a role in the type of responses gathered. Further, those who attended our clinic and completed the survey may not be representative of the Hispanic rheumatology patient population as a whole. However, based on our average clinic census, those who took the survey are similar to our rheumatic disease clinic population as a whole in demographic characteristics and proportions of primary rheumatic disease diagnoses. As this survey was largely descriptive in nature, we also did not have a control group. We would also caution that the Hispanic community in the US is highly diverse, so results from this particular Los Angeles County population cannot necessarily be extrapolated to all Hispanic populations. While the sample size for the current study was limited, the effects shown are within the range of minimum detectable differences achievable with sufficient statistical power (80%). The differences presented here are also comparable to those presented in similar studies, showing that these associations may still be meaningful in highlighting patient groups that are more willing to participate in future studies and use diet to manage their rheumatic disease symptoms (25,28).

Our data demonstrate that our population is very willing to implement dietary changes to affect their underlying rheumatic disease. These foundational data can be used for future hypothesis generation (i.e., determination of those individuals who are most receptive to a dietary intervention; derivation of the most culturally acceptable dietary intervention) studies, as well as intervention implementation in a similar population. This first survey-based study in a primarily Hispanic rheumatic disease patient population can help pave the way for culturally acceptable and evidence-based nutritional interventions for this unique patient population.

ACKNOWLEDGMENTS

We thank Sara Madrigal, Dr. Stavros Savvas, Dr. Martha Delgado, and Janine Souffrant, RD, for their attentive workup on Spanish translation and proofreading our survey, as well as the 6 patients who assisted with refining the survey. We also thank Dr. Richard Panush, Dr. William Stohl, and Dr. Michael Weisman for their input on this manuscript, and Krishan Parekh for his input on literature review.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Wise.

Acquisition of data. Lee, Wise.

Analysis and interpretation of data. Rodman, Hsu, Wise.

REFERENCES

1. Nahin RL, Barnes PM, Stussman BJ. Expenditures on complementary health approaches: United States, 2012. Natl Health Stat Report 2016;95:1–11.
2. Khajehdehi P, Zanjaninejad B, Aflaki E, et al. Oral supplementation of turmeric decreases proteinuria, hematuria, and systolic blood pressure in patients suffering from relapsing or refractory lupus nephritis: a randomized and placebo-controlled study. J Ren Nutr 2012;22:50–7.
3. Proudman SM, James MJ, Spargo LD, et al. Fish oil in recent onset rheumatoid arthritis: a randomised, double-blind controlled trial within algorithm-based drug use. Ann Rheum Dis 2015;74:89–95.
4. Shamekhi Z, Amani R, Habibagahi Z, et al. A randomized, double-blind, placebo-controlled clinical trial examining the effects of green tea extract on systemic lupus erythematosus disease activity and quality of life. Phytother Res 2017;31:1063–71.
5. Zhang C, Ge C, Wang J, et al. Fish oil enhanced the efficacy of low-dose cyclophosphamide regimen for proliferative lupus nephritis: a randomized controlled double-blind trial. Food Nutr Res 2021;65.
6. Kang JH, Xu H, Choi SE, et al. Obesity increases the incidence of new-onset lupus nephritis and organ damage during follow-up in patients with systemic lupus erythematosus. Lupus 2020;29:578–86.
7. Levitsky A, Brismar K, Hafstrom I, et al. Obesity is a strong predictor of worse clinical outcomes and treatment responses in early rheumatoid arthritis: results from the SWEFOT trial. RMD Open 2017;3:e000458.
8. Liu Y, Hazlewood GS, Kaplan GG, et al. Impact of obesity on remission and disease activity in rheumatoid arthritis: a systematic review and meta-analysis. Arthritis Care Res (Hoboken) 2017;69:157–65.
9. Patterson SL, Schmajuk G, Jafri K, et al. Obesity is independently associated with worse patient-reported outcomes in women with systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2019;71:126–33.
10. Schulman E, Bartlett SJ, Schieir O, et al. Overweight, obesity, and the likelihood of achieving sustained remission in early rheumatoid arthritis: results from a multicenter prospective cohort study. Arthritis Care Res (Hoboken) 2018;70:1185–91.
11. Ye D, Mao Y, Xu Y, et al. Lifestyle factors associated with incidence of rheumatoid arthritis in US adults: analysis of National Health and Nutrition Examination Survey database and meta-analysis. BMJ Open 2021;11:e038137.
12. Thomsen RS, Nilsen TI, Haugeberg G, et al. Adiposity and physical activity as risk factors for developing psoriatic arthritis: longitudinal data from a population-based study in Norway. Arthritis Care Res (Hoboken) 2021;73:432–41.
13. Robinson GA, Mcdonnell T, Wincup C, et al. Diet and lupus: what do the patients think? Lupus 2019;28:755–63.
14. Tedeschi SK, Frits M, Cui J, et al. Diet and rheumatoid arthritis symptoms: survey results from a rheumatoid arthritis registry. Arthritis Care Res (Hoboken) 2017;69:1920–5.
15. Contreras G, Lenz O, Pardo V, et al. Outcomes in African Americans and Hispanics with lupus nephritis. Kidney Int 2006;69:1846–51.
16. Burgos PI, McGwin G Jr, Pons-Estel GJ, et al. US patients of Hispanic and African ancestry develop lupus nephritis early in the disease course: data from LUMINA, a multiethnic US cohort (LUMINA LXXIV). Ann Rheum Dis 2011;70:393–4.

17. Lewis MJ, Jawad AS. The effect of ethnicity and genetic ancestry on the epidemiology, clinical features and outcome of systemic lupus erythematosus [review]. Rheumatology (Oxford) 2017;56 Suppl 1: i67–77.
18. Ugarte-Gil MF, Alarcon GS, Seet AM, et al. Association between race/ethnicity and COVID-19 outcomes in systemic lupus erythematosus (SLE) patients from the United States: data from the COVID-19 Global Rheumatology Alliance. Arthritis Care Res (Hoboken) 2023; 75:53–60.
19. Ocampo-Piraquive V, Nieto-Aristizabal I, Canas CA, et al. Mortality in systemic lupus erythematosus: causes, predictors and interventions [review]. Expert Rev Clin Immunol 2018;14:1043–53.
20. Petersen R, Pan L, Blanck H. Racial and ethnic disparities in adult obesity in the United States: CDC's tracking to inform state and local action. Prev Chron Dis 2019;16:E46.
21. Aguayo-Mazzucato C, Diaque P, Hernandez S, et al. Understanding the growing epidemic of type 2 diabetes in the Hispanic population living in the United States [review]. Diabetes Metab Res Rev 2019;35: e3097.
22. DeSalvo JC, Skiba MB, Howe CL, et al. Natural product dietary supplement use by individuals with rheumatoid arthritis: a scoping review. Arthritis Care Res (Hoboken) 2019:71:787–97.
23. Gardiner P, Sadikova E, Filippelli AC, et al. Medical reconciliation of dietary supplements: don't ask, don't tell. Patient Educ Couns 2015; 98:512–7.
24. Tarn DM, Karlamangla A, Coulter ID, et al. A cross-sectional study of provider and patient characteristics associated with outpatient disclosures of dietary supplement use. Patient Educ Couns 2015;98: 830–6.
25. Knippenberg A, Robinson GA, Wincup C, et al. Plant-based dietary changes may improve symptoms in patients with systemic lupus erythematosus. Lupus 2022;31:65–76.
26. De Vries JH, Dijkhuizen M, Tap P, et al. Patient's dietary beliefs and behaviours in inflammatory bowel disease. Dig Dis 2019:37:131–9.
27. Karnoe A, Pedersen LM, Karnsen S, et al. How people with multiple sclerosis experience the influence of nutrition and lifestyle factors on the disease. Disabil Rehabil 2020;42:3504–15.
28. Afifi L, Danesh MJ, Lee KM, et al. Dietary behaviors in psoriasis: patient-reported outcomes from a U.S. national survey. Dermatol Ther (Heidelb) 2017;7:227–42.
29. Pham T, Sokol H, Hailoua B, et al. Immune-mediated inflammatory diseases and nutrition: results from an online survey on patients' practices and perceptions. BMC Nutr 2021;7:38.
30. Schmeding A, Schneider M. Fatigue, health-related quality of life and other patient-reported outcomes in systemic lupus erythematosus [review]. Best Pract Res Clin Rheumatol 2013;27:363–75.
31. Petrocchi V, Visintini E, De Marchi G, et al. Patient experiences of systemic lupus erythematosus: findings from a systematic review, meta-summary, and meta-synthesis. Arthritis Care Res (Hoboken) 2022;74:1813–21.
32. Charoenwoodhipong P, Harlow SD, Marder W, et al. Dietary omega polyunsaturated fatty acid intake and patient-reported outcomes in systemic lupus erythematosus: the Michigan Lupus Epidemiology and Surveillance Program. Arthritis Care Res (Hoboken) 2020;72: 874–81.
33. Arriens C, Hynan LS, Lerman RH, et al. Placebo-controlled randomized clinical trial of fish oil's impact on fatigue, quality of life, and disease activity in systemic lupus erythematosus. Nutr J 2015;14:82.
34. Philippou E, Petersson SD, Rodomar C, et al. Rheumatoid arthritis and dietary interventions: systematic review of clinical trials. Nutr Rev 2021;79:410–28.
35. Mozaffarian D, Wilson PW, Kannel WB. Beyond established and novel risk factors: lifestyle risk factors for cardiovascular disease. Circulation 2008;117:3031–8.
36. Coras R, Martino C, Gauglitz J, et al. Rheumatoid arthritis improvement after exposure to an anti-inflammatory "ITIS" diet is associated with changes of gut microbiome and systemic metabolome [abstract]. Arthritis Rheumatol 2020;72 Suppl 10. URL: https://acrabstracts.org/abstract/rheumatoid-arthritis-improvement-after-exposure-to-an-anti-inflammatory-itis-diet-is-associated-with-changes-of-gut-microbiome-and-systemic-metabolome/.
37. Müller H, de Toledo HF, Resch KL. Fasting followed by vegetarian diet in patients with rheumatoid arthritis: a systematic review [review]. Scand J Rheumatol 2001;30:1–10.
38. Danve A, Sehra ST, Neogi T. Role of diet in hyperuricemia and gout [review]. Best Pract Res Clin Rheumatol 2021;35:101723.
39. Dehner C, Fine R, Kriegel MA. The microbiome in systemic autoimmune disease: mechanistic insights from recent studies [review]. Curr Opin Rheumatol 2019;31:201–7.
40. Lindberg NM, Stevens VJ, Halperin RO. Weight-loss interventions for Hispanic populations: the role of culture. J Obes 2013;2013:542736.
41. Whittemore R. Culturally competent interventions for Hispanic adults with type 2 diabetes: a systematic review [review]. J Transcult Nurs 2007;18:157–66.
42. Mikhail N, Wali S, Ziment I. Use of alternative medicine among Hispanics. J Altern Complement Med 2004;10:851–9.
43. Malika NM, Desai AK, Belliard JC. Herbal use and medical pluralism among Latinos in Southern California. J Community Health 2017;42: 949–55.
44. Ho DV, Nguyen J, Liu MA, et al. Use of and interests in complementary and alternative medicine by Hispanic patients of a community health center. J Am Board Fam Med 2015;28:175–83.
45. Ortiz BI, Shields KM, Clauson KA, et al. Complementary and alternative medicine use among Hispanics in the United States [review]. Ann Pharmacother 2007;41:994–1004.

Arthritis Care & Research
Vol. 75, No. 10, October 2023, pp 2215–2222
DOI 10.1002/acr.25104

AMERICAN COLLEGE of RHEUMATOLOGY
Empowering Rheumatology Professionals

REVIEW

Identifying and Managing Nociplastic Pain in Individuals With Rheumatic Diseases: A Narrative Review

Anne E. Murphy,[1] Deeba Minhas,[1] Daniel J. Clauw,[1] and Yvonne C. Lee[2]

Chronic pain is a burdensome and prevalent symptom in individuals with rheumatic disease. The International Association for the Study of Pain classifies pain into 3 descriptive categories: nociceptive, neuropathic, and nociplastic. These categories are intended to provide information about the mechanisms underlying the pain, which can then serve as targets for drug or non-drug treatments. This review describes the 3 types of pain as they relate to patients seen by rheumatology health care providers. The focus is on identifying individuals with nociplastic pain, which can either occur in isolation as in fibromyalgia, or as a comorbidity in individuals with primary autoimmune conditions, such as rheumatoid arthritis and systemic lupus erythematosus. Practical information about how rheumatology health care providers can approach and manage chronic pain is also provided.

Introduction

Patients who are referred to rheumatology frequently report pain as a primary symptom. However, rheumatology health care providers often lack the training to adequately assess and manage pain (1). The objective of this narrative review is to provide an overview of the 3 main categories of pain, highlighting the proposed mechanistic pathways associated with these phenotypes, and to provide rheumatology health care providers with specific tools to aid in the assessment and management of pain among patients with rheumatic diseases.

Types of pain and pain mechanisms

In 2016, an international community of pain researchers classified pain into 3 broad categories: nociceptive, neuropathic, and nociplastic (Figure 1) (2). These categories are defined by clinical descriptors, which suggest possible mechanistic contributors to each individual's pain experience. The intent of classifying pain into these 3 categories was to facilitate communication between stakeholders (e.g., clinicians, patients, researchers) and to enhance approaches for identifying and/or classifying appropriate participants in mechanistic studies. Information from these studies could then lead to the development of new management strategies and guidelines.

Nociceptive pain. Nociceptive pain derives from tissue injury, with subsequent sensation of pain by nociceptors. It is usually well-localized and can be precisely described by patients. Examples include pain from inflammation of joint tissues in active rheumatoid arthritis (RA), tissue ischemia in Raynaud's phenomenon, or pericardial inflammation in pericarditis. Pain resulting from palpation of actively inflamed joints in RA demonstrates the localized nature of nociceptive pain. In general, nociceptive pain responds to peripherally directed treatments such as nonsteroidal antiinflammatory drugs, injections, and surgical interventions, and for acute nociceptive pain, opioids may also be effective.

Neuropathic pain. Neuropathic pain occurs with injury or insult to a peripheral or central nerve. Examples include carpal tunnel syndrome, diabetic or nutritional neuropathy, injury, and/or mononeuritis multiplex. Pain and paresthesia typically follow the distribution of peripheral nerves in a dermatomal distribution. Descriptions of neuropathic pain often include lancinating, episodic, numb, or tingling qualities. In addition to targeting any underlying inflammatory process, treatments directed locally at nerves (e.g., surgery, injections, or topical treatments) or medications that target the central nervous system (CNS) may be useful.

Nociplastic pain. Nociplastic pain describes pain characterized by altered nociceptive processing (e.g., hypersensitivity), suggestive of dysregulation of CNS pain processing pathways (2,3). This term is essentially synonymous with older terms such

[1]Anne E. Murphy, MD, Deeba Minhas, MD, Daniel J. Clauw, MD: University of Michigan Medical School, Ann Arbor; [2]Yvonne C. Lee, MD, MMSc: Northwestern University Feinberg School of Medicine, Chicago, Illinois.

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.25104&file=acr25104-sup-0001-Disclosureform.pdf.

Address correspondence via email to Anne E. Murphy, MD, at aemurph@med.umich.edu.

Submitted for publication October 10, 2022; accepted in revised form February 9, 2023.

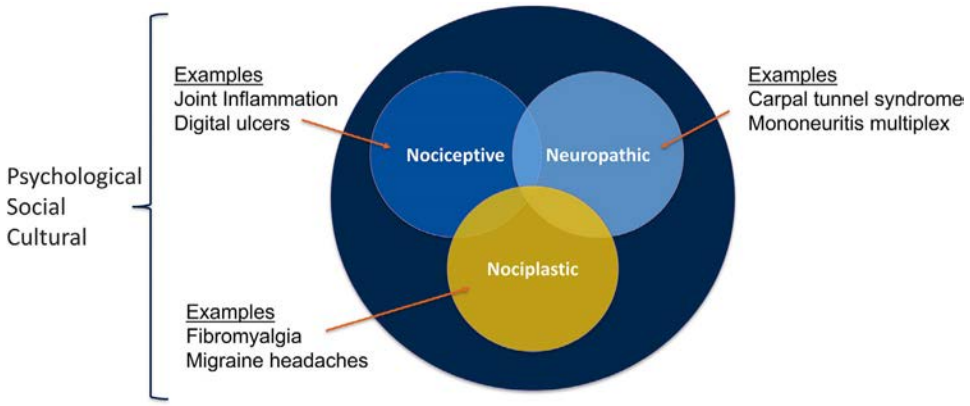

Figure 1. Pain can be classified into 3 main categories: nociceptive, neuropathic, and nociplastic. Each individual's overall pain experience may include ≥1 of these types of pain, which are also modulated by psychological, social, and cultural factors.

as central sensitization or centralized pain. Nociplastic pain is a broad term, which likely encompasses many different CNS pathways that lead to amplified processing of pain signals, decreased inhibition of pain, or both (4). Prototypical nociplastic pain conditions include both widespread (e.g., fibromyalgia [FM]) and localized conditions (e.g., chronic temporomandibular pain disorders [TMDs], chronic primary bladder pain syndrome, irritable bowel syndrome [IBS], tension headaches, and chronic migraine headaches), which are often referred to as chronic overlapping pain conditions (COPCs). In addition to pain, these conditions are frequently associated with fatigue, memory problems, poor sleep quality, and/or mood disturbances. Individuals with nociplastic pain are also often sensitive to nonpainful sensory stimuli (e.g., sensitivity to noises, odors, bright lights). Of these primary nociplastic COPCs, rheumatologists are most frequently asked to diagnose FM, either as an independent condition (i.e., primary FM) or as a condition secondary to an existing autoimmune condition (i.e., secondary FM). What had previously been referred to as primary FM and secondary FM may be associated with different mechanistic pathways, which have been termed "top-down" and "bottom-up," respectively (Figure 2) (5). Treatments include nonpharmacologic therapies and medications that target top-down and bottom-up CNS pathways.

Top-down mechanisms. Top-down mechanisms lead to pain amplification in part via alterations in the descending pain modulatory pathways (6). These pathways include areas in the brain, such as the periaqueductal gray and rostroventromedial medulla, which have descending projections to the spinal cord. The most well-studied descending pathways are endogenous analgesic pathways, which inhibit pain. When these pathways are dysregulated, there is diminished efficacy of the inhibitory pathways, resulting in pain amplification. There are also descending faciliatory pain pathways that increase the gain on pain processing, and these systems are often found to be hyperactive in nociplastic pain states. Nociplastic pain arising from dysregulated top-down mechanisms occurs in the absence of peripheral nociceptive input (e.g., joint inflammation), is frequently familial, and tends to develop at a younger age (7). Studies have shown a stronger genetic component to nociplastic pain in younger as compared to older individuals (8). Additionally, brain imaging in children with nociplastic pain shows similar changes to adults with FM, even preceding the development of pain symptoms (9).

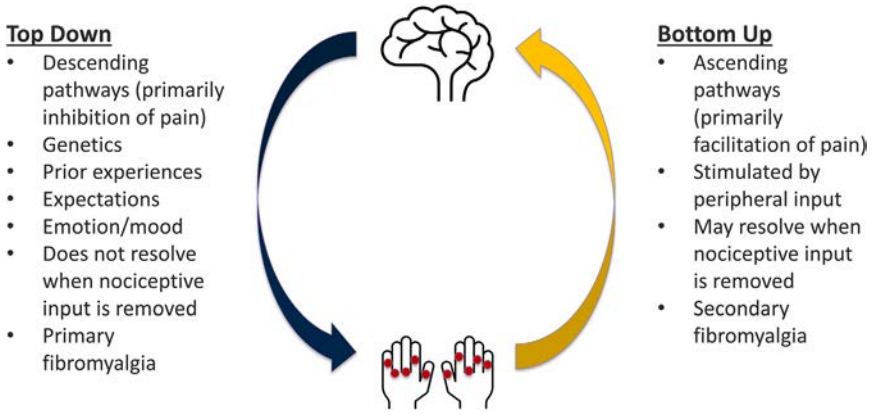

Figure 2. Mechanisms underlying nociplastic pain may be broadly categorized into 2 categories: top-down and bottom-up. Top-down processes are thought to be aberrant in patients with primary fibromyalgia. On the other hand, bottom-up processes are thought to drive secondary fibromyalgia. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25104/abstract.

Bottom-up mechanisms. Bottom-up mechanisms are ascending pathways in the CNS, which are stimulated by peripheral inputs, leading to pain facilitation. Unlike pain associated with top-down processes, pain associated with bottom-up pathways may, at least initially, respond to peripherally directed treatments (10). If the peripheral stimulus is not terminated, however, growth and reorganization of synaptic connections in the nervous system may occur, resulting in a change in the way the CNS functions. This process is termed "neuroplasticity" and may explain the development of secondary FM in patients with rheumatic diseases (5).

Nociplastic pain in patients with rheumatic diseases

Patients with rheumatic conditions and their health care providers often assume that all of their pain is nociceptive, a direct result of joint inflammation. In reality, multiple types of pain may contribute to the overall pain experience. For example, nociplastic pain affects a significant subgroup of patients with rheumatic diseases. Because the term 'nociplastic pain' is relatively new and most studies do not directly assess pain sensitivity, it is not possible to provide prevalence rates of nociplastic pain per se. However, several studies have examined the prevalence of FM, which is considered the prototypical widespread nociplastic pain condition.

In patients with rheumatic diseases, the prevalence of FM ranges from 10% to 48% (Table 1). In comparison, the prevalence of FM is ~2–6% in the general population (11,12). Importantly, the definitions of FM differed in these studies. Before 2010, research studies commonly used the American College of Rheumatology (ACR) 1990 criteria for the classification of FM (13), which required widespread pain in combination with tenderness at ≥11 of 18 tender point sites. However, tender points were infrequently assessed in clinical practice and often not performed correctly. To eliminate the requirement for a tender point examination and place greater focus on symptoms, the ACR published the 2010 preliminary diagnostic criteria for FM (14), which were based on widespread pain, additional symptoms (e.g., fatigue, unrefreshing sleep, cognitive symptoms), and a health care provider's assessment of a detailed list of somatic symptoms. These criteria, however, were not practical for survey-based research. Thus, investigators proposed modifications (15) to the 2011 ACR survey criteria for FM, which replaced the provider's assessment of somatic symptoms with 3 self-reported symptoms (headaches, lower abdominal pain/cramps, and depression). A limitation of these criteria, however, was that patients with regional (not generalized) pain could be diagnosed with FM if they reported multiple areas of pain in the same region. Thus, in 2016, investigators suggested adding a criterion (16), which requires pain in ≥4 of 5 body regions (left upper, right upper, left lower, right lower, and axial). These changes in the definition of FM likely explain some variability in the reported prevalence rates.

Table 1. Prevalence of fibromyalgia (FM) in rheumatic diseases

Rheumatic disease	Prevalence of FM, %	Reference citation
Rheumatoid arthritis	10–48	15,37,71,72,77
Systemic lupus erythematosus	10–22	73,74
Sjögren's syndrome	12–31	73,75,76
Seronegative spondyloarthritis	11–12	73,77
Systemic sclerosis	23–30	78

The variation notwithstanding, the prevalence of FM is clearly higher among patients with systemic rheumatic diseases than the general population, leading to the question of whether FM is really a comorbidity or an intrinsic part of the disease. Support for FM as a comorbidity comes from functional magnetic resonance imaging (fMRI) studies demonstrating that the neurobiological basis of secondary FM is similar to that of primary FM (17). Patients with these conditions exhibit increased functional connectivity between the insula and default mode network, and functional connectivity between the insula and default mode network is associated with severity of FM symptoms (18). This study, however, was cross-sectional and included patients with an average disease duration of >11 years. Thus, it is not clear whether these neurobiological findings were present prior to the development of RA or whether these findings developed in response to peripheral stimuli associated with the development of autoimmunity and inflammation.

Recent fMRI studies suggest that nociplastic pain, associated with FM, may indeed be an intrinsic part of systemic rheumatic diseases, resulting from peripheral stimuli (e.g., systemic inflammation) driving bottom-up processes that trigger the development of nociplastic pain. For example, a recent post hoc analysis showed that sedimentation rate was positively correlated with functional connectivity between the insula and the left inferior parietal lobule (a key component of the default mode network) in patients with RA and concomitant FM but not in patients with RA alone (19). These results suggest that systemic inflammation may lead to changes in functional CNS pathways, leading to the development of nociplastic pain in patients with RA.

Further evidence for the role of peripheral stimuli driving the bottom-up processes comes from a recent study reporting that passive transfer of IgG from patients with primary FM leads to increased sensitivity to noxious mechanical and cold stimulation in mice (20), suggesting that autoreactive antibodies could serve as peripheral, pain-eliciting stimuli. However, these findings do not explain many of the key features of nociplastic pain (e.g., sleep and memory problems and sensitivity to sensory stimuli), and thus the relevance is unclear.

Clinical teaching points in nociplastic pain

The remainder of this review details a practical framework for identifying and managing nociplastic pain among patients in the

outpatient rheumatology setting. At this time, there is no single test or biomarker that can be used to diagnose FM or nociplastic pain. While research studies involving fMRI and single-photon–emission computed tomography have shown differential brain connectivity and neurotransmitter levels between patients with nociplastic pain and matched healthy controls, neither imaging technique is currently utilized for clinical diagnosis (17–19). History and physical examination are the backbone of diagnosis.

History

Assess pain distribution. The location of pain can help distinguish inflammatory versus noninflammatory etiologies. In patients with rheumatic diseases, such as RA and systemic lupus erythematosus (SLE), nociceptive pain tends to localize to joints affected by inflammation and is often associated with joint swelling (21). Neuropathic pain, on the other hand, is characterized by a dermatomal or stocking–glove distribution. In comparison, nociplastic pain is typically more widespread. The distribution of pain can be easily assessed using a body map. For example, a body map is included in the Fibromyalgia Survey Questionnaire, a 1-page assessment based on the 2011 modified ACR preliminary diagnostic criteria for FM (15,22).

Because autoimmune and other diseases typically cause inflammation in certain body regions, studies have also suggested that pain in specific body areas that are not often affected by these processes, such as the upper arms or upper legs, is relatively specific for nociplastic causes of pain. Additionally, widespread pain, often involving the mid-lower back, neck, and arms, had good predictive value in diagnosing nociplastic pain syndromes (21).

Assess for non-pain symptoms. If widespread pain is identified, health care providers should evaluate other symptoms common in nociplastic pain syndromes, including poor sleep, memory problems, fatigue, sensitivity to non-nociceptive stimuli (e.g., visual, auditory, tactile), and mood disturbances. Compared to healthy controls, individuals with FM have a 4-fold higher composite symptom burden, whereas those with RA or SLE (without known coexisting FM) have symptom burdens averaging between healthy control and FM populations (21,23). Of all non-pain symptoms, fatigue and difficulty sitting for prolonged periods were reported as most prominent in those with FM compared to other populations (24). Assessing patients for non-pain symptoms can be a critical and helpful step to identify the presence of nociplastic pain; and these also often represent alternate therapeutic targets.

Past medical history. A large number of comorbid illnesses or allergies may be a clue to nociplastic pain states, such as FM. Patients with FM report more comorbid illnesses (4.5 versus 3.1) and severe allergies (32.6% versus 14.6%) than those with RA (25,26). Further investigation into reported drug allergies may reveal negative skin prick/patch testing results and normal IgE levels, consistent with drug hypersensitivities but not true allergies (27). Thus, these may likely represent hypersensitivities, a hallmark of nociplastic pain states.

Identification of coexisting COPCs can also alert the clinician to increased likelihood of nociplastic pain elsewhere. COPCs encompass several regional nociplastic pain conditions, including tension headaches, migraine headaches, IBS, TMD, and chronic pelvic pain. A systematic review reported that the lifetime prevalence of COPCs (including IBS, migraines, and TMD) in patients with FM ranged from 44% to 57% (28). In comparison, general population rates are estimated at between 5% and 11% for these conditions (29–32).

Medication history. While widespread pain is a hallmark of nociplastic pain, health care providers should continue to consider the possibility of alternate causes of pain. For example, medications may rarely cause myopathies associated with widespread pain due to widespread muscle damage (33–36). In the general population, statins are the medications that most commonly induce myopathy. Additionally, use of amiodarone and immune checkpoint inhibitors should be evaluated. Specific to rheumatic diseases, hydroxychloroquine and glucocorticoids should be considered as potential causes of myopathies. These latter myopathies, however, are usually painless.

Family history. Because nociplastic pain has a genetic component, it is useful to obtain a family history that includes COPCs, including FM. The heritability of FM is higher among individuals ≤50 years of age (23.5%) compared to those >50 years (8.6%). It is theorized that individuals who develop FM at younger ages tend to have a phenotype driven by top-down mechanisms, which are more heritable than phenotypes driven by bottom-up mechanisms (8). Thus, it may be particularly important to assess family history among individuals <50 years of age.

Physical examination

Tender joint count (TJC) and swollen joint count (SJC) difference. The quality of pain may provide additional clues to its source. Compared to individuals with RA and SLE, individuals with FM more often report tenderness, which is likely related to the widespread hypersensitivity that occurs with nociplastic pain (21). While the assessment of FM tender points has largely fallen out of favor, the clinician may leverage the often calculated TJC and SJC to differentiate nociceptive and nociplastic pain in patients with inflammatory arthritis, such as RA. In 2010, Pollard et al reported that a difference in TJC minus SJC of ≥7 predicted the presence of ≥11 tender points, with sensitivity ranging from 72% to 83% and specificity ranging from 80% to 98% (37). In 2014, Kristensen et al proposed a similar measure, the ratio of SJC to TJC as a measure of noninflammatory pain (38). Only 23% of individuals with low SJC to TJC ratios (<0.5) achieved the ACR 50% improvement criteria compared to 39% of individuals with moderate SJC to TJC ratios (0.5–1.0) and 40% of

individuals with high SJC to TJC ratios (>1.0). Thus, health care providers may consider using these thresholds (TJC minus SJC ≥7; ratio of SJC to TJC <0.5) as potential indicators of nociplastic pain in patients with RA.

Trigger points. Distinguishing between local/regional myofascial pain and FM is important because it can impact treatment. When assessing for myofascial pain, clinicians should differentiate tender points from trigger points. Tender points are defined as localized areas of tenderness in a muscle, muscle–tendon junction, fat pad, or bursal region and are classically associated with FM (39). Trigger points are hyperirritable points located within a taut band of skeletal muscle that can either be active, spontaneously inducing pain, or latent. These areas can cause local tenderness, referred pain, and autonomic changes when compressed (40). Trigger points are common and often related to poor body mechanics, postural abnormalities, and chronic strain from repetitive microtrauma seen especially with sedentary, office-based lifestyles (41). Studies have found that 90% of healthy adults reported at least 1 latent trigger point in the scapular positioning muscles (41), and 77.7% reported at least 1 latent trigger point in the lower limbs (42). Trigger points are important to note because they can be significant pain generators. They may respond well to various techniques of myofascial release, including trigger point injections done with dry needling or injectate such as a local anesthetic. In combination with a physical therapy program, trigger point injections can provide long-term relief (40). Muscle relaxants have also been shown to help muscle spasm, but their exact mechanisms in FM are unknown (43).

Bony tenderness. If pain and tenderness in bony regions are identified on examination, alternative diagnoses should be considered, including metabolic bone disease, periostitis, or osteomalacia, encompassing conditions such as hyperparathyroidism and Paget's disease and metabolic, inherited, neoplastic, or paraneoplastic conditions (44,45).

Treatment of nociplastic pain

Patient education. Productive patient–clinician communication is a necessary component of effective pain management. Discussions are often challenging, frustrating, and unproductive for both the patient and provider, highlighting an area in need of improvement (46,47). In the office, it can be helpful to educate the patient on the role of acute pain as an evolutionary local, self-limited, protective measure, while also introducing the idea that unchecked pain can cause a prolonged state of fight-or-flight, contributing to continuous stress chemicals and altered nociception.

Validation and patient buy-in. Clinicians should validate that pain is a very real and very personal experience, while emphasizing that multiple factors (e.g., biological, psychological, social) can influence the intensity and impact of pain. As discussed earlier, pain may or may not be intrinsically related to the underlying autoimmune process. This likely differs from patient to patient and across time within individual patients. Regardless, pain is always an important part of the patient experience.

For patients with an underlying systemic rheumatic disease, it is important to discuss that pain may not always be a direct reflection of inflammation and/or joint damage. Several studies have shown that patients with RA and FM have higher mean Disease Activity Score in 28 joints (DAS28) scores than RA patients without FM, with DAS28 scores often remaining above the lower limit for mild RA disease activity (DAS28 score ≤3.2). These elevations appear to be driven by TJC and patient global assessments, as opposed to SJC and acute-phase reactants (48,49). Since composite disease activity measures appear to be impacted by both nociceptive and nociplastic pain, clinicians should consider the potential influences of nociplastic pain on disease activity assessments and, when nociplastic pain is the primary issue, guide patients to understand that managing the nociplastic pain may be more effective than changing or intensifying immunosuppressive therapy.

Investigators and clinicians who work with patients with SLE have found it useful to categorize signs and symptoms as type 1 or type 2 manifestations (50–52), and this concept may also be applied to other systemic rheumatic diseases. Type 1 manifestations are directly related to inflammatory and autoimmune processes, whereas type 2 manifestations are symptoms (e.g., pain associated with FM) that are often multifactorial in origin and may or may not be directly related to underlying inflammatory processes. Type 2 symptoms can be fluctuant (e.g., coming and going as type 1 symptoms flare and remit), likely representing an inflammatory etiology, or persistent despite type 1 inactivity, more likely representing an FM-like phenotype. Using this framework for conceptualizing disease manifestations may help improve patient–provider communication by validating symptoms while also recognizing differing possible origins of symptoms (e.g., noninflammatory versus inflammatory).

If nociplastic pain is present, confirming its presence, naming it, and describing its clinical picture may help validate the patient's pain experience (53). Specifically, patients with an FM diagnosis have reported dismissive attitudes and lack of acceptance and support from their family, friends, and peers. Such dismissiveness can have a substantial impact on patients who are already distressed (53). In these cases, it is often helpful to acknowledge that skepticism about FM is partly from lack of knowledge regarding its etiology, lack of diagnostic testing, and lack of definitive treatment. It may help patients to understand that FM is not an all-or-none diagnosis. The severity of FM symptoms exists on a continuum and may fluctuate over time. Patients with nociplastic pain may have isolated hyperalgesia and/or a range of symptoms consistent with sensory amplification (e.g., sounds, smells, light), not just pain (54).

Managing priorities and expectations. Patients with chronic pain can feel overwhelmed and singularly focus all their efforts on pain reduction, risking increased frustration, depression, use of ineffective coping strategies, decreased activity levels, and

disengagement with their plan of care (55). Collaborative creation of goals between the patient and provider, with the patient playing an active role in their treatment, increases the likelihood that patients will adhere to physician recommendations and successfully improve their functioning (56). It may help patients avoid the paradigm that pain must be resolved prior to increasing their physical activity.

Nonpharmacologic interventions

Exercise. Exercise and cognitive behavioral therapy (CBT) have the greatest evidence for success in FM (57,58). For pain conditions with a predominance of nociplastic pain, clinicians should consider prescribing aerobic, muscle strengthening, and mind-body exercises (e.g., qigong or tai chi). Supervised sessions with 50- to 60-minute duration, 2–3 times a week, for ≥13 weeks are recommended (59).

CBT. There is a strong bidirectional link between mood disorders and persistent pain. High levels of pretreatment pain catastrophizing (e.g., "This is the worst pain," "I can think of nothing else," and "There's nothing I can do") are highly associated with poor treatment outcomes for pain-relieving interventions (60). Fortunately, pain catastrophizing is modifiable, and patients can be referred for CBT, acceptance commitment therapy, and other psychologic and support services (61).

Sleep. If nociplastic pain is identified, sleep quality may be a high-yield area to address (62). Poor sleep is a predictor of subsequent pain and is noted in 90% of FM patients (63). In studies, nonrestorative sleep was the strongest predictor of chronic widespread pain (64). The severity of sleep disturbance correlates with pain severity, reduced pain inhibition, and fatigue (65). Treatments include adherence to strict sleep hygiene and referral to a sleep clinic or CBT for insomnia in the case of persistent sleep issues.

Support. Higher levels of social support, defined as the perceived available resources from others in social networks, diminish pain severity and improve adjustment in chronic pain conditions (66). It is important to identify the specific type of support that is needed. For example, this could include spousal support, referral to a subspecialist, and/or referral to physical or occupational therapy to address fear of movement.

Tracking. Patients should be encouraged to track progress of self-management behaviors and corresponding symptom response, i.e., "What are behaviors that have the patient feeling good; what are triggers to flares?" On follow-up visits, it is important to assess progress toward goals and barriers. As the patient's situation improves or changes, treatment priorities and goals may change as well.

Pharmacologic interventions

Data regarding the effectiveness of pharmacologic treatments for nociplastic pain among patients with systemic inflammatory diseases are limited. As such, we refer readers to published guidelines for the management of primary FM (67–69).

In this review, we limit our discussion to studies that have examined the efficacy of medications for primary FM in populations with rheumatic disease. The medications with the best evidence base for treating nociplastic pain include low nighttime doses of tricyclic antidepressants (including cyclobenzaprine, which many providers are unaware is a tricyclic), serotonin and norepinephrine reuptake inhibitors (SNRIs; such as duloxetine and milnacipran), and gabapentinoids. Our research group conducted a randomized blinded crossover trial of the SNRI milnacipran for the treatment of pain in patients with RA (70). While the results of the primary analysis were null (no difference between treatment periods), a subgroup analysis including only participants with ≤1 swollen joint revealed significantly lower pain intensity when participants were treated with milnacipran compared to placebo. These results indicate that pain due to refractory inflammation is not effectively treated by treatments for nociplastic pain, whereas pain that remains after effective treatment of inflammation may be reduced using medications targeted at nociplastic pain pathways.

Conclusions

Chronic pain can be burdensome for patients and a difficult symptom for clinicians to treat adequately. Patients with rheumatic disease may have multiple causes and types of pain. Specifically, patients with rheumatic diseases have higher rates of FM, the classic nociplastic pain condition, compared to the general population. In addition, neuroimaging studies support the concept of neuroplasticity, leading to the development of nociplastic pain in patients with RA. While neuroplasticity may lead to nociplastic pain, it is also possible that positive stimuli (e.g., medications, CBT, and exercise) may trigger changes in the CNS that lead to the improvement of nociplastic pain. This review consolidates current evidence on pain mechanisms and serves as a tool for clinicians to use in the office when approaching chronic pain.

AUTHOR CONTRIBUTIONS

All authors drafted the article, revised it critically for important intellectual content, and approved the final version to be published.

REFERENCES

1. American College of Rheumatology Pain Management Task Force. Report of the American College of Rheumatology Pain Management Task Force. Arthritis Care Res (Hoboken) 2010;62:590–9.
2. Kosek E, Cohen M, Baron R, et al. Do we need a third mechanistic descriptor for chronic pain states? Pain 2016;157:1382–6.
3. Kosek E, Clauw D, Nijs J, et al. Chronic nociplastic pain affecting the musculoskeletal system: clinical criteria and grading system. Pain 2021;162:2629–34.

4. Clauw DJ, Arnold LM, McCarberg BH. The science of fibromyalgia. Mayo Clin Proc 2011;86:907–11.
5. Fitzcharles MA, Cohen SP, Clauw DJ, et al. Nociplastic pain: towards an understanding of prevalent pain conditions. Lancet 2021;397: 2098–110.
6. Ossipov MH, Morimura K, Porreca F. Descending pain modulation and chronification of pain. Curr Opin Support Palliat Care 2014;8: 143–51.
7. Coles ML, Weissmann R, Uziel Y. Juvenile primary fibromyalgia syndrome: epidemiology, etiology, pathogenesis, clinical manifestations and diagnosis. Pediatr Rheumatol Online J 2021;19:22.
8. Dutta D, Brummett CM, Moser SE, et al. Heritability of the fibromyalgia phenotype varies by age. Arthritis Rheumatol 2020;72:815–23.
9. Kaplan C, Schrepf A, Mawla I, et al. Neurobiological antecedents of multisite pain in children. Pain 2022;163:e596–603.
10. Schrepf A, Moser S, Harte SE, et al. Top down or bottom up? An observational investigation of improvement in fibromyalgia symptoms following hip and knee replacement. Rheumatology (Oxford) 2020; 59:594–602.
11. Walitt B, Nahin RL, Katz RS, et al. The prevalence and characteristics of fibromyalgia in the 2012 National Health Interview Survey. PLoS One 2015;10:e0138024.
12. Wolfe F, Ross K, Anderson J, et al. The prevalence and characteristics of fibromyalgia in the general population. Arthritis Rheum 1995;38: 19–28.
13. Wolfe F, Smythe HA, Yunus MB, et al. The American College of Rheumatology 1990 criteria for the classification of fibromyalgia: report of the multicenter criteria committee. Arthritis Rheum 1990;33:160–72.
14. Wolfe F, Clauw DJ, Fitzcharles MA, et al. The American College of Rheumatology preliminary diagnostic criteria for fibromyalgia and measurement of symptom severity. Arthritis Care Res (Hoboken) 2010;62:600–10.
15. Wolfe F, Clauw DJ, Fitzcharles MA, et al. Fibromyalgia criteria and severity scales for clinical and epidemiological studies: a modification of the ACR Preliminary Diagnostic Criteria for Fibromyalgia. J Rheumatol 2011;38:1113–22.
16. Wolfe F, Clauw DJ, Fitzcharles MA, et al. 2016 Revisions to the 2010/2011 fibromyalgia diagnostic criteria. Semin Arthritis Rheum 2016;46:319–29.
17. Napadow V, LaCount L, Park K, et al. Intrinsic brain connectivity in fibromyalgia is associated with chronic pain intensity. Arthritis Rheum 2010;62:2545–55.
18. Basu N, Kaplan CM, Ichesco E, et al. Neurobiologic features of fibromyalgia are also present among rheumatoid arthritis patients. Arthritis Rheumatol 2018;70:1000–7.
19. Kaplan CM, Schrepf A, Ichesco E, et al. Association of inflammation with pronociceptive brain connections in rheumatoid arthritis patients with concomitant fibromyalgia. Arthritis Rheumatol 2020;72:41–6.
20. Goebel A, Krock E, Gentry C, et al. Passive transfer of fibromyalgia symptoms from patients to mice. J Clin Invest 2021;131:e144201.
21. Friend R, Bennett RM. Distinguishing fibromyalgia from rheumatoid arthritis and systemic lupus in clinical questionnaires: an analysis of the revised Fibromyalgia Impact Questionnaire (FIQR) and its variant, the Symptom Impact Questionnaire (SIQR), along with pain locations. Arthritis Res Ther 2011;13:R58.
22. Häuser W, Jung E, Erbslöh-Möller B, et al. Validation of the fibromyalgia survey questionnaire within a cross-sectional survey. PLoS One 2012;7:e37504.
23. Bennett RM, Friend R, Jones KD, et al. The revised Fibromyalgia Impact Questionnaire (FIQR): validation and psychometric properties [erratum in Arthritis Res Ther 2009;11:415]. Arthritis Res Ther 2009; 11:R120.
24. Bergman MJ, Shahouri SH, Shaver TS, et al. Is fatigue an inflammatory variable in rheumatoid arthritis (RA)? Analyses of fatigue in RA, osteoarthritis, and fibromyalgia [erratum in J Rheumatol 2010; 39:454]. J Rheumatol 2009;36:2788–94.
25. Wolfe, F, Anderson, J, Harkness, D, et al. A prospective, longitudinal, multicenter study of service utilization and costs in fibromyalgia. Arthritis Rheum 1997;40:1560–70.
26. Wolfe F, Hawley DJ. Evidence of disordered symptom appraisal in fibromyalgia: increased rates of reported comorbidity and comorbidity severity. Clin Exp Rheumatol 1999;17:297–303.
27. Fleming KC, Volcheck MM. Central sensitization syndrome and the initial evaluation of a patient with fibromyalgia: a review. Rambam Maimonides Med J 2015;6:e0020.
28. Kleykamp BA, Ferguson MC, McNicol E, et al. The prevalence of psychiatric and chronic pain comorbidities in fibromyalgia: an ACTTION systematic review. Semin Arthritis Rheum 2021;51:166–74.
29. Wilson S, Roberts L, Roalfe A, et al. Prevalence of irritable bowel syndrome: a community survey. Br J Gen Pract 2004;54:495–502.
30. Canavan C, West J, Card T. The epidemiology of irritable bowel syndrome. Clin Epidemiol 2014;6:71–80.
31. Natoli JL, Manack A, Dean B, et al. Global prevalence of chronic migraine: a systematic review. Cephalalgia 2010;30:599–609.
32. Slade G, Durham J. Prevalence, impact, and costs of treatment for temporomandibular disorders. Appendix C. In: Yost O, Liverman CT, English R, et al, editors. Temporomandibular disorders: priorities for research and care. Washington (DC): National Academies Press; 2020.
33. Doughty CT, Amato AA. Toxic myopathies. Continuum (Minneap Minn) 2019;25:1712–31.
34. Estes ML, Ewing-Wilson D, Chou SM, et al. Chloroquine neuromyotoxicity. Clinical and pathologic perspective. Am J Med 1987;82: 447–55.
35. Stanton MM, Samii L, Leung G, et al. Amiodarone-induced neuromyopathy in a geriatric patient. BMJ Case Rep 2020;13:e236620.
36. Moreira A, Loquai C, Pföhler C, et al. Myositis and neuromuscular side-effects induced by immune checkpoint inhibitors. Eur J Cancer 2019;106:12–23.
37. Pollard LC, Kingsley GH, Choy EH, et al. Fibromyalgic rheumatoid arthritis and disease assessment. Rheumatology (Oxford) 2010;49: 924–8.
38. Kristensen LE, Bliddal H, Christensen R, et al. Is swollen to tender joint count ratio a new and useful clinical marker for biologic drug response in rheumatoid arthritis? Results from a Swedish cohort. Arthritis Care Res (Hoboken) 2014;66:173–9.
39. Borg-Stein J, Stein J. Trigger points and tender points: one and the same? Does injection treatment help? Rheum Dis Clin North Am 1996;22:305–22.
40. Appasamy M, Lam C, Alm J, et al. Trigger point injections. Phys Med Rehabil Clin N Am 2022;33:307–33.
41. Lucas K, Rich P, Polus B. How common are latent myofascial trigger points in the scapular positioning muscles? J Musculoskelet Pain 2008;16:279–86.
42. Zuil-Escobar J, Martinez-Cepa C, Martin-Urrialde J, et al. The prevalence of latent trigger points in lower limb muscles in asymptomatic subjects. PMR 2016;8:1055–64.
43. Brown BR Jr, Womble J. Cyclobenzaprine in intractable pain syndromes with muscle spasm. JAMA 1978;240:1151–2.
44. Glaser DL, Kaplan FS. Osteoporosis: definition and clinical presentation. Spine (Phila Pa 1976) 1997;22 Suppl:12S–6S.
45. Francis RM, Selby PL. Osteomalacia. Baillieres Clin Endocrinol Metab 1997;11:145–63.

46. Upshur CC, Bacigalupe G, Luckmann R. "They don't want anything to do with you": patient views of primary care management of chronic pain. Pain Med 2010;11:1791–8.
47. Simon LS. Relieving pain in America: a blueprint for transforming prevention, care, education, and research. J Pain Palliat Care Pharmacother 2012;26:197–8.
48. Ton E, Bakker MF, Verstappen SM, et al. Look beyond the disease activity score of 28 joints (DAS28): tender points influence the DAS28 in patients with rheumatoid arthritis. J Rheumatol 2012;39: 22–7.
49. Andersson ML, Svensson B, Bergman S. Chronic widespread pain in patients with rheumatoid arthritis and the relation between pain and disease activity measures over the first 5 years. J Rheumatol 2013; 40:1977–85.
50. Pisetsky DS, Clowse ME, Criscione-Schreiber LG, et al. A novel system to categorize the symptoms of systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2019;71:735–41.
51. Rogers JL, Eudy AM, Pisetsky D, et al. Using clinical characteristics and patient-reported outcome measures to categorize systemic lupus erythematosus subtypes. Arthritis Care Res (Hoboken) 2021;73: 386–93.
52. Pisetsky DS, Eudy AM, Clowse ME, et al. The categorization of pain in systemic lupus erythematosus. Rheum Dis Clin North Am 2021;47: 215–28.
53. Arnold LM, Gebke KB, Choy EH. Fibromyalgia: management strategies for primary care providers. Int J Clin Pract 2016;70:99–112.
54. Sluka KA, Clauw DJ. Neurobiology of fibromyalgia and chronic widespread pain. Neuroscience 2016;338:114–29.
55. Filoramo MA. Improving goal setting and goal attainment in patients with chronic noncancer pain. Pain Manag Nurs 2007;8:96–101.
56. Schrooten MG, Vlaeyen JW, Morley S. Psychological interventions for chronic pain: reviewed within the context of goal pursuit. Pain Manag 2012;2:141–50.
57. Williams DA, Cary MA, Groner KH, et al. Improving physical functional status in patients with fibromyalgia: a brief cognitive behavioral intervention. J Rheumatol 2002;29:1280–6.
58. Goldenberg DL, Burckhardt C, Crofford L. Management of fibromyalgia syndrome. JAMA 2004;292:2388–95.
59. Ferro Moura Franco K, Lenoir D, Dos Santos Franco FR, et al. Prescription of exercises for the treatment of chronic pain along the continuum of nociplastic pain: a systematic review with meta-analysis. Eur J Pain 2021;25:51–70.
60. Craner JR, Sperry JA, Evans MM. The relationship between pain catastrophizing and outcomes of a 3-week comprehensive pain rehabilitation program. Pain Med 2016;17:2026–35.
61. Schütze R, Rees C, Smith A, et al. How can we best reduce pain catastrophizing in adults with chronic noncancer pain? A systematic review and meta-analysis. J Pain 2018;19:233–56.
62. Finan PH, Goodin BR, Smith MT. The association of sleep and pain: an update and a path forward. J Pain 2013;14:1539–52.
63. Wagner JS, DiBonaventura MD, Chandran AB, et al. The association of sleep difficulties with health-related quality of life among patients with fibromyalgia. BMC Musculoskelet Disord 2012;13:199.
64. Aili K, Andersson M, Bremander A, et al. Sleep problems and fatigue as predictors for the onset of chronic widespread pain over a 5-and 18-year perspective. BMC Musculoskelet Disord 2018;19:390.
65. Choy EH. The role of sleep in pain and fibromyalgia. Nat Rev Rheumatol 2015;11:513–20.
66. Lopez-Martinez AE, Esteve-Zarazaga R, Ramirez-Maestre C. Perceived social support and coping responses are independent variables explaining pain adjustment among chronic pain patients. J Pain 2008;9:373–9.
67. Macfarlane GJ, Kronisch C, Dean LE, et al. EULAR revised recommendations for the management of fibromyalgia. Ann Rheum Dis 2017;76:318–28.
68. Fitzcharles MA, Ste-Marie PA, Goldenberg DL, et al. 2012 Canadian Guidelines for the diagnosis and management of fibromyalgia syndrome: executive summary. Pain Res Manag 2013;18:119–26.
69. Sommer C, Alten R, Bär KJ, et al. Drug therapy of fibromyalgia syndrome: updated guidelines 2017 and overview of systematic review articles. Schmerz 2017;31:274–84. In German.
70. Lee YC, Massarotti E, Edwards RR, et al. Effect of milnacipran on pain in patients with rheumatoid arthritis with widespread pain: a randomized blinded crossover trial. J Rheumatol 2016;43:38–45.
71. Wolfe F, Hauser W, Hassett AL, et al. The development of fibromyalgia – I: examination of rates and predictors in patients with rheumatoid arthritis (RA). Pain 2011;152:291–9.
72. Joharatnam N, McWilliams DF, Wilson D, et al. A cross-sectional study of pain sensitivity, disease activity assessment, mental health and fibromyalgia status in rheumatoid arthritis. Arthritis Res Ther 2015;17:11.
73. Haliloglu S, Carlioglu A, Akdeniz D, et al. Fibromyalgia in patients with other rheumatic diseases: prevalence and relationship with disease activity. Rheumatol Int 2014;34:1275–80.
74. Wolfe F, Petri M, Alarcón GS, et al. Fibromyalgia, systemic lupus erythematosus (SLE), and evaluation of SLE activity. J Rheumatol 2009; 36:82–8.
75. Dubost JJ, Couderc M, Pereira B, et al. Concomitant fibromyalgia in primary Sjögren's syndrome in the French ASSESS cohort: comparison of the ACR 1990 and ACR 2016 criteria, FiRST questionnaire and physician's opinion. Clin Exp Rheumatol 2021;39 Suppl 133: 140–5.
76. Choi BY, Oh HJ, Lee YJ, et al. Prevalence and clinical impact of fibromyalgia in patients with primary Sjögren's syndrome. Clin Exp Rheumatol 2016;34 Suppl 96:S9–13.
77. Fan A, Pereira B, Tournadre A, et al. Frequency of concomitant fibromyalgia in rheumatic diseases: monocentric study of 691 patients. Semin Arthritis Rheum 2017;47:129–32.
78. Perrot S, Peixoto M, Dieudé P, et al. Patient phenotypes in fibromyalgia comorbid with systemic sclerosis or rheumatoid arthritis: influence of diagnostic and screening tests. Screening with the FiRST questionnaire, diagnosis with the ACR 1990 and revised ACR 2010 criteria. Clin Exp Rheumatol 2017;35 Suppl 105:35–42.

Arthritis Care & Research
Vol. 75, No. 10, October 2023, pp 2223–2225

American College of Rheumatology
Empowering Rheumatology Professionals

LETTERS

DOI 10.1002/acr.25071

Rheumatoid arthritis disease activity and hospitalized infection in a large US registry: comment on the article by Yun et al

To the Editor:

Dr. Yun and colleagues are to be commended for their elegant observational study of the effect of rheumatoid arthritis (RA) disease activity on the chance of hospitalized infection (1). They found that disease activity increases the risk of hospitalized infection and identified an independent effect of other factors, including glucocorticoids (GCs) with a dose effect.

I'm interested in disentangling the effects of GCs from that of disease activity. The two, of course, have a complicated relationship, which leads to confounding by indication in most, if not all, observational studies on the adverse effects of GCs (2). This dataset has potential to study the relationship in more detail.

The problem I see is twofold, cross-sectionally and longitudinally. In the first example we can observe two patients with the same disease activity. One patient has been prescribed a GC and the other has not. The patient taking the GC has a "hidden" higher level of disease activity that would only become apparent if the medication were stopped. In the second example, we can observe a patient with stable disease activity but an underlying unstable dose of GC. In this case we are seeing a patient with a varying disease activity level masked by the changes in the GC dose. In both cases, simply controlling for disease activity and GC dose (by including them as factor in the model) does not solve this problem.

Quite apart from this, RA patients taking GCs may be regarded as having more severe disease, possibly introducing the same kind of heterogeneity in the study population that led the authors to exclude patients with high disease activity.

If I understand it correctly, the analysis included GC dose as a time-varying factor, but perhaps the authors should explore introducing a lag period so changes in disease activity after dose adjustments are better reflected.

Also, it would be interesting to introduce an interaction term (GC dose × disease activity) to account for their interdependence. This could result in a more accurate "pure" GC effect that is less confounded by indication.

Of course, the only real way to prevent confounding by indication is to do a trial. In our recent Glucocorticoid LOw-dose in Rheumatoid Arthritis (GLORIA) trial, we found a limited risk of 5 mg/day prednisolone in RA patients ages >65 years, but we did confirm the finding of increased risk of infection (3). In fact, the relative risk for infection classified as a serious adverse event (mostly for hospitalization) in the GC group was 1.62 compared to placebo, almost identical to the 1.61 found in the study by Yun et al.

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.25071&file=acr25071-sup-0001-Disclosureform.pdf.

Maarten Boers, MD, PhD, MSc
eds@amsterdamumc.nl
Amsterdam University Medical Centers,
Vrije Universiteit,
Amsterdam, The Netherlands

1. Yun H, Chen L, Roy JA, et al. Rheumatoid arthritis disease activity and hospitalized infection in a large US registry. Arthritis Care Res (Hoboken) 2023;75:1639–47.
2. Boers M. Observational studies on glucocorticoids are harmful! Lupus Sci Med 2017;4:e000219.
3. Boers M, Hartman L, Opris-Belinski D, et al. Low dose, add-on prednisolone in patients with rheumatoid arthritis aged 65+: the pragmatic randomised, double-blind placebo-controlled GLORIA trial. Ann Rheum Dis 2022;81:925–36.

DOI 10.1002/acr.25075

Reply

To the Editor:

We thank Dr. Boers for his interest and comments (1) on our recently published study evaluating the association between disease activity and hospitalized infection among patients with RA in the CorEvitas registry (2). Dr. Boers raised an important question, and we are pleased to have the opportunity to respond.

As Dr. Boers mentioned, disentangling the independent effects of GCs on disease activity and subsequent adverse events is fraught with challenges. For this reason, we used marginal structural models to address the confounding that may blur the indication for GC use and downstream effects (3,4). This method can account for the time-varying interrelationship between disease activity, RA medication use (including biologics and GC dose), and censoring. Marginal structural models incorporate not only current disease activity and GC dose, but also previous disease activity and other treatments. While Dr. Boers raises concerns that patients taking GCs may have "hidden" disease activity, incorporating previous disease activity and treatments should help address this concern.

Dr. Boers astutely suggests that there should be a lagged period associated with GC dose, and in fact this is the temporal sequence of how information was captured in the CorEvitas

registry and the analysis was conducted. At each visit, both patients and physician were required to document their medication use (including GC dose) prior to the visit and record their disease activity measured at that visit. Serious infections requiring hospitalization were captured using linked Medicare data, occurring after the visit.

The possibility of introducing an interaction term (GC dose × disease activity) to account for their interdependence was also suggested, but we believe that the interpretation of such an interaction term might be problematic. Since our main independent variable of interest is disease activity, interpreting the main effect in the presence of a significant interaction term would be challenging. Additionally, since the purpose of the current study was to evaluate whether disease activity was associated with hospitalized infection, after introducing an interaction term, the research question would become whether the association between disease activity and infection was different for patients with or without steroid use. Finally, because only 253 patients (<10%) in this analysis used glucocorticoids at a dose of >5–10 mg/day, this interaction term may have been underpowered.

We appreciate the findings from the trial that Dr. Boers and colleagues summarized and contrasted to our work. In the GLORIA trial (5), the relative risk for serious infection in the patients taking GCs (versus placebo) was 1.62, almost identical to the risk of 1.61 observed in our study. Several other points of commonality deserve mention. Mean age was 72 years in the GLORIA trial participants versus 69 years in our analysis; patients in the GLORIA trial at baseline had moderate disease activity (Disease Activity Score in 28 joints: mean 4.4 units), and we required all patients at baseline to have moderate disease activity (Clinical Disease Activity Index score: 10–22). Notably, the rate of serious infections was similar in both studies (GLORIA: 7.3 per 100 patient-years; our registry-based analysis: 7.9 per 100 patient-years.

We agree with Dr. Boers that randomized clinical trials remain the gold standard for evidence, but it is impractical to randomly assign patients to different disease activity groups and not possible to conduct a randomized controlled trial to address every clinical question, especially in the setting of rare outcomes as demonstrated by the lack of statistical significance for serious infection in the GLORIA trial. While even advanced causal inference methods such as marginal structural models may be affected by unmeasured confounding, the similar results in our study and in the GLORIA trial lend additional credibility to our work and help justify the ability of rigorous analyses with large datasets to provide information to help guide clinical care.

Dr. Curtis's work was supported by NIH/National Institute of Arthritis and Musculoskeletal and Skin Diseases (grant P30-AR-072583). Dr. George's work was supported by NIH/National Institute of Arthritis and Musculoskeletal and Skin Diseases (grant K23-AR-073931).

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.25075&file=acr25075-sup-0001-Disclosureform.pdf.

Huifeng Yun, MD, PhD
yunn@uab.edu
Jeffrey R. Curtis, MD, MS, MPH
University of Alabama at Birmingham
Birmingham
Michael D. George, MD, MPH
Hospital of the University of Pennsylvania
Philadelphia
Jeffrey Greenberg, MD, MPH
CorEvitas LLC
Waltham, Massachusetts
Leslie R. Harrold, MD, MPH
CorEvitas LLC
Waltham, Massachusetts
and University of Massachusetts Medical School
Boston

1. Boers M. Rheumatoid arthritis disease activity and hospitalized infection in a large U.S. registry: some additional suggestions for analyses: comment on the article by Yun et al [letter]. Arthritis Care Res (Hoboken) 2023;75:2223.
2. Yun H, Chen L, Roy JA, et al. Rheumatoid arthritis disease activity and hospitalized infection in a large U.S. registry. Arthritis Care Res (Hoboken) 2023;75:1639–47.
3. Hernan MA, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. Epidemiology 2000;11:561–70.
4. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology 2000;11:550–60.
5. Boers M, Hartman L, Opris-Belinski D, et al. Low dose, add-on prednisolone in patients with rheumatoid arthritis aged 65+: the pragmatic randomised, double-blind placebo-controlled GLORIA trial. Ann Rheum Dis 2022;81:925–36.

DOI 10.1002/acr.25110

Systemic lupus, immunosuppressives, COVID-19 vaccination, and antibody response: comment on the article by Petri et al

To the Editor:

We would like to share ideas on the article, Effect of Systemic Lupus Erythematosus and Immunosuppressive Agents on COVID-19 Vaccination Antibody Response by Petri et al (1), recently published in *Arthritis Care & Research*. A cohort of health care professionals was employed as a comparison group when Petri et al collected data on COVID-19 infection, immunization history, and COVID-19 antibodies in the Hopkins Lupus Cohort. The authors found that, despite receiving background immunosuppressive medication, patients with systemic lupus erythematosus exhibited lower levels of vaccination immunoglobulin G (IgG) than health care professionals (1). Mycophenolate, tacrolimus, and belimumab dramatically decreased IgG response after

immunization, according to Petri et al. The efficiency of the vaccination was increased by delaying mycophenolate for a week; this delay had a positive clinical impact on vaccine response without triggering clinical flare reactions (1).

To correctly interpret the findings, several factors must be considered. One of the potential complicating factors that could have influenced the outcomes of the initial booster dose was an unusually mild reaction. Without specialized laboratory testing, a link between asymptomatic COVID-19 and the absence of symptoms cannot be established, however, in the absence of specialized laboratory testing, a link between asymptomatic COVID-19 and the absence of clinical symptoms may exist (2). A silent COVID-19 infection must be ruled out if neither the recent clinical symptoms nor the current clinical signs are present. Cross contamination with an undiagnosed SARS-Co-V2 infection cannot be ruled out completely. Each person's immune system appears to react differently to the COVID-19 vaccine, depending on inherited genetic variability (3). The results and clinical suggestions of the current investigation need to be confirmed by additional clinical research to the current report.

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.25110&file=acr25110-sup-0001-Disclosureform.pdf.

Amnuay Kleebayoon
amnuaykleebai@gmail.com
Private Academic Consultant
Samraong, Cambodia
Viroj Wiwanitkit
Chandigarh University
Punjab, India
and Joesph Ayobabalola University
Ikeji-Arakeji, Nigeria

1. Petri M, Joyce D, Haag K, et al. Effect of systemic lupus erythematosus and immunosuppressive agents on COVID-19 vaccination antibody response. Arthritis Care Res (Hoboken) 2023;75:1878–85.
2. Joob B, Wiwanitkit V. Letter to the editor: coronavirus disease 2019 (COVID-19), infectivity, and the incubation period. J Prev Med Public Health 2020;53:70.
3. Čiučiulkaitė I, Möhlendick B, Thümmler L, et al. GNB3 c.825c>T polymorphism influences T-cell but not antibody response following vaccination with the mRNA-1273 vaccine. Front Genet 2022;13:932043.

DOI 10.1002/acr.25109

Reply

To the Editor:

In our study (1), we excluded both health care workers and patients with systemic lupus erythematosus who had a history of COVID-19 infection so the observed antibody levels could be attributed to vaccinations. We agree with the implicit point that a small number of patients with a history of unrecognized COVID-19 infection might have been included in both groups.

Asymptomatic COVID-19 infection has been reported in immunocompromised individuals with systemic autoimmune diseases (2), but comprehensive studies are lacking. The frequency of asymptomatic COVID-19 infections among those screened for COVID-19 was very low in one study, 0.25% (3). We agree that observational studies of COVID-19 have limitations. However, we see no reason why asymptomatic infections would have biased our findings regarding the relative antibody levels between the two groups. In addition, this could not have affected any of the within-lupus analyses results.

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.25109&file=acr25109-sup-0001-Disclosureform.pdf.

Michelle Petri, MD, MPH
mpetri@jhmi.edu
Daniel Joyce, BS
Kristin Haag, MS
Andrea Fava, MD
Daniel W. Goldman, PhD, MCS
Diana Zhong, MD
Shaoming Xiao, MS
Aaron Milstone, MD, MHS
Johns Hopkins University School of Medicine
Baltimore, Maryland
Laurence S. Magder, PhD, MPH
University of Baltimore School of Medicine
Baltimore, Maryland

1. Petri M, Joyce D, Haag K, et al. Effect of systemic lupus erythematosus and immunosuppressive agents on COVID-19 vaccination antibody response. Arthritis Care Res (Hoboken) 2023;75:1878–85.
2. Fung M, Babik JM. COVID-19 in immunocompromised hosts: what we know so far. Clin Infect Dis 2021;72:340–50.
3. Ma Q, Liu J, Liu Q, et al. Global percentage of asymptomatic SARS-CoV-2 infections among the tested population and individuals with confirmed COVID-19 diagnosis: a systematic review and meta-analysis. JAMA Netw Open 2021;4:e2137257.