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 guide-lines, health care economics, health care policy, educational, social, and public health issues, and future trends in rheumatology practice.

Volume 74 No. 2 February 2022

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 74 • February 2022 • NO. 2

Cover image: The figure on the cover (from Weiner-Well et al, page 171) is a radiograph of the chest showing mildly increased interstitial markings and a right-sided infiltrate.

Arthritis Care & Research
Vol. 74, No. 2, February 2022, pp 165–170
DOI 10.1002/acr.24368

American College of Rheumatology
Empowering Rheumatology Professionals

CLINICOPATHOLOGIC CONFERENCE

A Zebra at the Rodeo: Dyspnea, Hematuria, and a Family History of Arthritis

Javier S. Cabrera-Pérez, Justin Branch, Anaid Reyes, Mini Michael, Karen W. Eldin, Manuel Silva-Carmona, and Tiphanie P. Vogel

CASE PRESENTATION

Chief symptoms

An 18-year-old woman was admitted for evaluation and management of symptomatic anemia after presenting to her primary care physician with fatigue and shortness of breath.

History of present illness

An 18-year-old, previously healthy woman was admitted for urgent evaluation and management of symptomatic anemia after presenting to her primary care physician with fatigue and shortness of breath. The initial outpatient evaluation was remarkable for hypochromic, microcytic anemia with a hemoglobin level of 6.7 gm/dl.

Three months prior to her admission, the patient developed a persistent cough without wheezing or other symptoms of a preceding upper respiratory infection. Her respiratory symptoms progressed to include dyspnea and gradually worsened despite treatment with amoxicillin and bronchodilators. There was no chest pain. In addition, she reported fatigue that began one month prior to presentation to her physician while she was showing swine at a local livestock competition and rodeo. Her mother recalled the patient having a pale appearance upon returning from the rodeo, though the patient was able to attend school and carry out her afterschool activities.

In the week prior to admission, the patient had noticed her urine was dark but dismissed this as a side effect of the iron tablets she was taking. The day prior to admission, she noticed that both feet appeared swollen and were aching. She reported no previous instances of joint swelling or any fevers, night sweats, rashes, oral or nasal ulcers, alopecia, Raynaud's phenomenon, photosensitivity, epistaxis, hearing loss, dysphagia, or numbness.

Past medical, social, and family history

The patient had no significant medical history prior to presentation, except one hospitalization as an infant for respiratory syncytial virus bronchiolitis. She denied tobacco, alcohol, or recreational drug use and denied sexual activity. She described being very active and working on her family farm every morning and interacting with pigs, chickens, goats, cattle, and lambs.

At the time of her presentation, the patient's younger brother had been under treatment for 5 years for polyarticular juvenile idiopathic arthritis (JIA) with onset at the age of 8 years. Her parents and older sister are healthy. Further questioning revealed her paternal grandmother had severe rheumatoid arthritis (RA) and died from RA-related complications at age 50 years, though specific medical details were not available. Additionally, a paternal uncle was also noted to have a reported diagnosis of RA. The patient also had two older paternal half-brothers; one died young after a motor vehicle collision and the other had a history of JIA during his childhood but was not under medical treatment as an adult. The patient is Mexican American, and her family has been in Texas for more than three generations.

Physical examination

The patient was afebrile and vital signs were unremarkable on presentation. She was a well-developed, well-nourished, tired-appearing female teenager in no apparent distress, but with noticeable pallor. There was no rash, jaundice, or icterus. Oropharynx and nasal mucosa were clear. On auscultation, fine inspiratory and expiratory crackles were heard, most prominently in the right lung base, but she had normal respiratory effort with good air entry and no wheezing. Her extremities were notable for bilateral periarticular ankle swelling, but no synovitis or effusions

Javier S. Cabrera-Pérez, MD, PhD, Justin Branch, BS, Anaid Reyes, BA, Mini Michael, MD, Karen W. Eldin, MD (current address: McGovern Medical School, University of Texas Health Science Center, Houston), Manuel Silva-Carmona, MD, Tiphanie P. Vogel, MD, PhD: Baylor College of Medicine and Texas Children's Hospital, Houston, Texas.

No potential conflicts of interest relevant to this article were reported.

Address correspondence to Tiphanie P. Vogel, MD, PhD: Department of Pediatrics, Division of Rheumatology, Baylor College of Medicine, 1102 Bates Street Suite 330, Houston, TX 77030. Email: tpvogel@bcm.edu.

Submitted for publication December 9, 2019; accepted in revised form June 23, 2020.

were present. The remainder of her physical examination was unremarkable.

Laboratory evaluation, imaging, and interventions

Initial laboratory evaluations are shown in Table 1. In addition to anemia, results were notable for the presence of systemic inflammation (erythrocyte sedimentation rate [ESR] of 91 mm/hour) and multiple positive serologic tests, including antinuclear antibodies (ANAs), rheumatoid factor (RF), anti–cyclic citrullinated peptide (CCP) antibodies, antineutrophil cytoplasmic antibodies (ANCAs), and myeloperoxidase (MPO) antibodies. The ANCA immunofluorescence pattern was inconclusive. Antibodies against red blood cells (RBCs), double-stranded DNA, and extractable nuclear antigens were not detected, and levels of complement 3 and C4 were within normal limits. Urinalysis revealed significant proteinuria 4.6 mg protein/mg creatinine and hematuria (RBCs too numerous to count) with active sediment, including granular and hyaline casts.

Computed tomography of the chest showed patchy ground glass and reticular opacities with septal thickening within the right middle lobe and lingula and throughout the basal segments of both lower lobes (Figure 1). This was superimposed on diffuse, thin-walled cystic changes throughout the lungs (Figure 1). Therefore, the radiographic findings were concerning for both pulmonary hemorrhage and nonspecific interstitial pneumonia, suggestive of the presence of a chronic process despite her recent onset of symptoms. Bronchoalveolar lavage samples of the right middle lobe and lingula revealed blood-tinged fluid. Samples showed abnormal cellularity that comprised 65–80% macrophages (normal range 80–90%), 15–25% lymphocytes (normal range 5–10%), and 5–10% neutrophils (normal range <5%). There were abundant RBCs in the sample, and iron staining revealed occasional hemosiderin-laden macrophages.

Table 1. Laboratory results*

Variables	Initial value	Normal range
WBC, 10^3/μl	8.7	4.5–13.5
Hemoglobin, gm/dl	6.7	12.0–16.0
Hematocrit, %	23.4	36.0–45.0
Mean corpuscular volume, fl	66.3	78.0–95.0
Platelet count, 10^3/μl	377	150–450
Absolute neutrophil count, 10^3/μl	5.1	1.8-8
Absolute lymphocyte count, cells/μl	2,728	1,000–3,900
Absolute reticulocyte count, 10^6/μl	0.191	0.029–0.990
Direct antiglobulin test (Coombs' test)	Negative	Negative
Erythrocyte sedimentation rate, mm/hour	91	<20
C-reactive protein, mg/dl	<0.5	<1
ANA titer	1:1,280	<1:80
ANCA titer	1:320	Negative
Anti-MPO titer, AU/ml	207	<19
Rheumatoid factor, IU/ml	344	<14
Anti-CCP, units	27	0–19
C3 complement, mg/dl	102	86–182
C4 complement, mg/dl	25	17–51

* ANA = antinuclear antibody; ANCA = antineutrophil cytoplasmic antibody; anti-CCP = anti–cyclic citrullinated peptide; anti-MPO = antimyeloperoxidase antibody; WBC = white blood cell.

CASE SUMMARY

The patient is an 18-year-old, previously healthy woman who presented with fatigue and shortness of breath in the setting of anemia and systemic inflammation. She was found to have hematuria and evidence of chronic cystic interstitial lung disease (ILD) and diffuse alveolar hemorrhage (DAH), overall concerning for a pulmonary-renal syndrome. These conditions were observed in the setting of the patient having multiple positive findings on autoantibody testing and a strong family history of autoimmune disease.

DIFFERENTIAL DIAGNOSIS

ANCA-associated vasculitis. The major ANCA-associated systemic vasculitides, microscopic polyangiitis (MPA) and granulomatosis with polyangiitis (GPA), are the result of small vessel vascular inflammation and typically present with pulmonary hemorrhage, glomerulonephritis, and the presence of ANCA autoantibodies (1). In MPA, ANCAs are typically found in a perinuclear immunofluorescence pattern with MPO specificity performed by enzyme-linked immunosorbent assay; in GPA, ANCAs are typically cytoplasmic with proteinase 3 (PR3) specificity. The granulomatous inflammation underlying GPA can also lead to several classic findings—including sinusitis, hearing loss, and pulmonary nodules—which were notably absent in this patient.

Patients with ANCA-associated vasculitis often describe a prodromal illness preceding the onset of fulminant disease by up to several months. While pulmonary and renal findings are the predominant clinical features in ANCA-associated vasculitis, many patients also experience skin, ocular, neurologic, and musculoskeletal involvement, with 40% of patients experiencing inflammatory polyarthritis (2). ANCA-associated vasculitis is uncommon, with ~20 cases per million people, and incidence increases with age, peaking in the seventh decade of life.

Glomerulonephritis in ANCA-associated vasculitis is pauci-immune with focal necrotizing crescents. However, up to half of biopsies have an immune complex deposition, a finding that is associated with increased proteinuria and a higher percentage of crescents (1,3). While the diagnosis of ANCA-associated vasculitis is often confirmed by positive findings on serologic testing, MPO-ANCA can be found in other disorders such as systemic lupus erythematosus (SLE) (1,4).

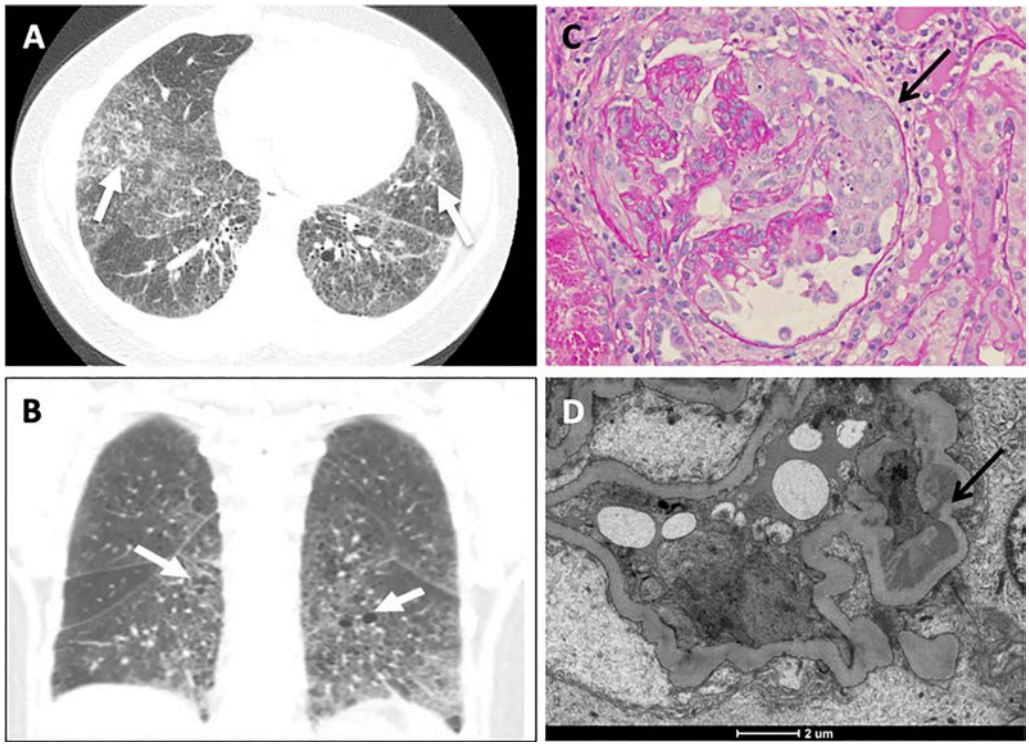

Figure 1. Computed tomography (CT) of the chest and renal histologic findings. **A** and **B**, Patchy ground glass and reticular lung opacities (**arrows**) (**A**) with thin-walled cystic changes found throughout the lungs (**arrows**) (**B**). **C**, Renal biopsy section with periodic acid–Schiff staining, with cellular crescent formation (**arrow**). Original magnification × 400. **D**, Representative electron micrograph from the renal biopsy highlights scattered mesangial electron-dense deposits (**arrow**).

Anti–glomerular basement membrane disease. Anti–glomerular basement membrane (GBM) disease, or Goodpasture's syndrome, is a rare small vessel pulmonary-renal syndrome, with ~1.5 cases per million people reported (5). It is caused by pathogenic autoantibodies against type IV collagen lining the basement membrane of vessels in the lungs and kidney. The classic presentation of this disorder is typically seen in men in their fifth or sixth decades of life, but the clinical presentation of pulmonary hemorrhage and rapidly progressive glomerulonephritis could be otherwise undistinguishable from ANCA-associated vasculitis. In fact, patients with anti–GBM disease often also have positive findings on serologic testing for ANCAs, especially MPO-ANCA. The majority of patients with anti–GBM disease experience prodromal symptoms (i.e., fever, malaise, fatigue) as part of their presentation, but frank polyarthritis is uncommon (6). Overall, the prodromal phase is shorter (days to weeks) than in ANCA-associated vasculitis, unless the patient has dual anti–GBM and ANCA positivity. Dual-positive patients also have an increased rate of relapse compared to anti-GBM single-positive patients. Diagnosis of anti–GBM disease is made by the detection of anti–GBM antibodies in the peripheral blood and/or along the basement membrane of a tissue biopsy, usually renal, showing a linear pattern of IgG immunofluorescence in a patient with renal and/or pulmonary disease.

Systemic lupus erythematosus. As a "great imitator," SLE should be included in the differential when presented with this clinical presentation—namely, a woman of reproductive age with fatigue, anemia, nephritis, and pulmonary hemorrhage who is found to have a positive ANA titer (7). Pulmonary involvement with serositis is a recognized complication in SLE, and cumulatively occurs in over half of lupus patients. SLE with ILD or DAH is more rare, and only 2% of lupus patients have DAH (7). Lupus patients with DAH are more likely to have active nephritis; however, they are also more likely to have positive findings for antiphospholipid antibodies and hypocomplementemia. It is worth commenting that in hypocomplementemic lupus patients, diffuse lung opacities could represent significant infection as the result of decreased fixation and opsonization of microbes. Other common findings in active SLE, including mucocutaneous manifestations and arthritis (found in one- to two-thirds of patients, depending on the method of diagnosis [8]), were absent in this patient.

COPA syndrome. Recently, Watkin and colleagues described five families with an autosomal dominant genetic lung disease caused by mutations in the *COPA* gene (9). Many of the patients also had other autoimmune disease, mainly inflammatory arthritis (75% of patients) and kidney disease (25% of patients),

including glomerulonephritis. Six mutations in the WD-40 domain of COPα, absent from population databases, have been reported to cause COPA syndrome. The presence of a mutation is predicted to lead to impaired retrograde transport between the Golgi and endoplasmic reticulum of cells, a process vital for further posttranslational processing (10). Although the exact pathophysiologic mechanisms behind this mutation are not yet clear, the impairment in cell trafficking increases endoplasmic reticulum stress, which leads to the production of inflammatory Th17 cells and may increase interferon signaling (11), suggesting an additional autoinflammatory component to this rare disorder.

Patients with COPA syndrome tend to have symptoms that manifest clinically with childhood-onset pulmonary disease, but COPA syndrome can also present in adults. Findings on imaging can be concerning for nonspecific interstitial pneumonia and/or DAH, and follicular bronchiolitis has also been observed in lung biopsy samples (12). There is usually a strong family history of autoimmune disease, such as treatment refractory RA, but asymptomatic carriers of *COPA* mutations have also been described. COPA syndrome also displays variable expressivity in regard to extrapulmonary manifestations, even within the same family. Those affected tend to have a chronic relapsing–remitting disease course despite receiving extensive immunosuppressant therapies (9,12).

It should be noted that various viral, bacterial, and parasitic infections can also cause pulmonary hemorrhage and/or glomerulopathy. Therefore, infection was carefully excluded. Anti–GBM disease was thought to be a less likely diagnosis based on the age and the sex of the patient. Similarly, SLE presenting with ILD and pulmonary hemorrhage, but in the absence of musculoskeletal and cutaneous features and with normal complement levels, was considered to be a less likely diagnosis than ANCA-associated vasculitis. The decision was made to proceed with a renal biopsy to help distinguish between these entities and determine a diagnosis.

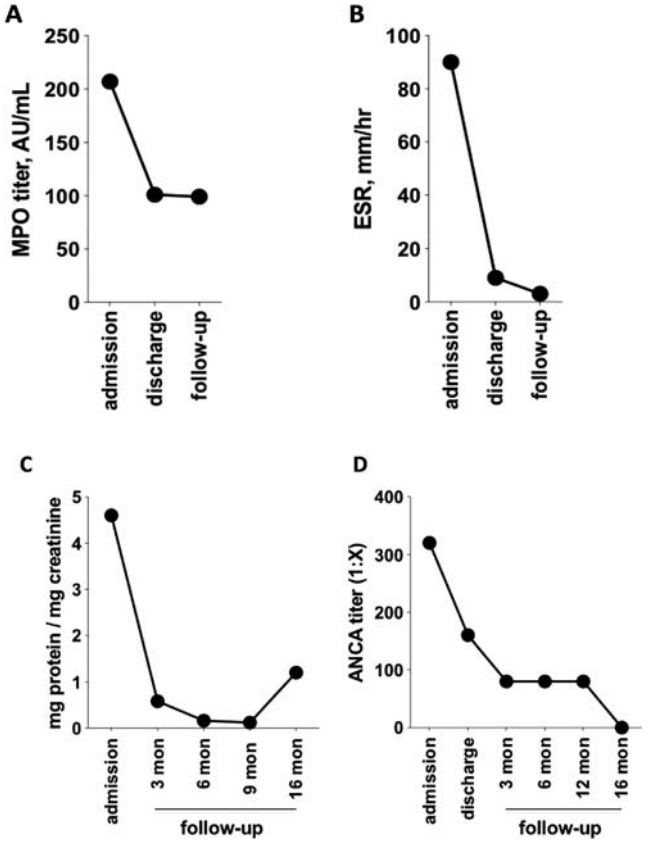

Figure 2. Serial laboratory values of the patient. **A** and **B**, Titers of serum myeloperoxidase (MPO) antibodies and erythrocyte sedimentation rate (ESR) showing rapid response to initiation of therapy. **C** and **D**, Urine protein/creatinine ratio and antineutrophil cytoplasmic antibody (ANCA) titers demonstrating steady improvement and stabilization over time, from clinical presentation to recent follow-up.

CLINICAL COURSE

A percutaneous biopsy of the kidney showed crescentic glomerulonephritis (up to 40% predominantly cellular crescents) on a background of global and segmental glomerulosclerosis (15% of glomeruli) and mild interstitial fibrosis and tubular atrophy (Figure 1). Immunofluorescence staining revealed low-to-moderate intensity for mesangial IgM, IgG, and C3. Staining for C1q and IgA was negative. Immune complex deposition was confirmed by electron microscopy (Figure 1). Immunofluorescence testing was negative for linear IgG, and results were negative for serum anti–GBM antibodies. Furthermore, no tubuloreticular inclusions were appreciated on ultrastructural examination to suggest lupus nephritis.

After receiving three intravenous pulses of methylprednisolone at doses of 1,000 mg each, the patient was started on a slow glucocorticoid taper. Additionally, after being given a working diagnosis of ANCA-associated vasculitis MPA subtype, she received 375 mg/m^2 of rituximab in addition to six cycles of therapeutic plasma exchange, which were well-tolerated except for one episode of facial hives. Significant improvement in laboratory parameters was observed immediately after onset of therapy. Particularly, ESR and MPO titers were promptly lower (Figure 2). Creatinine levels were elevated at the patient's hospital admission (0.94 mg/dl), but quickly declined and stabilized over the course of her hospitalization (baseline 0.61–0.72 mg/dl). She received a periprocedural transfusion of packed RBCs early in the course of hospitalization. The symptoms and clinical findings of anemia subsequently improved and stabilized during the patient's hospital admission but did not normalize until four weeks after discharge. As the patient responded to induction well, she was discharged home with the following prescribed care instructions: a renal diet, a course of rituximab at vasculitis dosing (4 total doses of 375 mg/m^2) in addition to oral glucocorticoids, lisinopril (to control proteinuria and mild hypertension), and prophylactic trimethoprim/sulfamethoxazole.

Due to the presence of rheumatic disease spanning three generations of a single family, including individuals with

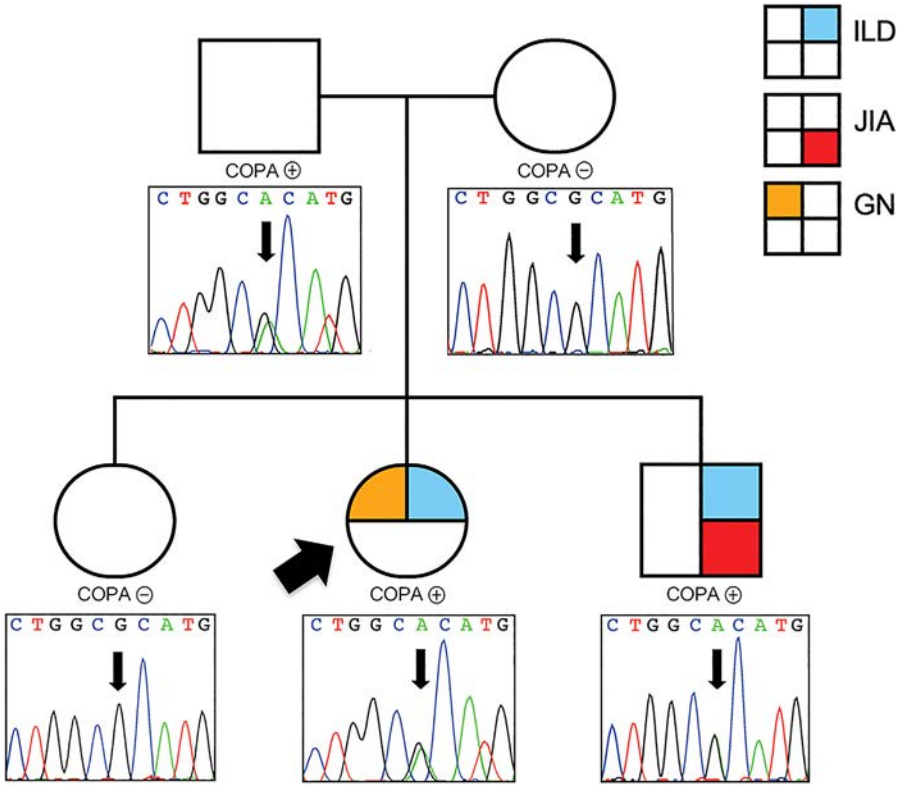

Figure 3. Inheritance of the *COPA* gene variant in available family members of the patient. The patient (**large arrow**) was found to have a pathogenic mutation (**small arrows**) in the *COPA* gene (c.G689A, p.R233H), which was inherited from her asymptomatic father. The patient's brother, who had been diagnosed with polyarticular juvenile idiopathic arthritis (JIA), was subsequently tested and found to carry the mutation as well. Additional clinical evaluation of the brother revealed the presence of interstitial lung disease (ILD), but no sign of kidney involvement. Additional paternal relatives were unavailable or declined genetic testing. There were no pertinent disease findings reported in the patient's maternal relatives. GN = glomerulonephritis.

early-onset, treatment-refractory, and fatal disease, there was suspicion for a genetic disorder, particularly COPA syndrome. Trio whole-exome sequencing was obtained during the patient's admission. Results returned 8 weeks after discharge with a known pathogenic mutation in the *COPA* gene (p.R233H), inherited from the patient's asymptomatic father (Figure 3).

DISCUSSION

This patient with COPA syndrome and DAH with crescentic glomerulonephritis had a presentation indistinguishable from ANCA-associated vasculitis. Arriving at a final diagnosis of ANCA-associated vasculitis in a patient with multi-organ involvement can be challenging. The ANCA-associated vasculitis with MPO can present with findings identical to this patient, and this was the patient's working diagnosis. However, this clinical presentation is not exclusive to MPA, and it is well known that MPO-ANCA is a nonspecific finding that can be present in other inflammatory conditions (7). In this case, the patient's family history led to the correct diagnosis.

Treatment decisions for this patient who had COPA syndrome with clinical, laboratory, imaging, and pathologic features consistent with MPO-positive vasculitis were made prior to her genetic diagnosis based on a diagnosis of ANCA-associated vasculitis (1). Other COPA patients have been treated with a wide spectrum of aggressive immunosuppressive therapies, and some have needed organ transplantation for refractory disease (9,12). Our patient responded rapidly to treatment with glucocorticoids, rituximab, and therapeutic plasma exchange, with complete resolution of her clinical symptoms, anemia, systemic inflammation, and acute kidney injury within 3 to 6 weeks.

Hydroxychloroquine is a frequently used immune modulator that increases the pH of intracellular vacuoles and is suspected to impact protein modification within the Golgi apparatus in addition to other effects (13). Given that it is both generally well-tolerated and could be predicted to directly impact the pathway involved in the pathogenesis of COPA syndrome, hydroxychloroquine was added to the patient's regimen following her genetic diagnosis. After six months, the patient was empirically re-dosed with rituximab, as her B cells had repopulated by then, and completed a year-long taper of oral glucocorticoids. She is currently well 1.5 years after diagnosis and is receiving hydroxychloroquine monotherapy. She has evidence of pulmonary and renal scaring based on imaging of the

chest and persistent proteinuria (Figure 2), but her laboratory parameters are normal, including negative results for ANCAs in the serum (Figure 2). She is closely monitored for any changes in clinical or laboratory parameters.

In animal studies, ANCAs have been confirmed to be pathogenic upon adoptive transfer (1). ANCA antibodies lead to neutrophil activation and subsequent complement pathway induction. Despite ongoing investigations, the inciting or predisposing factors to developing pathogenic ANCA autoantibodies remain elusive. In COPA syndrome, the initial insult is a genetic mutation, ultimately leading to immune dysregulation and the subsequent production of autoantibodies—not just ANCA, but typically also ANA and RF. COPA syndrome has only recently been described. As additional cases are reported, the full phenotypic spectrum and natural history of this disease will become clearer.

Renal biopsy findings in this patient demonstrated moderate intensity immunofluorescence (IgM, IgG, and C3) and immune complex deposition by electron microscopy. Given that immune complexes are present in many cases of "pauci-immune" ANCA-associated glomerulonephritis (3), these findings were felt to be atypical but compatible with that initial diagnosis. Renal biopsy histopathology in COPA syndrome is variable. Reports have described immune complex–mediated disease with histologic impressions of IgA nephropathy and lupus nephritis (12,14), and necrotizing lesions and/or cellular crescents are common (9). Therefore, COPA syndrome should be considered in patients presenting with a pulmonary-renal syndrome.

As exemplified in this case, there can be remarkable similarities between ANCA-associated vasculitis and COPA syndrome. However, upon close inspection, there may be features present in a case of COPA syndrome that do not fit with "classic" ANCA-associated vasculitis, such as in this patient who had multiple additional positive results on serologic testing (ANA, RF, and CCP), an inconclusive immunofluorescence staining pattern despite high titer MPO antibodies, and immune complex deposition in the kidney. There are prognostic and family planning implications in this genetic disease, for which genetic counseling is suggested. Therefore, clinicians should consider a diagnosis of COPA syndrome in the right context. We recommend consideration of COPA syndrome when confronted with a case of childhood-onset ILD, particularly those with concurrent glomerulonephritis presenting as a pulmonary-renal syndrome, when a strong familial history of ILD, inflammatory arthritis, and/or glomerulonephritis has been elicited, or when the patient has a treatment-refractory course of disease.

FINAL DIAGNOSIS

COPA syndrome.

ACKNOWLEDGMENTS

The authors would like to thank the patient and her family members for allowing us to share their story.

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. Vogel had full access to all data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Cabrera-Pérez, Silva-Carmona, Vogel.
Acquisition of data. Cabrera-Pérez, Branch, Reyes, Eldin.
Analysis and interpretation of data. Cabrera-Pérez, Michael, Eldin, Silva-Carmona, Vogel.

REFERENCES

1. Geetha D, Jefferson JA. ANCA-associated vasculitis: core curriculum. Am J Kidney Dis;75:124–37.
2. Houben E, Bax WA, van Dam B, Slieker WA, Verhave G, Frerichs FC, et al. Diagnosing ANCA-associated vasculitis in ANCA positive patients: a retrospective analysis on the role of clinical symptoms and the ANCA titre. Medicine (Baltimore) 2016;95:e5096.
3. Haas M, Eustace JA. Immune complex deposits in ANCA-associated crescentic glomerulonephritis: a study of 126 cases. Kidney Int 2004; 65:2145–52.
4. Choi HK, Liu S, Merkel PA, Colditz GA, Niles JL. Diagnostic performance of antineutrophil cytoplasmic antibody tests for idiopathic vasculitides: metaanalysis with a focus on antimyeloperoxidase antibodies. J Rheumatol 2001;28:1584–90.
5. Segelmark M, Hellmark T. Anti-glomerular basement membrane disease: an update on subgroups, pathogenesis and therapies. Nephrol Dial Transplant 2019;34:1826–32.
6. McAdoo SP, Pusey CD. Anti-Glomerular Basement Membrane Disease. Clin J Am Soc Nephrol 2017;12:1162–72.
7. Lopez Velazquez M, Highland KB. Pulmonary manifestations of systemic lupus erythematosus and Sjogren's syndrome. Curr Opin Rheumatol 2018;30:449–64.
8. Zayat AS, Mahmoud K, Md Yusof MY, Mukherjee S, D'Agostino MA, Hensor EM, et al. Defining inflammatory musculoskeletal manifestations in systemic lupus erythematosus. Rheumatology (Oxford) 2019;58:304–12.
9. Watkin LB, Jessen B, Wiszniewski W, Vece TJ, Jan M, Sha Y, et al. COPA mutations impair ER-Golgi transport and cause hereditary autoimmune-mediated lung disease and arthritis. Nat Genet 2015;47:654–60.
10. Letourneur F, Gaynor EC, Hennecke S, Demolliere C, Duden R, Emr SD, et al. Coatomer is essential for retrieval of dilysine-tagged proteins to the endoplasmic reticulum. Cell 1994;79:1199–207.
11. Volpi S, Tsui J, Mariani M, Pastorino C, Caorsi R, Sacco O, et al. Type I interferon pathway activation in COPA syndrome. Clin Immunol 2018; 187:33–6.
12. Tsui JL, Estrada OA, Deng Z, Wang KM, Law CS, Elicker BM, et al. Analysis of pulmonary features and treatment approaches in the COPA syndrome. ERJ Open Res 2018;4.
13. Fox RI. Mechanism of action of hydroxychloroquine as an antirheumatic drug. Semin Arthritis Rheum 1993;23 Suppl 1:82–91.
14. Boulisfane-El Khalifi S, Viel S, Lahoche A, Fremond ML, Lopez J, Lombard C, et al. COPA syndrome as a cause of lupus nephritis. Kidney Int Rep 2019;4:1187–9.

Arthritis Care & Research
Vol. 74, No. 2, February 2022, pp 171–178
DOI 10.1002/acr.24532

American College of Rheumatology
Empowering Rheumatology Professionals

CLINICOPATHOLOGIC CONFERENCE

Caught Red-Handed

Yonit Wiener-Well,[1] Philip D. Levin,[1] Ephraim Sagi,[1] Eldad Ben-Chetrit,[2] and Eli Ben-Chetrit[1]

CASE PRESENTATION

Chief symptoms

A 54-year-old female patient was admitted in April 2020 for diarrhea, skin rash, and shock.

History of present illness

The patient presented to the emergency room with a mild headache, nausea, vomiting, and diarrhea that began several days earlier. She also complained of diffuse arthralgia affecting the knees, elbows, and shoulders and reported a skin rash on her face, phalanges of both hands, and calves that worsened on the day of admission. Five days earlier, she had experienced a high fever (40°C). She denied exposure to cold environments, animals, or recent travel. She lived in an urban region and did not notice any insect bites. She did recall cleaning fish with possible skin puncture by fish bones several days prior to admission. She denied sore throat, cough, loss of taste or smell, and exposure to individuals who were confirmed to have COVID-19.

Medical, social, and family history

The patient denied comorbidities and chronic medication use. She was not aware of any connective tissue diseases or vascular problems such as Raynaud's phenomenon. She lived with her family, none of whom had symptoms similar to hers. The patient was married and the mother of 12 children. She denied any extramarital relationship. She denied tobacco or illicit substance use.

Physical examination

On physical examination, the patient was alert and oriented without nuchal rigidity. Her blood pressure was 83/51 mm Hg, pulse rate 88 beats per minute, temperature 36.7°C, and O_2 saturation 100% on ambient air. An erythematous skin rash was evident on the fingers and thumbs of both hands, and to a lesser extent, on the palms. The rash blanched with pressure (Figure 1). A discrete maculopapular rash was noted on the lateral sides of her calves and did not extend to her feet. Faint erythema of the face was also present. There were no signs of arthritis. The patient looked pale and mildly tachypneic. Her lungs were clear on auscultation, and heart sounds were regular with no murmurs. Her abdomen was soft with no organomegaly. No leg edema or acrocyanosis was observed. Gynecologic examination was unremarkable, and neurologic examination was normal.

Laboratory and radiographic evaluation

Laboratory findings are shown in Table 1. Levels of hemoglobin were normal. Mild leukocytosis, thrombocytopenia, and mildly elevated results on liver function tests with high levels of creatine phosphokinase were noted. The patient's C-reactive protein (CRP) level was 37.7 mg/dl (reference range 0–0.5 mg/dl). On the second day of admission, elevated levels of brain natriuretic peptide (BNP) (1,434 pg/ml [normal level <100 pg/ml]) and serum troponin (2,809 ng/ml [reference range 0–20 ng/ml]) were detected.

Radiographs of the chest showed mildly increased interstitial lung markings and a right-sided infiltrate (Figure 2). Findings on echocardiogram (EKG) were unremarkable. EKG showed mild-to-moderate left ventricular (LV) dysfunction (ejection fraction of 45%) with mild pericardial and pleural effusions.

Results on the following tests were negative: throat swab specimen by polymerase chain reaction (PCR) for influenza, respiratory syncytial virus, adenovirus, and enteroviruses; blood PCR for parvovirus, stool culture specimens for *Shigella*, *Salmonella*, *Campylobacter*, *Clostridioides difficile*, and rotavirus; five consecutive blood cultures, serology for *Legionella* species, *Mycoplasma*

[1]Yonit Wiener-Well, MD, Philip D. Levin, MB, BChir, Ephraim Sagi, MD, Eli Ben-Chetrit, MD: Shaare Zedek Medical Center, Jerusalem, Israel; [2]Eldad Ben-Chetrit, MD: Hadassah Hebrew University Medical Center, Jerusalem, Israel.

Drs. Wiener-Well and Levin contributed equally to this work.

No potential conflicts of interest relevant to this article were reported.

Address correspondence to Eli Ben-Chetrit, MD, Infectious Disease Unit, Shaare-Zedek Medical Center and Faculty of Medicine, Hebrew University of Jerusalem, Shmu'el Bait Street 12, Jerusalem 9103102, Israel. Email: elibc@szmc.org.il.

Submitted for publication June 24, 2020; accepted in revised form December 1, 2020.

Figure 1. An erythematous skin rash on the fingers and thumb of both hands (**A**), and to a lesser extent, of the palms of the hands (**B**).

pneumonia, Q fever, *Rickettsia conorii, Brucella melitensis*; rapid plasma reagin (RPR); cytomegalovirus; Epstein–Barr virus; hepatitis B virus; hepatitis C virus; and HIV. Results on testing performed for the detection of rheumatoid factor, antinuclear antibody (ANA), antineutrophil cytoplasmic antibody, cryoglobulins, and anticardiolipin antibody (including β_2GPI) were also negative. Serum levels for complement 3 (C3) and C4 were normal. However, the patient's levels of serum ferritin (2,016 ng/ml [normal range 11–205 ng/ml]), fibrinogen (684 mg/dl [normal range 100–500 mg/dl], D-dimer (10,652 ng/ml [normal range 0–500 ng/ml]), and dilute Russell's viper venom time were also elevated.

Three nasopharyngeal swab specimens using different real-time PCR (RT-PCR) testing techniques taken on admissions days 1, 2, and 4 were all negative for SARS–CoV-2 (SeeGene, GeneXpert, and CDC kits).

Table 1. Laboratory values during the patient's hospitalization

Laboratory test	Reference range	Admission	First week	Second week/discharge
WBCs, $\times 10^3$/μl	3.6–10	10.7	17.3	7.1
Neutrophils, %	50–75	96	92.7	72
Hemoglobin, gm/dl	12–16	14.3	10.3	10.6
Hematocrit, %	36–46	39.8	29.7	32.7
Platelet count, $\times 10^3$/μl	150–450	128	181	522
Sodium, mEq/liter	135–145	127	134	138
Potassium, mEq/liter	3.6–5	4.0	4.2	3.5
Calcium, mg %	8.4–10.5	7.5	7.3	8.0
BUN, mg/dl	9–20	19	13	7
Creatinine, mg/dl	0.52–1.04	1.03	0.53	0.50
Glucose, mg/dl	65–105	135	227	98
Alkaline phosphatase, IU/liter	38–150	72	107	80
GGT, IU/liter	12–43	49	44	37
AST, IU/liter	5–34	69	43	25
ALT, IU/liter	9–52	36	21	21
LDH, IU/liter	125–220	362	339	237
CRP, mg/dl	0–0.5	37.75	32.26	7.98
Albumin, gm/dl	3.2–5.2	–	–	2.6
CK, IU/liter	40–150	256	167	–
Troponin, ng/liter	0–20	216	2,809	14
BNP, pg/ml	<100	724.4	1,434	137.7
Cholesterol, mg/dl	100–190	92	–	–
Triglycerides, mg/dl	35–160	226	–	–
Ferritin, ng/ml	11–205	2,016	–	–
Fibrinogen, mg/dl	200–500	552	494	345
D-dimer, ng/ml	0–500	12,144	2,406	1,967

* ALT = alanine aminotransferase; AST = aspartate aminotransferase; BNP = brain natriuretic peptide; BUN = blood urea nitrogen; CK = creatine kinase; CRP = C-reactive protein; GGT = gamma glutamyl transferase; LDH = lactate dehydrogenase; WBCs = white blood cells.

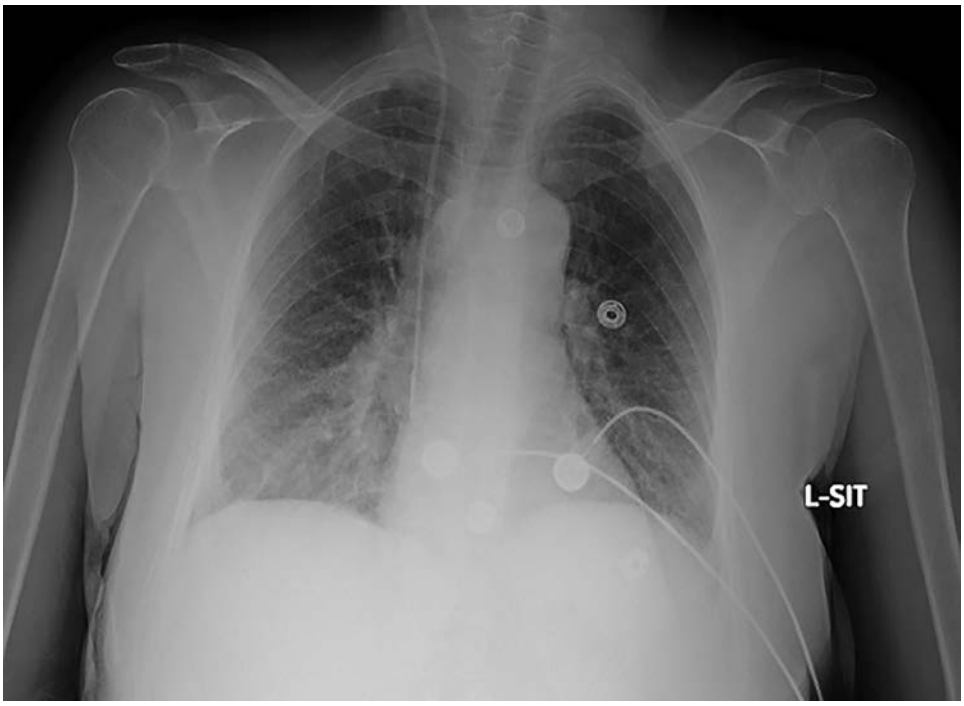

Figure 2. Radiographs of the chest showing mildly increased interstitial markings and a right-sided infiltrate.

CASE SUMMARY

A 54-year-old previously healthy female patient presented with headache, nausea and vomiting, diarrhea, diffuse arthralgia, and an erythematous skin rash over her face, the distal part of her hands, and both calves, with a high fever of five days' duration. Physical examination was notable for hypotension. Laboratory evaluation revealed increased levels of CRP, BNP, troponin, ferritin, and D-dimer. Radiographs of the chest revealed pneumonitis and an EKG showed reduced left ventricular function. Blood and stool cultures showed no growth. PCR testing of nasopharyngeal and throat swabs was negative for enteroviruses, respiratory viruses, and SARS–CoV-2. Serologic testing for atypical respiratory pathogens, rickettsiae, syphilis, hepatitis viruses, Epstein–Barr virus, and cytomegalovirus were all negative.

DIFFERENTIAL DIAGNOSIS

Fever, rash, and hypotension raised the initial differential diagnosis between severe infection and autoimmune disease. Infectious causes considered in the differential included viral or bacterial infections characterized by the combination of gastroenteritis and myocarditis such as toxic shock syndrome (TSS), rickettsial disease, *Vibrio vulnificus*, and COVID-19. Inflammatory or autoimmune disorders were also considered including systemic lupus erythematosus (SLE), sarcoidosis, cryoglobulinemia, and Kawasaki-like disease.

Viral or bacterial infection. Presentation with fever, diarrhea, vomiting, and shock may suggest severe gastroenteritis with fluid loss caused by bacterial pathogens such as *Shigella* species, *Salmonella* species, *Campylobacter* species, or *Clostridioides difficile*. Diarrhea is also commonly caused by viral agents such as enteroviruses (coxsackievirus and echovirus, among others) and rotavirus (1,2). Elevated levels of troponin and BNP along with the EKG findings of decreased LV function and mild pericardial fluid were consistent with a diagnosis of myocarditis. Therefore, enteroviruses, parvovirus, and influenza were also considered. The extremely elevated level of CRP supported a diagnosis of infection. However, the patient denied exposure to ill persons and the local influenza season was already over. Furthermore, the rash on both hands is not a common feature of these infections and repeated negative blood, throat, and stool cultures and PCR did not support the presence of infection.

The history of possible skin puncture by a fish bone raised the suspicion of *Vibrio vulnificus* infection (3). Ingestion of this virulent bacterium can cause acute gastroenteritis and sepsis. It can also cause necrotizing wounds if inoculated through open skin or puncture wounds from the spines of fish. However, negative blood cultures along with the absence of comorbidities known to be risk factors for severe disease (mainly chronic liver disease and immunocompromised states), pneumonitis, and the presence of an atypical rash did not point to this diagnosis of *V vulnificus* infection.

COVID-19 is an infectious disease caused by the SARS–CoV-2 virus. The main clinical features of acute infection with SARS–CoV-2 are fever, sore throat, cough, shortness of breath, hypoxia, diarrhea, and loss of taste and smell. Severe COVID-19 can deteriorate to multiple organ failure (respiratory, cardiac, and kidney failure).

Our patient had fever, headache, malaise, arthralgia, pneumonia, hypotension, myocardial involvement, and a skin rash affecting mainly her hands. In the middle of the pandemic, the possibility of COVID-19 infection was raised early in the course of illness; however, three consecutive PCR tests of nasopharyngeal specimens for SARS–CoV-2 were negative.

Toxic shock syndrome. TSS is caused by bacterial toxins and typically manifests in otherwise healthy individuals (4). Symptoms may include fever, malaise, confusion, skin rash, hypotension, multiple organ failure, and death. There may also be symptoms related to the specific underlying infection, particularly necrotizing fasciitis or pneumonia. TSS is typically caused by *Staphylococcus aureus* or group A *Streptococcus*. Its characteristic rash resembles a sunburn and can involve any region of the body, especially the palms of the hands and soles of the feet. In patients who survive TSS, the rash desquamates after 10 to 21 days. Our patient fulfilled almost all criteria for a diagnosis of TSS (5). She had high fever, hypotension, and a skin rash as well as diarrhea, arthralgia, and mild transaminitis but no renal, hematologic, neurologic, or muscular involvement. There was, however, multiple cultures showed no evidence of staphylococcal or streptococcal infection, including cultures taken prior to therapy with antibiotic. Physical examination did not reveal wounds, insect bites, cellulitis, or a soft tissue abscess, and the patient reported no use of a tampon. Finally, the acral rash was not typical for TSS.

Rickettsial infection. Tick-borne rickettsial infections are caused by intracellular bacteria and belong to the spotted fever group of the genus *Rickettsia* (6). The *Rickettsia conorii* subspecies *conorii*, responsible for causing Mediterranean spotted fever, is endemic in southern Europe. The *Rickettsia conorii* subspecies *israelensis* is transmitted by ticks infected with *Rhipicephalus sanguineus*. Dogs have been suggested to be a competent reservoir of this pathogen. The clinical manifestations of rickettsial infections are similar to those of Mediterranean spotted fever, which includes fever, flu-like symptoms, and a rash spreading to the palms of the hands and soles of the feet that is either maculopapular or petechial in nature. Compared to Mediterranean spotted fever, however, the eschar at the inoculation site is usually lacking, and more significant gastrointestinal manifestations are typically observed. In recent years, atypical and serious life-threatening presentations have been reported. Our patient presented with a febrile illness accompanied by rash, gastrointestinal symptoms, and headache. However, there was no exposure to dogs or other animals, and the palmar erythema was not consistent with a typical rickettsial rash. Finally, serologic testing for rickettsia was ultimately negative.

Systemic lupus erythematosus. SLE can affect the joints (arthritis), brain, lungs (basal fibrosis), kidneys (glomerulonephritis), blood vessels (vasculopathy with thrombosis), and skin (malar rash, chilblain-like erythema) (7). Chilblain lupus erythematosus presents with a rash that mainly affects the acral surfaces, such as the fingers, toes, ears, and nose. The rash is characterized by tender plaques that have a purple discoloration. The patient's presentation included a marked purple rash of both her hands and arthralgia involving her hands and knees. However, her age at presentation was relatively older compared to the individuals who typically present with this condition, and the rash was not that of chilblains. She also had nausea, vomiting, and diarrhea, which are not common presenting features of lupus. Finally, results of an ANA test were also negative. Considering the relatively older age of the patient at disease onset, her presentation with gastrointestinal symptoms, and the negative ANA test, a diagnosis of SLE seemed very unlikely.

Sarcoidosis. Sarcoidosis is a chronic granulomatous inflammatory disease (8). It may involve the lungs (hilar lymphadenopathy and fibrosis), heart (dilated and restrictive cardiomyopathy), eyes (uveitis), and the skin (lupus pernio and erythema nodosum). Lupus pernio is a chronic raised indurated lesion of the skin and is often violaceous. It appears on the nose, ears, cheeks, lips, forehead, and the digits. It is pathognomonic of sarcoidosis. Our patient's illness shared some clinical features with sarcoidosis. She had a pernio-like purple rash, arthralgia, heart, and lung involvement. However, the lung involvement was limited to the middle lobe with no typical hilar lymphadenopathy. The cardiac involvement was perimyocarditis, rather than restrictive or dilated cardiomyopathy.

Cryoglobulinemia. Cryoglobulins can precipitate as clumps leading to vascular thrombosis and peripheral ischemia, including gangrene of the fingers and toes (9). The disease is associated with various malignant, infectious, or autoimmune diseases that are the underlying cause for the production of the cryoglobulins. Since pernio or chilblains are caused by vascular damage following exposure to cold, it is reasonable to look for cryoglobulins in patients with pernios. However, in a retrospective study, only 2 of 79 patients with pernio had cryoglobulins (10). When these patients were further evaluated, it was found that both were positive for ANAs, and one patient had Raynaud's phenomenon, suggesting an underlying autoimmune disease. Our patient had an acral skin rash. However, testing for cryoglobulins was negative, and computed tomography (CT) of the total body showed no evidence of a malignant disorder, and serologic testing for hepatitis C virus (often associated with cryoglobulinemia) was negative.

Kawasaki disease. Kawasaki disease, or mucocutaneous lymph node syndrome, is an illness mainly affecting children under the age of five years, and its etiology remains poorly understood (11). The disorder affects the mucus membranes, lymph nodes, and blood vessels, which causes vasculitis. Patients are usually febrile for at least five days. Later, they develop conjunctivitis, cracked lips with diffuse erythema of the oral mucosa (strawberry tongue), redness, swelling and induration of the hands and/or feet, and cervical lymphadenopathy. Uveitis may be present as well (12). Involvement of the coronary arteries can lead to aneurysms.

Our patient presented with high fever, arthralgia, and acral erythema—all of which may suggest a possible diagnosis of Kawasaki disease. However, the disease typically affects children, and there was no lymphadenopathy or erythema of the lips or oral cavity observed in the patient. The lung involvement and the unique rash as well as diarrhea are not typical features of Kawasaki disease. Finally, echocardiography did not reveal any coronary abnormalities (dilatations or aneurysms), which are typical of Kawasaki disease.

Kawasaki-like illness or multisystem inflammatory syndrome in children (MIS-C) with SARS–CoV-2 has been reported since the emergence of COVID-19 (13,14). This syndrome has also been termed pediatric inflammatory multisystem syndrome temporally associated with SARS–CoV-2 (PIMS-TS) and may fulfill the full or partial criteria for Kawasaki disease. The syndrome manifests with fever, marked inflammation (high levels of CRP and ferritin), and may affect multiple organs including the heart (LV dysfunction, myocarditis, coronary artery aneurysms, and conduction abnormalities), lungs (pneumonia), kidneys (acute renal failure), the gastrointestinal tract (diarrhea), and the skin (rash). Neurologic symptoms (headaches, encephalopathy, neurologic deficits) may also develop. Testing results may be positive for current or recent SARS–CoV-2 infection (by RT-PCR serologic testing), and medical history is usually consistent with COVID-19 exposure within a month prior to the onset of symptoms (Table 2) (15). In contrast to Kawasaki disease, which affects mostly infants and young children, patients with MIS-C may be older (5–14 years) and have more intense inflammation and greater myocardial injury than patients with Kawasaki disease. MIS-C is more common among African American or Hispanic children (in contrast to Kawasaki, which is more common among infants of Asian descent) (14,16). Although our patient was older (in terms of MIS-C), the constellation of symptoms during the COVID-19 pandemic may have implied MIS in an adult (MIS-A). Evidence of recent infection with SARS–CoV-2 may have supported this diagnosis.

Table 2. MIS-C case definition as suggested by the CDC*

1. An individual ages <21 years presenting with fever†, laboratory evidence of inflammation‡, and evidence of clinically severe illness requiring hospitalization, with multisystem (≥2) organ involvement (cardiac, renal, respiratory, hematologic, gastrointestinal, dermatologic, or neurologic); ***AND***
2. No alternative plausible diagnoses; ***AND***
3. Positive for current or recent SARS–CoV-2 infection by RT-PCR, serologic or antigen testing, or exposure to an individual who had suspected or confirmed COVID-19 within 4 weeks prior to the onset of symptoms.

* CDC = Centers for Disease Control and Prevention; CRP = C-reactive protein; ESR = erythrocyte sedimentation rate; LDH = lactic acid dehydrogenase; IL-6 = interleukin-6; MIS-C = multisystem inflammatory syndrome in children; RT-PCR = real-time polymerase chain reaction. The case definition for MIS-C was obtained from a CDC Health Alert Network publication (Ref. 15). Of note, some individuals may fulfill full or partial criteria for Kawasaki disease, but should be reported if they meet the case definition for MIS-C; additionally, MIS-C should be a diagnosis considered in any pediatric death with evidence of SARS–CoV-2 infection.
† Fever of >38.0°C for 24 hours or longer or reported subjective fever lasting 24 hours or longer.
‡ Including, but not limited to, one or more of the following: elevated levels of CRP, ESR, fibrinogen, procalcitonin, D-dimer, ferritin, LDH, or IL-6; elevated number of neutrophils; reduced number of lymphocytes; and low level of albumin.

CLINICAL COURSE

Our patient was hospitalized in a negative pressure isolation room in the intensive care unit with strict airborne precautions. Fluid resuscitation and inotropic support (noradrenaline) were started along with broad-spectrum antibiotics (ceftriaxone, doxycycline, and clindamycin). Ceftriaxone was later changed to ceftazidime to cover possible infection with *Vibrio vulnificus*. On the second day of her hospitalization, the patient became more tachypneic and hypoxemic and required 2 liters per minute of oxygen support by nasal cannula. Serial radiographs of the chest showed pulmonary congestion and bilateral pleural effusions. A pulmonary CT scan disclosed dense alveolar involvement of the middle lobe, suggesting the presence of pneumonia or atelectasis (Figure 3). Bilateral moderate pleural effusions were also evident. There was no hilar lymphadenopathy. No pulmonary emboli were detected. Periportal and subcutaneous edema was noted in the abdomen.

EKG findings of mild-to-moderate LV dysfunction and mild pericardial effusion supported the diagnosis of perimyocarditis with cardiac failure and probable capillary leak syndrome as the cause of the low serum albumin level, abdominal subcutaneous edema, and pleuropericardial effusion. The patient's respiratory status worsened, and she needed high flow nasal cannula oxygen at a flow of 50 liters per minute with a fraction of inspired oxygen level of 0.45. She received noradrenaline to maintain blood pressure, 300 mg daily of hydrocortisone, and low-dose furosemide.

By day 10, the patient's condition had improved remarkably. Her blood pressure had normalized, she was weaned off supplemental oxygen, electrolyte levels had normalized, and BNP, D-dimer, and fibrinogen had gradually decreased to normal levels. Follow-up radiographs of the chest showed improvement. The skin of her hands desquamated (Figure 4), and the rash on the patient's calves and face disappeared.

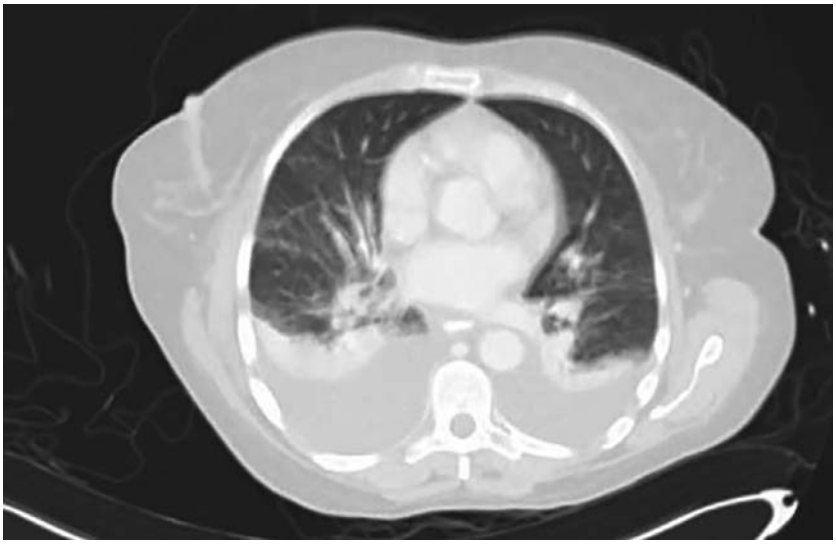

Figure 3. A pulmonary computed tomography scan of the chest disclosing dense alveolar involvement of the middle lobe, suggesting the presence of pneumonia or atelectasis as well as bilateral pleural effusions.

The patient was assessed by a dermatologist who raised the possibility of COVID-19 infection and a syndrome similar to that described as "COVID toes." In the absence of an alternative diagnosis, serum samples for IgG/IgM (Xiamen Wiz Biotech) were sent to a laboratory for COVID-19 testing on day 10 of the patient's hospital admission and were weakly positive. Two days later, repeat serologic testing was performed using the Food and Drug Administration–approved Abbott architect immunoassay (Abbott Diagnostics). Testing results were strongly positive for SARS–CoV-2 IgG antibodies suggesting recent COVID-19 infection. The patient was discharged from the hospital 17 days after admission. A week later, she was readmitted for blurred vision. Ophthalmologic examination revealed mild bilateral anterior uveitis. She received topical steroids and experienced full recovery of the uveitis.

Figure 4. Skin desquamation of the hands during resolution of illness.

DISCUSSION

Our patient presented with hypotension and acral skin erythema of the hands—two uncommon features at the onset of symptomatic COVID-19 infection. These clinical manifestations together with an absent history of SARS–CoV-2 exposure and repeated negative results on PCR tests for the virus made the diagnosis of acute COVID-19 infection less plausible and enhanced the search for other diagnoses.

Shock in a patient with rash, high fever, headache, and malaise suggested the diagnosis of septic shock. However, the lack of positive cultures (blood, urine, and stool) did not support this diagnosis. The occurrence of concomitant perimyocarditis could have supported the diagnosis of an underlying viral infection. Yet PCR testing for common viral agents associated with perimyocarditis was negative. Fever, malaise, arthralgia, rash, and perimyocarditis may also suggest vasculitis. Therefore, a diagnosis of SLE was also considered, but the predominant symptom of shock and the negative serologic test results for ANAs precluded a diagnosis of SLE.

While immediate therapy with volume resuscitation, inotropic support, and antibiotics was indicated, a dilemma arose in regard to adjunctive systemic glucocorticoids for a possible autoimmune/autoinflammatory process such as vasculitis, particularly as infection had not been ruled out. Systemic glucocorticoids have a major role in the treatment of autoimmune diseases for immediate suppression of inflammation and recovery. Additionally, the patient had a very low level of serum albumin with signs of anasarca suggesting "capillary leak syndrome," which may also be an indication for treatment with glucocorticoids. Further, at the time of this patient's admission, the role of steroids in treating COVID-19 was controversial (17,18). Since no objective

support for infection was evident, hydrocortisone was administered intravenously at a dose of 300 mg/day, possibly contributing to the patient's recovery.

We suggest that the patient did not have an acute COVID-19 infection when she was admitted, as repeat PCR testing was negative. SARS–CoV-2 anti-IgG antibodies were identified, so it is possible that the rash, perimyocarditis, and shock on presentation were actually "postinfectious autoimmune features" of SARS–CoV-2 infection. This parallels recent reports of Kawasaki-like illness or MIS-C/MIS-A among children and adults with COVID-19 (14,19–21). The Centers for Disease Control and Prevention has suggested criteria for PIMS-TS or MIS-C (Table 2) (15). The condition of the patient described here fulfilled all criteria except age and demonstrated the possibility of MIS in adults. The late onset of bilateral anterior uveitis further supports the notion of a late autoimmune phenomenon, although no reports of COVID-19–related uveitis in humans have been published.

To date, 27 cases of MIS in adults ages 16–50 years have been reported (21–31). These individuals presented with Kawasaki-like multisystem inflammatory syndrome in the setting of recent SARS–CoV-2 infection based on exposure history and serologic testing. Their symptoms included fever, rash, conjunctivitis, erythema or cracking of the lips, cervical lymphadenopathy, rash, and diarrhea. Cardiovascular involvement (LV dysfunction, myocarditis, shock) was reported in the majority of patients (21–24). Seventeen patients were treated with glucocorticoids. Intravenous immunoglobulin (IVIG) and/or tocilizumab were administered in nine patients, with IVIG and tocilizumab considered as potential treatments for MIS-A and MIS-C (32). In the patient discussed here, IVIG and/or tocilizumab were not administered because by the time MIS-A was considered as a diagnosis, the patient's condition had improved, and the absence of coronary artery aneurysms per EKG precluded the need for IVIG.

Eight (30%) of the 27 patients reported with MIS-A had negative results on PCR testing and positive results on SARS–CoV-2 antibody testing. Notably, in a report by Godfred-Cato et al (33), 45% of 440 children with MIS-C had a negative result on SARS–CoV-2 PCR testing. Almost all patients had positive findings for SARS–CoV-2 antibodies, implying that MIS-A and MIS-C represent postinfectious processes.

The possibility of false-negative PCR results in a symptomatic patient who also has positive results for SARS–Cov-2 IgG antibodies (usually evident after day 7 of infection) is very unlikely. The first serologic assay that was performed was borderline positive. This was the colloidal gold immunochromatography assay (Xiamen Wiz Biotech), which tests IgM/IgG antibodies (without differentiation) and has a sensitivity and specificity of 71.1% and 96.2%, respectively. Two days later, repeat serologic testing was performed using the Food and Drug Administration–approved Abbott architect assay (Abbott Diagnostics), which tested strongly positive. The assay tests only IgG antibodies and has a sensitivity and specificity greater than 99% when performed 14 days or more after symptoms start (34). The positive IgG results on serologic testing along with PCR tests that were repeatedly negative suggest that our patient had had a recent COVID-19 infection from which she had recovered. We suggest that the phenomena described are late immune-mediated manifestations of COVID-19 infection.

The pathogenesis of MIS is not clear. Emerging phenotypes include a combination of Kawasaki disease, TSS, and macrophage activation syndrome or hemophagocytic lymphohistiocytosis syndrome—all reflect dysregulated immune responses (35). Skin biopsy findings in one patient had features typical for Kawasaki disease (nonspecific sparse inflammatory infiltrate) and features suggestive of TSS (few intraepidermal neutrophils with necrotic keratinocytes), providing histologic support for a distinct inflammatory syndrome (22). Cardiac endothelitis and small vessel vasculitis, as well as thrombotic microvascular injury in the lungs, skin, and other organs, and complement deposition in vessels that suggest intense complement activation have been described previously (27). Unfortunately, a skin biopsy was not performed in our patient. However, the desquamation of the skin at the sites of the rash during resolution of illness suggests a Kawasaki-like or TSS-like disorder (as noted above) and supports a diagnosis of MIS. We suggest the notion that the patient was admitted during the postinfectious phase of the disease, with her clinical manifestations related to an immune response to COVID-19 infection.

FINAL DIAGNOSIS

Multisystem inflammatory syndrome in adults, a Kawasaki-like disease temporally related to COVID-19 infection.

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. Eli Ben-Chetrit had full access to all data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Wiener-Well, Levin, Eldad Ben-Chetrit, Eli Ben-Chetrit.

Acquisition of data. Wiener-Well, Eldad Ben-Chetrit, Eli Ben-Chetrit.

Analysis and interpretation of data. Levin, Sagi, Eldad Ben-Chetrit, Eli Ben-Chetrit.

REFERENCES

1. Bányai K, Estes MK, Martella V, Parashar UD. Viral gastroenteritis. Lancet 2018;392:175–86.
2. Moss PJ, McKendrick MW. Gastrointestinal Infection. Curr Opin Infect Dis 1997;10:402–7.
3. Phillips KE; Satchell KJ. *Vibrio vulnificus*: from oyster colonist to human pathogen. PLoS Pathog 2017;13:e1006053.

4. Gottlieb M, Long B, Koyfman A. The evaluation and management of toxic shock syndrome in the emergency department: a review of the literature. J Emerg Med 2018;54:807–14.
5. Centers for Disease Control and Prevention. Toxic shock syndrome (other than streptococcal) (TTS) 2011 case definition. URL: https://ndc.services.cdc.gov/case-definitions/toxic-shock-syndrome-2011/.
6. Parola P, Paddock CD, Socolovschi C, Labruna MB, Mediannikov O, Kernif T, et al. Update on tick-borne rickettsioses around the world: a geographic approach. Clin Microb Rev 2013;26:657–702.
7. Tunnicliffe DJ, Singh-Grewal D, Kim S, Craig JC, Tong A. Diagnosis, monitoring, and treatment of systemic lupus erythematosus: a systematic review of clinical practice guidelines. Arthritis Care Res (Hoboken) 2015;67:1440–52.
8. Grunewald J, Grutters JC, Arkema EV, Saketkoo LA, Moller DR, Müller-Quernheim J. Sarcoidosis. Nat Rev Dis Primers 2019;5:45.
9. Desbois AC, Cacoub P, Saadoun D. Cryoglobulinemia: an update in 2019. Joint Bone Spine 2019;86:707–13.
10. Yang X, Perez OA, English JC III. Adult perniosis and cryoglobulinemia: a retrospective study and review of the literature. J Am Acad Dermatol 2010;62:e21–22.
11. Dietz SM, van Stijn D, Burgner D, Levin M, Kuipers IM, Hutten BA, et al. Dissecting Kawasaki disease: a state-of-the-art review. Eur J Pediatr 2017;176:995–1009.
12. Germain BF, Moroney JD, Guggino GS, Cimino L, Rodriguez C, Bocanegra TS. Anterior uveitis in Kawasaki disease. J Pediatr 1980; 97:780–1.
13. Verdoni L, Mazza A, Gervasoni A, Martelli L, Ruggeri M, Ciuffreda M, et al. An outbreak of severe Kawasaki-like disease at the Italian epicentre of the SARS-CoV-2 epidemic: an observational cohort study. Lancet 2020;395:1771–8.
14. Whittaker E, Bamford A, Kenny J, Kaforou M, Jones CE, Shah P, et al. Clinical characteristics of 58 children with a pediatric inflammatory multisystem syndrome temporally associated with SARS-CoV-2. JAMA 2020;324:259–69.
15. Centers for Disease Control and Prevention. Multisystem inflammatory syndrome in children (MIS-C) associated with coronavirus disease 2019 (COVID-19). 2020. URL: https://emergency.cdc.gov/han/2020/han00432.asp.
16. Feldstein LR, Rose EB, Horwitz SM, Collins JP, Newhams MM, Son MB, et al. Multisystem inflammatory syndrome in U.S. children and adolescents. N Engl J Med 2020;383:334–46.
17. Alhazzani W, Møller MH, Arabi YM, Loeb M, Gong MN, Fan E, et al. Surviving sepsis campaign: guidelines on the management of critically ill adults with Coronavirus Disease 2019 (COVID-19). Intensive Care Med 2020;46:854–87.
18. Fadel R, Morrison AR, Vahia A, Smith ZR, Chaudhry Z, Bhargava P, et al. Early short-course corticosteroids in hospitalized patients with COVID-19. Clin Infect Dis 2020;71:2114–20.
19. Viner RM, Whittaker E. Kawasaki-like disease: emerging complication during the COVID-19 pandemic. Lancet 2020;395:1741–3.
20. Cheung EW, Zachariah P, Gorelik M, Boneparth A, Kernie SG, Orange JS, et al. Multisystem inflammatory syndrome related to COVID-19 in previously healthy children and adolescents in New York City. JAMA 2020;324:294–6.
21. Morris SB, Schwartz NG, Patel P, Abbo L, Beauchamps L, Balan S, et al. Case series of multisystem inflammatory syndrome in adults associated with SARS-CoV-2 infection—United Kingdom and United States, March–August 2020. MMWR Morb Mortal Wkly Rep 2020;69:1450–6.
22. Shaigany S, Gnirke M, Guttmann A, Chong H, Meehan S, Raabe V, et al. An adult with Kawasaki-like multisystem inflammatory syndrome associated with COVID-19. Lancet 2020;396: e8–10.
23. Sokolovsky S, Soni P, Hoffman T, Kahn P, Scheers-Masters J. COVID-19 associated Kawasaki-like multisystem inflammatory disease in an adult. Am J Emerg Med 2021;39:e1–253.
24. Cogan E, Foulon P, Cappeliez O, Dolle N, Vanfraechem G, De Backer D. Multisystem inflammatory syndrome with complete Kawasaki disease features associated with SARS-CoV-2 infection in a young adult: a case report. Front Med (Lausanne) 2020;7:428.
25. Jones I, Bell LC, Manson JJ, Last A. UCLH COVID response team. An adult presentation consistent with PIMS-TS. Lancet Rheumatol 2020; 2:e520–1.
26. Kofman AD, Sizemore EK, Detelich JF, Albrecht B, Piantadosi AL. A young adult with COVID-19 and multisystem inflammatory syndrome in children (MIS-C)-like illness: a case report. BMC Infect Dis 2020 https://doi.org/10.1186/s12879-020-05439-z.
27. Magro C, Mulvey JJ, Berlin D, Nuovo G, Salvatore S, Harp J, et al. Complement associated microvascular injury and thrombosis in the pathogenesis of severe COVID-19 infection: a report of five cases. Transl Res 2020;220:1–13.
28. Chau VQ, Giustino G, Mahmood K, Oliveros E, Neibart E, Oloomi M, et al. Cardiogenic shock and hyperinflammatory syndrome in young males with COVID-19. Circ Heart Fail 2020;13:e007485.
29. Oxley TJ, Mocco J, Majidi S, Kellner CP, Shoirah H, Singh P, et al. Large-vessel stroke as a presenting feature of Covid-19 in the young. N Engl J Med 2020;382:e60.
30. Fox SE, Lameira FS, Rinker EB, Vander Heide RS. Cardiac endotheliitis and multisystem inflammatory syndrome after COVID-19. Ann Intern Med 2020;173:1025–7.
31. Ventura MJ, Guajardo E, Clark EH, Bhairavarasu K, Kherallah RY, DiNardo AR, et al. Correspondence on 'Paediatric multisystem inflammatory syndrome temporally associated with SARS-CoV-2 mimicking Kawasaki disease (Kawa-COVID-19): a multicentre cohort' by Pouletty et al. Ann Rheum Dis 2020 doi: 10.1136/annrheumdis-2020-218959. E-pub ahead of print.
32. Henderson LA, Canna SW, Friedman KG, Gorelik M, Lapidus SK, Bassiri H, et al. American College of Rheumatology clinical guidance for multisystem inflammatory syndrome in children associated with SARS-CoV-2 and hyperinflammation in pediatric COVID-19: version 1. Arthritis Rheumatol 2020;72:1791–805.
33. Godfred-Cato S, Bryant B, Leung J, Oster ME, Conklin L, Abrams J, et al. COVID-19–associated multisystem inflammatory syndrome in children—United States, March–July 2020. MMWR Morb Mortal Wkly Rep 2020;69:1074–80.
34. Ryan A, Pepper G, Wener MH, Fink SL, Morishima C, Chaudhary A, et al. Performance characteristics of the Abbott Architect SARS-CoV-2 IgG assay and seroprevalence in Boise, Idaho. J Clin Microbiol 2020;58:e00941–20.
35. Panupattanapong S, Brooks EB. New spectrum of COVID-19 manifestations in children: Kawasaki-like syndrome and hyperinflammatory response. Cleve Clin J Med 2020 doi: https://doi.org/10.3949/ccjm.87a.ccc039. E-pub ahead of print.

Arthritis Care & Research
Vol. 74, No. 2, February 2022, pp 179–186
DOI 10.1002/acr.24446

American College of Rheumatology
Empowering Rheumatology Professionals

Quality of Care for Patients With Systemic Lupus Erythematosus: Data From the American College of Rheumatology RISE Registry

Gabriela Schmajuk,[1] Jing Li,[2] Michael Evans,[2] Christine Anastasiou,[2] Julia L. Kay,[2] and Jinoos Yazdany[2]

Objective. Although multiple national quality measures focus on the management and safety of rheumatoid arthritis, few measures address the care of patients with systemic lupus erythematosus (SLE). Our objective was to apply a group of quality measures relevant to the care of patients with SLE, and we used the American College of Rheumatology's Rheumatology Informatics System for Effectiveness (RISE) registry to assess nationwide variations in care.

Methods. The data derived from RISE and included patients who had ≥2 visits with SLE codes ≥30 days apart in 2017–2018. We calculated performance on 5 quality measures: renal disease screening, blood pressure assessment and management, hydroxychloroquine (HCQ) prescribing, safe dosing for HCQ, and prolonged glucocorticoid use at doses of >7.5 mg/day. We reported performance on these measures at the practice level. We used logistic regression to assess independent predictors of performance after adjusting for sociodemographic and utilization factors.

Results. We included 27,567 unique patients from 186 practices; 91.7% were female and 48% White, with a mean age of 53.5 ± 15.2 years. Few patients had adequate screening for the development of renal manifestations (39.5%). Although blood pressure assessment was common (94.4%), a meaningful fraction of patients had untreated hypertension (17.7%). Many received HCQ (71.5%), but only 62% at doses of ≤5.0 mg/kg/day. Some received at least moderate-dose steroids for ≥90 days (18.5%). We observed significant practice variation on every measure.

Conclusion. We found potential gaps in care for patients with SLE across the US. Although some performance variation may be explained by differences in disease severity, dramatic differences suggest that developing quality measures to address important health care processes in SLE may improve care.

INTRODUCTION

The movement to measure quality of care in rheumatology has accelerated in the past decade, with new quality measures being developed, especially for patients with conditions such as rheumatoid arthritis and gout (1). The primary purpose of measuring and reporting quality in these conditions is to facilitate evidence-based practice that can improve patient outcomes, and to encourage the accountability of providers, health systems, and health plans. Development of quality measures for systemic lupus erythematosus (SLE) has lagged, in large part because it is a heterogenous, multiorgan-system disease with few evidence-based guidelines.

In 2009, Yazdany et al published the first set of quality indicators for patients with SLE, which addressed lupus-specific processes of care, including timely diagnosis and treatment of lupus nephritis, appropriate serologic monitoring, teratogenic drug counseling, drug toxicity monitoring, glucocorticoid management, and sun avoidance counseling (2). As evidence has grown around the comorbid conditions associated with SLE, quality measures

The views expressed represent those of the authors and not necessarily those of the American College of Rheumatology.

Data collection was supported by the American College of Rheumatology's Rheumatology Informatics System for Effectiveness Registry. This study protocol was prepared by and the data analysis was executed by the University of California, San Francisco. The work of Drs. Schmajuk and Yazdany was supported by the Agency for Healthcare Research and Quality (1R01-HS-024412) and the Russell/Engleman Medical Research Center for Arthritis. Dr. Anastasiou's work was supported by the National Institute on Aging (NIH grant 5T32-AG-049663-04). The work of Ms. Kay and Dr. Yazdany was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases NIH grant (P30-AR-070155).

[1]Gabriela Schmajuk, MD, MSc: University of California, San Francisco, and San Francisco Veterans Affairs Medical Center, San Francisco, California; [2]Jing Li, MPH, Michael Evans, MA, Christine Anastasiou, MD, MAS, Julia L. Kay, BA, Jinoos Yazdany, MD, MPH: University of California, San Francisco.

Dr. Yazdany has received consulting fees from Eli Lilly and Company and AstraZeneca (less than $10,000 each). No other disclosures relevant to this article were reported.

Address correspondence to Gabriela Schmajuk, MD, MSc, Associate Professor In Residence, Department of Medicine (Rheumatology), University of California, San Francisco, San Francisco VA Medical Center, 4150 Clement Street 111R, San Francisco, CA 94121. Email: gabriela.schmajuk@ucsf.edu.

Submitted for publication June 25, 2020; accepted in revised form September 8, 2020.

SIGNIFICANCE & INNOVATIONS

- We calculated performance on 5 quality measures relevant to the outpatient care of patients with systemic lupus erythematosus (SLE): renal disease screening, blood pressure assessment and management, hydroxychloroquine (HCQ) prescribing, safe dosing for HCQ, and prolonged glucocorticoid use at doses of >7.5 mg/day.
- We found potential gaps in care for patients with SLE across the US. Although some performance variation may be explained by differences in disease severity, dramatic differences across practices suggest that developing quality measures to address important health care processes in SLE may improve quality of care.

that address cardiovascular disease, osteoporosis, and infectious risk (vaccinations) have also been considered applicable to this patient population (3). However, only 2 performance measures that address outcomes germane to patients with SLE have been tested using administrative data: in-hospital mortality and 30-day hospital readmission rate. Unfortunately, these measures are not relevant to the ambulatory setting, where most patients with SLE receive their care.

In this study, our objectives were to specify a series of quality measures for outpatients with SLE and to assess performance on these measures nationally using data from a large electronic health record (EHR)–based registry in the US.

PATIENTS AND METHODS

Quality measure specification. We defined a series of quality measures relevant to the outpatient care of patients with SLE based on existing evidence-based recommendations and taking into account the feasibility of assessing measures using structured data from the EHR (Table 1). The first 4 were process measures addressing important features of the care of patients with SLE, including screening for renal disease and hypertension, and the universal and safe use of hydroxychloroquine (HCQ) (4–6).

We defined a single intermediate outcome measure to address blood pressure control based on an existing National Quality Forum–endorsed measure (7): for patients with at least 2 blood pressure measurements recorded, we assessed whether systolic blood pressure was >140 mm Hg or diastolic blood pressure was >90 mm Hg on at least 2 occasions and ≥30 days apart (8).

We also defined an exploratory measure around glucocorticoid use that assessed whether patients were receiving moderate- or high-dose glucocorticoids at a dose of >7.5 mg prednisone (or equivalent) daily for at least 90 days during the calendar year. The rationale for this exploratory measure was to provide clinicians with a measure that could provide insight into the proportion of patients who could meet glucocorticoid criteria for the lupus low disease activity state (LLDAS) (9).

Data source. The data derived from the American College of Rheumatology's Rheumatology Informatics System for Effectiveness (RISE) registry. RISE is a national, EHR-enabled registry that passively collects data on all patients seen by participating practices, reducing the selection bias present in single-insurer claims databases (10). As of December 2018, RISE held validated data from 1,113 providers in 226 practices, representing approximately 32% of the US clinical rheumatology workforce.

Study population. Patients included in this study were age >18 years and had at least 2 SLE diagnoses ≥30 days apart, during calendar year 2017 (January 1 to December 31) or calendar year 2018 (January 1 to December 31). Patients with visits in both years were only included in the analysis of 2018 data (n = 12,292). We excluded patients from practices in which laboratory data were not available (patients [n = 189]; practices [n = 28]).

Quality measure assessment in the RISE registry. Each of the measures in Table 1 was assessed across all patients in the RISE registry who entered the study population, according to the denominator, numerator, and exclusion definitions above. In the primary analysis, renal disease screening could occur via urinalysis alone or a quantitative assessment of urine protein. In a sensitivity analysis, we required a quantitative assessment of urine protein (i.e., the numerator definition included quantitative assessment such as urine protein or a urine protein:creatinine ratio, but a patient with a urinalysis result alone would not enter the numerator). Safe HCQ dosing was defined as a dose of ≤5.0 mg/kg/day. We also examined HCQ doses of ≤6.5 mg/kg/day (see Supplementary Tables 1 and 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24446/abstract). Patients prescribed HCQ who were missing or had an invalid weight (i.e., weight below the first percentile or higher than the 99th percentile weight of the general US population) were counted as a "No Pass" (n = 431) (11).

For the prolonged glucocorticoid use measure, prednisone equivalents included oral cortisone, hydrocortisone, prednisolone, triamcinolone, methylprednisolone, dexamethasone, and betamethasone. Pill sizes (in milligrams) were calculated based on National Drug Code codes, where available, or drug name and route, and an equivalence dose table using prednisone as the reference. Due to the complexity of prednisone dosing, we used a commercially available tool (12) in combination with manual review to determine the total prednisone dose based on the medication instruction ("sig") fields. If a patient was given 2 prescriptions of different amounts during the same 90-day period, the total daily dose reflected the sum of the 2 amounts. If a patient

Table 1. Specification of proposed SLE quality measures*

Quality measure description	Denominator†	Numerator	Exclusions, exceptions	Measurement period
Renal disease screening: proportion of patients with SLE who had urinary screening for lupus nephritis at least once per year	Patients with ≥2 face-to-face encounters with ICD codes for SLE, ≥30 days apart	Patients with ≥1 documented urine study (urinalysis, urine protein, or urine protein:creatinine ratio)	ESRD (585.6, N18.6, Z99.2, CPT 90951- 90970)	1 calendar year (e.g., 1/1/2018–12/31/2018)
Blood pressure assessment: proportion of patients with SLE who had at least 2 blood pressure readings per year	Patients with ≥2 face-to-face encounters with ICD codes for SLE, ≥30 days apart	Patients with ≥2 blood pressure readings recorded ≥30 days apart	None	1 calendar year (e.g., 1/1/2018–12/31/2018)
Blood pressure uncontrolled: proportion of patients with SLE without adequate blood pressure control	Patients with ≥2 face-to-face encounters with ICD codes for SLE, ≥30 days apart, AND ≥2 blood pressure readings, ≥30 days apart	Patients with systolic blood pressure of >140 mm Hg or diastolic blood pressure of >90 mm Hg on ≥2 occasions, ≥30 days apart	None	1 calendar year (e.g., 1/1/2018–12/31/2018)
HCQ prescription: proportion of patients with SLE who were prescribed HCQ	Patients with ≥2 face-to-face encounters with ICD codes for SLE, ≥30 days apart	Patients with at least 1 prescription for HCQ	Toxic maculopathy of retina (H35.0, 381–383, 362.55) or poisoning, adverse effect of other specified systemic antiinfectives and antiparasitics (T37.8x, T37.9x, E931.4)	1 calendar year (e.g., 1/1/2018–12/31/2018)
Safe HCQ dosing: proportion of patients with SLE receiving HCQ prescribed doses associated with less retinal toxicity	Patients with ≥2 face-to-face encounters with ICD codes for SLE, ≥30 days apart, AND at least 1 prescription for HCQ	Patients prescribed ≤5.0 mg/kg/day of HCQ on their most recent prescription	None	1 calendar year (e.g., 1/1/2018–12/31/2018)
Glucocorticoid use of >7.5 mg/day for ≥90 days: proportion of patients with SLE who do not meet the Lupus Low Disease Activity Index glucocorticoid criteria (≤7.5 prednisone mg/day).	Patients with ≥2 face-to-face encounters with ICD codes for SLE, ≥30 days apart	Patients prescribed >7.5 mg of prednisone (or equivalent) for ≥90 days (not required to be continuous days)	None	1 calendar year (e.g., 1/1/2018–12/31/2018)

* CPT = Current Procedural Terminology; ESRD = end-stage renal disease; HCQ = hydroxychloroquine; ICD = International Classification of Diseases; SLE = systemic lupus erythematosus.
† SLE was defined using ICD codes 710.0, 710.00, or M32x (except M32.0).

was given 2 prescriptions of the same amount during the same 90-day period, this amount was considered an extension of the same prescription, so amounts were not summed. Patients with a "sig" field that only said "as directed" were assumed to be taking 1 pill once per day (n = 562), given that this dosage would likely be the most conservative (lowest dose) interpretation of the order. Patients without any glucocorticoids prescribed were considered to have a dose of "0" and counted as "Pass" for this measure. The total number of days with a dose of >7.5 mg was calculated during the calendar year; patients with ≥90 days were counted as a "No Pass" for the measure. The 90 days were not required to be continuous.

We defined a composite measure to assess performance on the combination of process measures listed in Table 1 (renal disease screening, blood pressure assessment, HCQ prescription, and safe HCQ dosing). Performance was calculated as the number of measures fulfilled divided by the number of measures for which each patient was eligible. All patients were eligible for the first 3 measures, and patients with at least 1 prescription for HCQ were assessed for all 4. Performance was aggregated by practice.

Covariates and clinical manifestations. We extracted information on patient and practice characteristics from RISE. Patient characteristics included age, sex, race/ethnicity, insurance, number of visits during the study period, Area Deprivation Index (an area-level measure of socioeconomic status [range 1–100], with lower scores meaning higher socioeconomic status) (13), Charlson comorbidity index (based on the Deyo protocol as a measure of comorbidity [14]), and functional status measure scores (including the Multidimensional Health Assessment

Table 2. Patient characteristics (n = 27,567)*

Characteristic	Value
Female	25,284 (91.7)
Age, mean ± SD years	53.5 ± 15.2
Race/ethnicity	
White	13,235 (48.0)
African American	5,168 (18.8)
Hispanic	2,633 (9.6)
Asian	609 (2.2)
Other/mixed	1,758 (6.4)
Unknown/declined	4,164 (15.1)
Insurance	
Private	9,783 (35.5)
Medicare	5,719 (20.8)
Any Medicaid	1,082 (3.9)
Other	1,506 (5.5)
Unknown	9,477 (34.4)
Area Deprivation Index, median (IQR)	46 (25–69)
Geographic division	
New England	438 (1.6)
Mid-Atlantic	2,601 (9.4)
East North Central	2,908 (10.6)
West North Central	1,867 (6.8)
South Atlantic	10,172 (36.9)
East South Central	3,337 (12.1)
West South Central	2,575 (9.3)
Mountain	1,191 (4.3)
Pacific	2,478 (9.0)
Clinical characteristics	
Number of visits, mean ± SD	3.9 ± 2.7
Charlson comorbidity index, mean ± SD	1.4 ± 1.1
ANA positive (n = 11,994)	8,414 (70.2)
Anti–double-stranded DNA positive (n = 17,908)	8,229 (46.0)
Lupus nephritis†	1,585 (5.7)
End-stage renal disease	151 (0.5)
Functional status assessment scores, mean ± SD	
MDHAQ (n = 5,324; range 0–10)	1.98 ± 2.4
HAQ (n = 2,597; range 0–3)	0.78 ± 0.7
HAQ-II (n = 739; range 0–3)	0.78 ± 0.7
Medications	
Hydroxychloroquine	19,647 (71.5)
Glucocorticoids	12,299 (44.6)
Biologics or JAK inhibitors‡	4,660 (16.9)
Methotrexate	2,713 (9.8)
Azathioprine	2,044 (7.4)
Mycophenolate or mycophenolic acid	2,029 (7.4)
Tacrolimus	23 (0.1)

* Values are the number (%) unless indicated otherwise. ANA = antinuclear antibody; HAQ = Health Assessment Questionnaire; IQR = interquartile range; MDHAQ = Multidimensional Health Assessment Questionnaire.
† Lupus nephritis was defined by International Classification of Diseases, Ninth Revision codes 580.0–586.0 and 791.0.
‡ Biologics included abatacept, belimumab, denosumab, rituximab, and other; JAK inhibitors included tofacitinib.

Questionnaire [MDHAQ], the Health Assessment Questionnaire [HAQ], and HAQ-II). Additional medication data were also extracted, including use of biologics (abatacept, belimumab, denosumab, rituximab, and other), JAK inhibitors (tofacitinib), mycophenolate or mycophenolic acid, azathioprine, methotrexate, and tacrolimus. Diagnoses were defined using International Classification of Diseases (ICD) codes for each of the following during the study period: SLE (710.0, 710.00, or M32x [except M32.0]); lupus nephritis (ICD codes 580.0–586.0 and 791.0); and end-stage renal disease (N18.6, 585.6, Z99.2, or Current Procedural Terminology code for dialysis 90951-90970) (15). We extracted information on antinuclear antibody (ANA) and anti–double-stranded DNA (anti-dsDNA) antibodies at any time prior to the measurement year; ANA and dsDNA were classified as positive if the results included "positive," "detected," or "reactive," or if titers were >1:40 for ANA or ≥1:40 for dsDNA antibodies.

Practice characteristics included practice type (single-specialty, solo practitioner, multispecialty, health system, and other), practice size (number of providers, number of eligible patients in each practice), EHR vendor, geographic division, and the number of years contributing data to RISE. The latter variable was used to account for the possibility that data completeness may improve the longer a practice participated in the registry.

Statistical analysis. Descriptive statistics were used to examine patient and practice characteristics. Patient-level quality measures were reported as the proportion of eligible individuals meeting criteria for the measures according to Table 1. Practice-level performance aggregated information from all patients seen within a given practice, examining the proportion of patients fulfilling each quality measure among all those eligible; interquartile ranges (IQRs) were reported. Practices reporting on <20 patients were excluded from the practice-level analyses. We used multilevel logistic regression models that included age, sex, race, insurance, Area Deprivation Index, number of visits, and geographic region to assess independent predictors of performance on each measure, accounting for clustering by practice. Analyses were performed using SAS software, version 9.4. The Western

Table 3. Practice characteristics (n = 186)*

Characteristic	Value
Practice type	
Single-specialty group practice	110 (59.1)
Solo practitioner	49 (26.3)
Multispecialty group practice	23 (12.4)
Health system	4 (2.2)
Number of providers per practice	
Median (IQR)	2 (1–5)
Range	1–35
Number of eligible patients in each practice	
Median (IQR)	104 (42–205)
Range	1–1,125
EHR vendor	
NextGen	75 (40.3)
eClinicalWorks	32 (17.2)
Amazing Charts	16 (8.6)
eMDs	10 (5.4)
Aprima	8 (4.3)
Other	45 (24.2)
Years contributing data to RISE	
Median (IQR)	2.68 (1.73–3.58)
Range	0.32–5.37

* Values are the number (%) unless indicated otherwise. EHR = electronic health record; IQR = interquartile range; RISE = Rheumatology Informatics System for Effectiveness.

Table 4. Quality measures, number of eligible patients, and overall performance

Quality measure	Eligible patients, no.	Overall performance, no. (%)	Practices included in practice-level analysis, no.*	Practice-level performance, 25th–75th percentile
Renal disease screening	27,369	10,823 (39.5)	164	4.1–60.9
Blood pressure assessment	27,567	26,037 (94.4)	165	96.7–100
Blood pressure uncontrolled	26,037	4,612 (17.7)	152	7.9–26.0
Hydroxychloroquine prescription	27,486	19,647 (71.5)	165	64.9–80.0
Safe hydroxychloroquine dosing	19,647	12,172 (62.0)	163	77.3–95.5
Prolonged glucocorticoid use of >7.5 mg	27,567	5,085 (18.5)	165	10.7–22.1
Composite process measure†	27,567	7,626 (27.7)	165	1.2–42.8

* Practice-level analysis included only practices reporting on ≥20 patients.
† Composite process measure includes renal disease screening, blood pressure assessment, hydroxychloroquine prescription, and safe hydroxychloroquine dosing.

Institutional Review Board and University of California, San Francisco Committee on Human Research approved this study.

RESULTS

There were 27,567 patients with SLE included in this study. The majority (91.7%) were female, with a mean ± SD age of 53.5 ± 15.2 years (Table 2). Almost half (48%) of the patients were White, 18.8% were African American, and 9.6% were Hispanic. Most patients had private or Medicare insurance (35% and 21%, respectively), with a small number of patients on Medicaid (3.9%); a large proportion of patients had unknown insurance coverage (34%). The mean ± SD number of visits was 3.9 ± 2.7 during the study period. The median for Area Deprivation Index was 46 (IQR 25–69). Patients had a mean ± SD Charlson comorbidity index score of 1.4 ± 1.1. Overall, mean ± SD scores of MDHAQ, HAQ, and HAQ-II were 2.0 ± 2.4, 0.8 ± 0.7, and 0.8 ± 0.7, respectively. A total of 71.5% of patients were receiving HCQ, 45% receiving glucocorticoids, and 17% receiving biologics or JAK inhibitors. Other medications used are listed in Table 2. In all, 1,585 patients (5.7%) had a diagnosis of lupus nephritis and 151 (0.5%) were diagnosed with end-stage renal disease.

Among the 186 practices represented, 59.1% were single-specialty groups, followed by 26.3% solo practitioners, and 12.4% multispecialty groups (Table 3). The median number of providers per practice was 2 (range 1–35; IQR 1–5) and the median of eligible patients per practice was 104. NextGen and eClinicalWorks made up almost 60% of the EHRs used by these practices (40.3% and 17.2%, respectively).

Performance on the proposed quality measures is shown in Table 4: fewer than 40% of patients with SLE had adequate

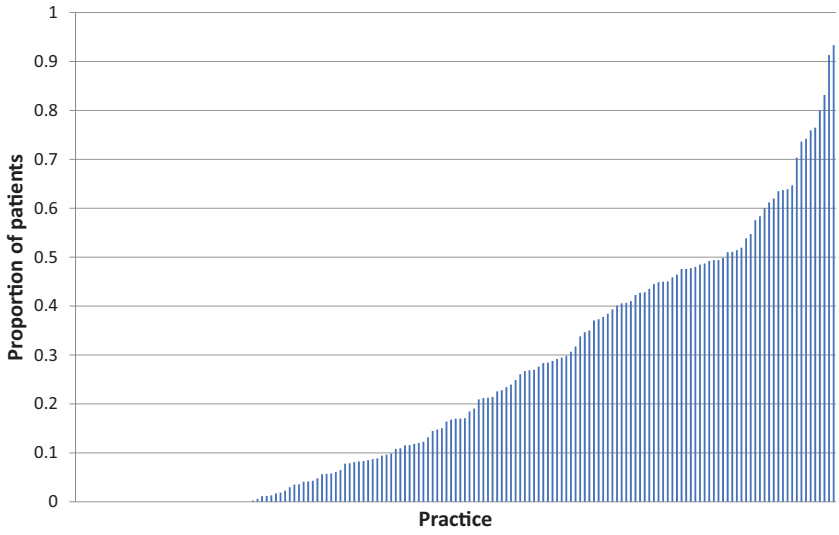

Figure 1. Proportion of patients with systemic lupus erythematosus in practices in the Rheumatology Informatics System for Effectiveness registry who passed the composite process measure, by practice (n = 165). Composite measures included renal disease screening, blood pressure assessment, hydroxychloroquine prescription, and safe hydroxychloroquine dosing (<5.0 mg/kg/day). Practices reporting on <20 patients were not included.

Table 5. Composite measure, patient level analysis clustering by practice (n = 27,251)*

	Unadjusted OR (95% CI)	Adjusted OR (95% CI)
Age, per 10 years	0.94 (0.93–0.96)†	0.95 (0.93–0.97)†
Male	1.31 (1.20–1.42)†	1.34 (1.23–1.47)†
Race/ethnicity		
White	Ref.	Ref.
Hispanic	1.13 (1.03–1.25)†	1.09 (0.98–1.22)
African American	1.37 (1.28–1.47)†	1.34 (1.24–1.45)†
Asian	1.11 (0.94–1.32)	1.06 (0.89–1.28)
Other/mixed	1.11 (0.99–1.24)	1.11 (0.99–1.25)
Unknown	1.23 (1.13–1.35)†	1.20 (1.10–1.32)†
Insurance		
Private	Ref.	Ref.
Medicare	1.08 (0.94–1.23)	0.92 (0.80–1.07)
Any Medicaid	1.04 (0.97–1.11)	0.95 (0.88–1.03)
Other	1.22 (1.08–1.38)†	1.12 (0.97–1.30)
Unknown	0.90 (0.82–0.99)†	0.81 (0.73–0.89)†
Geographic division		
New England	Ref.	Ref.
Mid-Atlantic	0.59 (0.17–2.02)	0.78 (0.25–2.44)
East North Central	2.59 (0.87–7.76)	2.47 (0.84–7.28)†
West North Central	4.60 (1.40–15.11)†	4.63 (1.43–14.97)
South Atlantic	1.74 (0.61–4.92)	1.66 (0.60–4.59)
East South Central	2.06 (0.68–6.22)	2.09 (0.71–6.16)
West South Central	2.44 (0.81–7.38)	2.17 (0.74–6.42)
Mountain	1.27 (0.34–4.84)	1.47 (0.42–5.20)
Pacific	0.97 (0.32–2.97)	0.93 (0.31–2.76)
Visits, no.	1.04 (1.03–1.05)†	1.04 (1.03–1.06)†
ADI	1.00 (1.00–1.00)†	1.00 (1.00–1.00)

* 95% CI = 95% confidence interval; ADI = Area Deprivation Index; OR = odds ratio; Ref. = reference.
† Patients missing ADI and from practices with <20 patients were not included in this analysis. Variables included in the multivariate models: age, sex, race/ethnicity, insurance, number of visits, geographic division, and ADI.

screening for renal disease. Although blood pressure screening was common (94.4%), a meaningful fraction of patients (17.7%) had undertreated hypertension. A total of 71.5% of patients had received at least 1 prescription for HCQ, and 38% were prescribed doses of >5.0 mg/kg/day. Nearly 20% of patients were receiving at least moderate-dose glucocorticoids for at least 90 days during the calendar year, signaling that they had not achieved LLDAS.

Analysis of the composite of the 4 process measures revealed that 27.7% of patients received all services for which they were eligible. Among patients with any kind of renal disease (n = 1,662), performance on the composite measure was 42.5%. As with the individual measures, we observed wide practice variation on the composite measure, ranging from 1% to 93.3% among practices reporting on at least 20 patients (Figure 1).

In a sensitivity analysis where we required a quantitative assessment of renal protein for the renal disease screening measure, overall performance was only 24.3% (6,645 of 27,369) with a practice performance median of 9.5% (IQR 0–33.9). Using this version of the renal disease screening measure resulted in a composite measure performance of 17.3%, with a practice performance median of 6.1% (IQR 0–23.5). In multilevel logistic regression models, we found that patients who were older, female, and White were less likely to receive all process measures for which they were eligible (Table 5). As expected, patients with fewer visits were less likely to receive all services.

DISCUSSION

This is the first nationwide examination of a series of electronically specified quality measures applicable to patients with SLE using a large EHR-based registry in the US. While some aspects of care were standardized across rheumatology practices, such as blood pressure monitoring, others demonstrated significant gaps in care, including moderate use of HCQ, low rates of screening for renal disease, and a significant portion of patients with uncontrolled hypertension. We also found that approximately one-fifth of patients received >7.5 mg of prednisone for >90 days, suggesting that they would not have achieved LLDAS.

The purpose of developing and assessing the measures defined here was 3-fold. First, some measures could be used for quality reporting. Existing rheumatology-specific measures address the care of rheumatoid arthritis and gout, but none specifically address SLE, a disease that disproportionately affects vulnerable populations, so including these measures is an important step in expanding quality programs. Second, there has been at least 1 study linking performance on process measures with reduced damage in SLE, so improving performance on these measures may reduce damage going forward (16). Third, some measures, especially the blood pressure control and prolonged glucocorticoid use measures, could be used for population health management across clinics or health systems and may facilitate the creation of tools that can be used directly to improve care. For example, implementing the prolonged use of the glucocorticoids measure in the RISE registry dashboard would facilitate the creation of reports showing lists of patients who may need closer follow-up or more aggressive glucocorticoid management plans.

We demonstrated the feasibility of assessing these measures by extracting information from structured fields in the EHR. Abstracting information about tests for urine protein, blood pressure and weight values, and medication doses was possible through structured EHR data fields. Calculations of prednisone dose presented a significant challenge, as this calculation required extraction of information from the medication instruction field (“sig”) where available, and many instructions read only “as directed.” To accomplish this calculation at scale and in real time, alternate methods that estimate dose based on the number of

pills dispensed might be easier, although such a method could sacrifice accuracy (17). Future work should test a variety of methods to accurately extract this information, including creating more standardized instruction options or having standardized fields where a clinician can designate whether a patient is receiving >7.5 mg prednisone/day at any given visit. We did not attempt to assess measures such as vaccination status, HCQ eye screening, or lipid monitoring. The feasibility of extracting this information, which may not be routinely documented in the rheumatology EHR at all, or captured only in the text of the clinical note, was substantially lower than those measures we did focus on. Future work should address these additional, important features of SLE care.

We observed significant variations in care across patients and practices. We found that patients who were older, female, and White were less likely to receive all services for which they were eligible, which likely reflects less intensive monitoring of patients with mild disease. Interestingly, practice variation in performance on the composite measure was not completely explained by these differences in patient case mix (unadjusted performance range 0–100%; adjusted performance range 3–63%) and may be due to differences in care provided, in documentation, or in workflows across practices. Although our data strongly suggest that there is a significant gap in the care of patients with SLE, the magnitude of the gap may be smaller than is reported here, reflecting inadequate EHR documentation. For example, some patients may have been screened or monitored for lupus nephritis or hypertension by clinicians outside the rheumatology practice, in which case these data would not have entered the participating rheumatologist's EHR. Work linking RISE data to administrative claims (e.g., Medicare claims) is ongoing and will improve our understanding of the magnitude of this underestimation. Nevertheless, most patients with SLE with access to rheumatology care (i.e., all patients included in this study), are likely to have HCQ prescribed by their rheumatologist.

Our finding that 70% of patients have at least 1 prescription for HCQ during the calendar year is similar to other recent reports of HCQ use, even among patients seeing a rheumatologist (18–20). Ultimately, inclusion of these quality measures in the RISE dashboard (or, potentially, in national pay-for-performance programs) will necessitate agreement from relevant stakeholders that these aspects of care are important to measure and improve. Moreover, improvement in these aspects of care will require accurate assessment of these measures, which may entail changes to documentation workflows at the practice level, and for RISE practices, more customized mapping of data elements by the registry clinical informatics team.

Most prior studies of quality of care in SLE have examined care for SLE outside of the specialty care setting. In these studies, racial/ethnic minorities were less likely to access subspecialty care for SLE, and those with low socioeconomic status were more likely to travel long distances to see a rheumatologist (21). Moreover, those with no health insurance were less likely to receive high-quality care (22). In the Medicaid population, those with low socioeconomic status were less likely to receive timely care for lupus nephritis and less likely to receive HCQ (17). We did not see previously observed differences in RISE data, suggesting that the largest sociodemographic disparities in health care may occur prior to patients accessing rheumatology care. Whether these observations remain consistent when more academic medical centers join the RISE registry, so that there will be greater diversity across socioeconomic status, will be interesting to see.

The main strength of this study is its description of the actual care received by patients; the data were derived from the RISE registry, were collected passively from the EHR, and reflect all patients seen in practices, thereby reducing selection bias. However, there are also several limitations: as mentioned above, the measures only capture care provided by the rheumatologist, so we may have underestimated the actual care received by patients across all of their providers. We were unable to capture reasons why care did not occur; for example, some patients may have declined HCQ or antihypertensives altogether. For the glucocorticoid measure, patients may have been prescribed prednisone for non-SLE conditions by nonrheumatology clinicians. Finally, RISE includes very few academic centers, so although it provides an important and unique picture of community-based rheumatology practice, data may not be generalizable to large health systems.

In summary, we evaluated a series of quality measures applicable to the care of patients with SLE. We found significant gaps in care among patients with SLE in a large US EHR-based registry. Implementing these measures to assess these gaps and feed information back to providers is likely to help improve the quality of care for patients with SLE.

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. Schmajuk 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. Schmajuk, Li, Yazdany.

Acquisition of data. Schmajuk, Li, Evans, Kay, Yazdany.

Analysis and interpretation of data. Schmajuk, Li, Anastasiou, Yazdany.

REFERENCES

1. National Quality Forum. NQF-endorsed measures for musculoskeletal conditions. 2015. http://www.qualityforum.org/Publications/2015/01/NQF-Endorsed_Measures_for_Musculoskeletal_Conditions.aspx.
2. Yazdany J, Panopalis P, Gillis JZ, Schmajuk G, MacLean CH, Wofsy D, et al. A quality indicator set for systemic lupus erythematosus. Arthritis Rheum 2009;61:370–7.

3. Feldman CH, Speyer C, Ashby R, Bermas BL, Bhattacharyya S, Chakravarty E, et al. Development of a set of lupus-specific ambulatory care–sensitive, potentially preventable adverse conditions: a Delphi consensus study. Arthritis Care Res (Hoboken) 2021; 73:146–57.
4. Fanouriakis A, Kostopoulou M, Alunno A, Aringer M, Bajema I, Boletis JN, et al. 2019 update of the EULAR recommendations for the management of systemic lupus erythematosus. Ann Rheum Dis 2019;78:736–45.
5. Ruiz-Irastorza G, Ramos-Casals M, Brito-Zeron P, Khamashta MA. Clinical efficacy and side effects of antimalarials in systemic lupus erythematosus: a systematic review. Ann Rheum Dis 2010;69:20–8.
6. Marmor MF, Kellner U, Lai TY, Melles RB, Mieler WF, American Academy of Ophthalmology. Recommendations on screening for chloroquine and hydroxychloroquine retinopathy (2016 revision). Ophthalmology 2016;123:1386–94.
7. Quality ID #236 (NQF): controlling high blood pressure. URL: https://qpp.cms.gov/docs/QPP_quality_measure_specifications/CQM-Measures/2019_Measure_236_MIPSCQM.pdf.
8. Tselios K, Sheane BJ, Gladman DD, Urowitz MB. Optimal monitoring for coronary heart disease risk in patients with systemic lupus erythematosus: a systematic review. J Rheumatol 2016;43:54–65.
9. Franklyn K, Lau CS, Navarra SV, Louthrenoo W, Lateef A, Hamijoyo L, et al. Definition and initial validation of a lupus low disease activity state (LLDAS). Ann Rheum Dis 2016;75:1615–21.
10. Yazdany J, Robbins M, Schmajuk G, Desai S, Lacaille D, Neogi T, et al. Development of the American College of Rheumatology's rheumatoid arthritis electronic clinical quality measures. Arthritis Care Res (Hoboken) 2016;68:1579–90.
11. Centers for Disease Control and Prevention. Adult BMI calculator. URL: https://www.cdc.gov/healthyweight/assessing/bmi/adult_bmi/english_bmi_calculator/bmi_calculator.html.
12. Elimu Informatics. Semantic normalization: beyond traditional terminology mapping. URL: https://www.elimu.io/semantic-normalization/.
13. University of Wisconsin Department of Medicine. Neighborhood atlas. URL: https://www.neighborhoodatlas.medicine.wisc.edu/.
14. Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. J Clin Epidemiol 1992;45:613–9.
15. Chibnik LB, Massarotti EM, Costenbader KH. Identification and validation of lupus nephritis cases using administrative data. Lupus 2010;19:741–3.
16. Yazdany J, Trupin L, Schmajuk G, Katz PP, Yelin EH. Quality of care in systemic lupus erythematosus: the association between process and outcome measures in the Lupus Outcomes Study. BMJ Qual Saf 2014;23:659–66.
17. Castillo F, Strait A, Evans M, Kay J, Gianfrancesco M, Izadi Z, et al. Deriving accurate prednisone dosing from electronic health records: analysis of a natural language processing tool for complex prescription instructions [abstract]. Arthritis Rheumatol 2019;71 Suppl 10.
18. Schmajuk G, Yazdany J, Trupin L, Yelin E. Hydroxychloroquine treatment in a community-based cohort of patients with systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2010;62:386–92.
19. Yazdany J, Feldman CH, Liu J, Ward MM, Fischer MA, Costenbader KH. Quality of care for incident lupus nephritis among Medicaid beneficiaries in the United States. Arthritis Care Res (Hoboken) 2014;66:617–24.
20. Xiong WW, Boone JB, Wheless L, Chung CP, Crofford LJ, Barnado A. Real-world electronic health record identifies antimalarial underprescribing in patients with lupus nephritis. Lupus 2019;28:977–85.
21. Yazdany J, Gillis JZ, Trupin L, Katz P, Panopalis P, Criswell LA, et al. Association of socioeconomic and demographic factors with utilization of rheumatology subspecialty care in systemic lupus erythematosus. Arthritis Rheum 2007;57:593–600.
22. Yelin E, Yazdany J, Tonner C, Trupin L, Criswell LA, Katz P, et al. Interactions between patients, providers, and health systems and technical quality of care. Arthritis Care Res (Hoboken) 2015;67:417–24.

Arthritis Care & Research
Vol. 74, No. 2, February 2022, pp 187–198
DOI 10.1002/acr.24431

American College of Rheumatology
Empowering Rheumatology Professionals

Systemic Lupus Erythematosus: Targeted Literature Review of the Epidemiology, Current Treatment, and Disease Burden in the Asia Pacific Region

Yoshiya Tanaka,[1] Sean O'Neill,[2] Mengtao Li,[3] I-Ching Tsai,[4] and Ya-Wen Yang[5]

Objective. To understand the epidemiology, current treatment, and disease burden of systemic lupus erythematosus (SLE) in the Asia Pacific (APAC) region.

Methods. A targeted literature review of published evidence on SLE in the APAC region was conducted, using the Medline database (2008–2018), conference proceedings, and other supplementary sources.

Results. The current review identified 70 studies conducted in China (n = 15), Japan (n = 13), Taiwan (n = 12), Korea (n = 9), Australia (n = 7), Hong Kong (n = 6), Singapore (n = 4), and multiple places within the APAC region (n = 4). Incidence rates (per 100,000 persons per year) ranged from 0.9–8.4, while prevalence rates ranged from 3.7–127 (per 100,000 persons); however, recent data were limited. Asian patients with SLE were reported to have higher disease severity, disease activity (higher SLE disease activity index scores), and organ damage accrual, along with increased morbidity, mortality, and susceptibility to renal involvement compared with other ethnicities in the APAC region. The risk of developing SLE is higher in the Asian population. Routinely used SLE therapies included belimumab, hydroxychloroquine, cyclophosphamide, tacrolimus, azathioprine, mycophenolate mofetil, and glucocorticoids; however, prescribing patterns varied across the region. Increased disease activity was associated with high economic burden and poor quality of life for SLE patients in the APAC region.

Conclusion. SLE remains a disease with a significant unmet medical need for an innovative therapy that is well-tolerated and effective for patients in the APAC region. Further evidence is required to better characterize the disease and fully capture the burden and impact of SLE in the APAC region. This review has highlighted where there is a paucity of data from patients across the APAC region.

INTRODUCTION

Systemic lupus erythematosus (SLE) is a chronic, multisystem autoimmune disease with an estimated 3.7 million prevalent cases worldwide (1,2). SLE is characterized by a wide spectrum of clinical manifestations, ranging from skin rashes to major organ damage (3). Other common symptoms of SLE include a characteristic red "butterfly" rash on the face, arthritis, and nephritis (3–5). Patients with SLE experience joint pain and fatigue, deficits in cognition and physical function, and have reduced health-related quality of life (HRQoL) compared with the general population (6,7). The mortality rate for patients with SLE is 67% higher than that of the general population (2).

Supported by Janssen: Pharmaceutical Companies of Johnson & Johnson.

[1]Yoshiya Tanaka, MD, PhD: School of Medicine, University of Occupational and Environmental Health, Japan, Kitakyushu, Japan; [2]Sean O'Neill, BMed, PhD, FRACP: Royal North Shore Hospital, The University of Sydney Northern Clinical School, St Leonards, New South Wales, Australia; [3]Mengtao Li, MD, PhD: Peking Union Medical College Hospital, Peking Union Medical College & Chinese Academy of Medical Sciences, National Clinical Research Center for Immunologic Diseases, Ministry of Science & Technology, and Key Laboratory of Rheumatology and Clinical Immunology, Ministry of Education, Beijing, China; [4]I-Ching Tsai, MHA: Janssen Pharmaceutical Companies of Johnson & Johnson, Taipei City, Taiwan; [5]Ya-Wen Yang, MD, MS, MBA: Global Medical Affairs, Janssen Pharmaceutical Companies of Johnson & Johnson.

Dr. Tanaka has received speaking fees and/or honoraria from Daiichi-Sankyo, Eli Lilly, Novartis, YL Biologics Ltd, Bristol Myers Squibb, Eisai, Chugai, AbbVie, Astellas, Pfizer, Sanofi, Asahi Kasei, GlaxoSmithKline, Mitsubishi Tanabe Pharma, Gilead Sciences, and Janssen Pharmaceuticals (more than $10,000 each) and research grants from AbbVie, Mitsubishi Tanabe Pharma, Chugai, Asahi Kasei, Eisai, Takeda and Daiichi-Sankyo. Dr. O'Neill has served as an investigator in clinical trials sponsored by Janssen, UCB, AstraZeneca and AbbVie and has received research grants from AbbVie and Janssen Pharmaceuticals. Dr. Li has received speaking fees and/or honoraria from Eli Lilly, Pfizer, AbbVie, Novartis, Janssen Pharmaceuticals and Sanofi (less than $10,000 each). Mrs. Tsai is an employee of Janssen Pharmaceuticals, and owns stock in Johnson & Johnson, the parent company of Janssen Pharmaceuticals. Dr. Ya-Wen Yang is an employee of Janssen Pharmaceutical companies of Johnson & Johnson and owns stock in Johnson & Johnson, the parent company of Janssen Pharmaceuticals.

Address correspondence to Ya-Wen Yang, MD, MS, MBA, Global Medical Affairs, Janssen Pharmaceutical Companies of Johnson & Johnson, 200 Tournament Drive, Horsham, PA 18902, Horsham, PA. Email: byang36@its.jnj.com.

Submitted for publication January 29, 2020; accepted in revised form August 18, 2020.

SIGNIFICANCE & INNOVATIONS

- The Asian population is at a higher risk of developing systemic lupus erythematosus (SLE) compared with other ethnicities and regions. SLE is associated with increased morbidity and mortality, as well as significant risk of early death compared with that in the general population.
- SLE carries a significant humanistic burden, with patients experiencing poor health-related quality of life and additional unmet needs that are not addressed by current physician-scored disease activity instruments.
- Further evidence is required to fully capture the impact of SLE in the Asia Pacific (APAC) region, e.g., real-world data of treatment-associated outcomes in SLE.
- SLE remains a disease with a significant unmet medical need for an innovative therapy that is well-tolerated and effective for patients in the APAC region.

The etiology of SLE is not fully understood; however, both genetic predisposition and environmental triggers are thought to be involved in activation of autoreactive T cells and overproduction of autoantibodies by B cells during the pathologic processes (5,8). The variable disease course and outcomes of SLE differ across racial/ethnic groups and sex, reflecting genetic, environmental, sociodemographic, and health care system differences in the Asia Pacific (APAC) region, as well as potential differences in treatment availability. The unpredictable disease course and organ involvement can make the diagnosis and treatment of SLE difficult (2,4).

Treatment decisions in SLE are driven by the organ systems involved, the severity of the disease, and the drug's safety profile. Although treatment recommendations differ globally, the goal of treatment is consistent across guidelines. Patients with more severe disease often require more intensive treatments such as medium to high dose glucocorticoids, immunosuppressants, and B cell modulators, where the goal is to induce and then sustain disease remission to avoid severe damage of vital organ systems (1,9,10). With the exception of hydroxychloroquine, glucocorticoids, some immunosuppressants (e.g., azathioprine), and belimumab, the majority of agents used to treat SLE are used off label (1), highlighting the need for more targeted therapies with proven efficacy for SLE.

Although there are a range of therapies available for SLE patients, symptom management remains unsatisfactory, and many treatments, particularly glucocorticoids, are associated with burdensome adverse effects (11–13). Patients receiving glucocorticoid treatment reported that the treatment constituted most of their organ damage at 15 years of follow-up (12). Glucocorticoid treatment is also an independent risk factor for cardiovascular events (13). There is, consequently, an unmet need for novel treatments that provide sustained disease control with minimal side effects, particularly for patients with severe SLE (4,14).

The Asian population in the APAC region is associated with a higher risk of developing SLE (10); however, data are limited on the epidemiology and burden of SLE across the region (2,3). The aim of this study was to review available published evidence associated with the epidemiology, current treatment, and disease burden of SLE in the APAC region.

MATERIALS AND METHODS

A targeted literature review (TLR) was conducted in accordance with the Cochrane Collaboration (15) and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Statement (16) guidelines. Ovid Medline was searched from January 2008 to September 2018, using a combination of free text keywords and Medical Subject Heading terms (see Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24431/abstract). Conference proceedings between 2015 and 2018, and inclusive and other supplementary sources (clinical guidelines, press releases, reference lists of included studies and identified systematic literature reviews [SLRs]) were searched. Randomized controlled trials (RCTs), observational, real-world studies, and SLRs that reported population-level data on the diagnosis, disease characteristics, treatment effect, mortality, comorbidities, epidemiology, direct and indirect costs, QoL, and unmet need associated with SLE in the APAC region were included. Studies that were cited in SLRs identified by this TLR were included as additional sources. An overview of the eligibility criteria and TLR methodology is shown (see Supplementary Appendix B, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24431/abstract).

RESULTS

Identification of included publications. A total of 70 unique publications (3,4,10,11,17–82) comprised of 27 clinical studies, 16 epidemiologic studies, and 27 studies reporting economic and humanistic burden and unmet need met the inclusion criteria (see Supplementary Appendix C, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24431/abstract). Most of the studies were conducted in China (n = 14) (11,19,21,23,28,29,38,43,47,54,59,78,79,82), Japan (n = 13) (31,33–37,40,58,65,67,74,75,77), and Taiwan (n = 12) (20,22,24,25,45,52,56,60–62,64,69), followed by Korea (n = 9) (17,26,32,42,44,53,63,72,76), Australia (n = 7) (4,18,39,49,50,57,66), Hong Kong (n = 6) (10,30,46,71,80,81), and Singapore (n = 4) (27,48,55,70); 4 studies were conducted in multiple places within the APAC region (3,41,51,73). Given the broad scope of the outcomes of interest, a high number of studies were identified; however, many reported laboratory findings, disease genetics, and pathology, and were

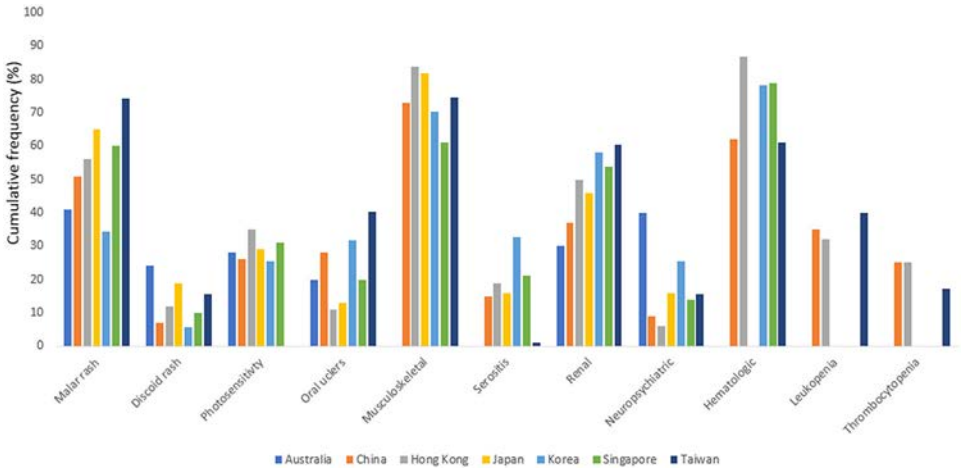

Figure 1. Frequency of key clinical manifestations in systemic lupus erythematosus patients across the Asia Pacific, by geographic region.

therefore excluded. Studies in pediatric populations were also excluded.

Disease characteristics of patients in included studies. Disease characteristics were reported for SLE patients in studies from Australia (n = 6) (4,18,39,50,57,66), China (n = 15) (11,19,21,23,28,29,38,43,47,54,59,78,79,82), Hong Kong (n = 6) (10,30,46,71,80,81), Japan (n = 13) (31,33–37,40,58,65,67,74, 75,77), Korea (n = 7) (17,26,41,42,44,63,72), Singapore (n = 4) (27,48,55,70), and Taiwan (n = 10) (20,22,24,25,45,52,61,62, 64,69) (see Supplementary Appendix C, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24431/abstract). The frequency of clinical manifestations by region is summarized in Figure 1 (20,22,24,25,45,52,61,62,64,69). Most patients were Chinese (35.0%) or Japanese (33.0%), followed by Taiwanese (17.0%), Korean (6.0%), and White (6.0%). In all of the studies that reported disease characteristics, prevalence of SLE was higher in female than in male patients (87.0% versus 13.0%). The mean age of included patients was 41.7 years (range 24–65 years). The mean age at SLE onset was 31.1 years (range 9.6–63.0 years), and was 32.6 years (range 11.3–75.0 years) at diagnosis. The mean disease duration was 5.8 years (range 0.5–30.6 years) across studies. SLE is characterized by episodes of disease flares, which, over time, can result in major organ damage, and fear of flares is a commonly reported unmet need of SLE patients in the APAC region (4).

Epidemiology of SLE. *Incidence and prevalence.* Incidence or prevalence of SLE was reported by 8 studies (see Supplementary Appendix C available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24431/abstract) (3,4,44,46,51–53,56). The incidence rates (per 100,000 persons per year) cited for the 7 APAC geographic regions ranged from 0.9 to 8.4, and the prevalence rates (per 100,000 persons per year) ranged from 3.7 to 127 (3,4,44,46,51–53,56). A relatively higher prevalence of SLE was seen in Aboriginal and Torres Strait Islander people (52.6–89.3) and pregnant women (127.0) in Australia (3,4) and in patients in Taiwan (21.1–97.5) (51,52,56) when compared with other populations in Australia (45.3) (3) and other APAC geographic regions, such as China (10.0–70.0) (51), Hong Kong (58.8) (3), Korea (18.8–21.7) (44,53), and Singapore (40.0) (51). The lowest prevalence and incidence of SLE was recorded in Japan (3.7–37.7 and 0.9–2.1, respectively) (3,51). A breakdown of the incidence and prevalence of SLE by ethnicity and age in each geographic region was not available from the literature, and the data do not allow a conclusion regarding whether SLE is more common in certain parts of the APAC region than in others.

Morbidity and mortality. In total, 11 studies were identified that reported mortality/morbidity (Table 1) and survival (Table 2) in the APAC region (3,24,29,42,45–47,54,55,83). The leading causes of death across the cited geographic regions were infections (27.6–69.0%), which were more common in Hong Kong (60%), Australia (67.0%), and Taiwan (69.0%) compared with other APAC geographic regions, followed by SLE activity (26.9–80.0%) and thromboembolism (8–30%) (3,24,29,42,45–47,54). Renal involvement was more common in the Asian population compared with White patients (83). Older age at disease onset, infection, autoimmune hemolytic anemia, thrombocytopenia, and pulmonary arterial hypertension were independent risk factors that were linked to SLE deaths (54). The risk of mortality is greater for SLE patients compared with the general population (see Supplementary Appendix C, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24431/abstract) (42,45,54,55).

Survival rates are summarized in Table 2. No significant difference in survival was found between male and female patients, and the age at death of the majority of the SLE patient population ranged from 30 to 44 years (54). There are some discrepancies in survival and mortality rates across the geographic regions, which might be a result of differences in access to health care and socioeconomic factors related to delay in treatment and poor treatment compliance; data on these discrepancies are lacking in the cited literature.

Table 1. Mortality of SLE patients by geographic region of interest*

Author, year (ref.)	Geographic region(s), no. of deaths	Ethnicity	SLE duration; age at death (years)	Cause of death (%)
Chen et al, 2006 (90)†	China, Shanghai, 16	NR	11.4 ± 5.2; 35.6 ± 13.1	Infection (31), renal (31), CVD (13), CBV (13), GI vasculitis (6), unknown (6)
Kim et al, 1999 (91)†	Korea, 40	Korean	3.9 ± 1.8; 33.8 ± 13.6	Infection (33), active SLE (25), CVD (18), CBV (10), hematologic (8), pulmonary (3), GI (3), unknown (3)
Chun and Bae, 2005 (92)†	Korea, 10	Korean	NR	Active SLE (80), infection (10), suicide (10)
Ichikawa et al, 1985 (93)†	Japan, 212	NR	NR;33.3 ± 11.3	Infection (35), active SLE (27), CVD (7), CBV (10), GI (7), suicide (6), malignancy (3) hepatic (2), other (3)
Pu et al, 2000 (94)†	Taiwan, 36	ATSI, White	NR	Infection (69), renal (17), pulmonary (14), CBV (6), CNS (6), malignancy (6), unknown (8)
Chang et al, 1998 (95)†	Taiwan, 15	Chinese	NR	Infection (33), renal (20), CNS (20), pulmonary (13), CBV (7), malignancy (7)
Mok et al, 2005 (96)†	Hong Kong, 30	Chinese	5.1 ± 5.9;43.8 ± 17.4	Infection (60), CBV (10), renal (7) CVD (6), malignancy (3), suicide (3), unknown (10)
Wong, 1992 (97)†	Hong Kong, 137	Chinese	NR	Active SLE (60), infection (40)
Lee et al, 1993 (98)†	Hong Kong, 137	Chinese	NR	CBV (45), renal (36), infection (23), CVD (9), suicide (5), GI (5)
Koh et al, 1997 (99)†	Singapore, 67	Chinese, Malaysian, Indian, other	4.0 (range 0.1–20.8); 35.1 ± 14	Active SLE (45), infection (40), thromboembolism (8), malignancy (6), CVD (2)
Bossingham, 2003 (100)†	Australia, 9	ATSI, European, Sikh	9.2; NR	Thromboembolism (or suspected) (67), active SLE or treatment complications (33)
Segasothy and Phillips, 2001 (101)†	Australia, 2	ATSI	1.2; 36	Infection (50), thromboembolism (50)
Anstey et al, 1993 (102)†	Australia, 9	ATSI	2.9; 30	Infection (67), CVD (22), renal (11)
Mu et al, 2018 (47)	China, 45	NR	2.6 (0.5–7.0); NR	Infection (31.1), renal failure (6.7), PAH (6.7), CBV (6.7), NPSLE (4.4), thrombocytopenia (4.4)
Wang et al, 2018 (54)	China, 16	Chinese	NR	Infection (2), severe active SLE (1), sudden cardiac death (1), CBV (1), unknown (2)
Yang et al, 2014 (55)	Singapore, 77	Chinese, Malaysian, Indian, other	NR	NR

* Values are the mean ± SD unless indicated otherwise. ATSI = Aboriginal and Torres Strait Islander people; CBV = cerebrovascular event; CNS = central nervous system; CVD = cardiovascular disease; GI = gastrointestinal; NPSLE = neuropsychiatric SLE; NR = not reported; PAH = pulmonary arterial hypertension; ref. = reference; SLE = systemic lupus erythematosus; SMR = standardized mortality ratio.
† Summarized in ref. 3.

Risk factors for and comorbidity associated with SLE. SLE patients from included studies had preexisting comorbidities at diagnosis, and the risk of developing multiple comorbidities was higher after diagnosis compared to matched controls. SLE is associated with a greater risk of cancer, cardiovascular, renal, liver, rheumatologic, and neurologic diseases as well as hypothyroidism, psychosis, and anemia. Risk factors identified by this TLR for neuropsychiatric SLE included higher SLE Disease Activity Index scores, antiphospholipid antibody positivity, absence of anti–double-stranded DNA antibody at SLE diagnosis and fewer years of education. Presence of neuropsychiatric SLE, especially focal central nervous system neuropsychiatric SLE, increased the risk of mortality in SLE patients (42). Age (30–49 years age group), female sex, and socioeconomic status were reported to be significant risk factors of developing SLE in the APAC region.

Disease burden of SLE. *Treatment outcomes of existing SLE therapies in the APAC region.* Of 27 clinical studies identified by this TLR, 20 reported treatment outcomes of existing therapies in the APAC region (see Supplementary Appendix C, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24431/abstract). A range of study designs were reported, including a pharmacokinetics/pharmacodynamics study, phase III clinical trials, and prospective and retrospective observational studies that assessed a variety of approved and experimental drug treatments (including cyclophosphamide [31,33], belimumab [38,41], rituximab [17,18,27,36,39], epratuzumab [37], tabalumab [35], azathioprine [19,21], hydroxychloroquine [22,24, 40], and glucocorticoids [10,20,26]). However, only hydroxychloroquine, cyclophosphamide, azathioprine, and mycophenolate mofetil (MMF) are routinely used in the APAC region

Table 2. Survival of SLE patients by geographic region*

Author, year (ref.)	Country, City/region	Study period	Ethnicity/age at onset, years	Cases, (no.)	Survival rates (%)		
					1-year	5-year	10-year
Chen et al, 2006 (90)†	China, Shanghai	1980–1998	NR	50	98 (from SLE onset), 98 (from SLE diagnosis)	98 (from SLE onset), 86 (from SLE diagnosis)	84 (from SLE onset), 76 (from SLE diagnosis)
Xie et al, 1998 (103)†	China, Shanghai	1959–1993	NR	566	93 (from SLE onset)	73 (from SLE onset)	60 (from SLE onset)
Kim et al, 1999 (91)†	Korea	1993–1997	Korean	544	98	94	NR
Iseki et al, 1994 (104)†	Japan, Okinawa	1972–1993	Korean	566	97 (from SLE diagnosis)	89 (from SLE diagnosis)	78 (from SLE diagnosis)
Kameda, 1988 (105)†	Japan, Fukuoka City	1975–1977	NR	103 (female)	NR	89	64
Funauchi et al, 2007 (106)†	Japan, Osaka	1975–2004	NR	306	NR	NR	89
Pu et al, 2000 (94)†	Taiwan	1988–1998	Chinese/<50	152	97	88	82
Pu et al, 2000 (94)†	Taiwan	1988–1998	Chinese/50–64	21	76	66	54
Pu et al, 2000 (94)†	Taiwan	1988–1998	Chinese/≥65	21	74	55	55
Chang et al, 1998 (95)†	Taiwan	1983–1996	Chinese (male)	72	85 (from SLE diagnosis)	76 (from SLE diagnosis)	75 (from SLE diagnosis)
Chang et al, 1998 (95)†	Taiwan	1983–1996	Chinese (female)	519	89 (from SLE diagnosis)	81 (from SLE diagnosis)	78 (from SLE diagnosis)
Mok et al, 2005 (96)†	Taiwan	1983–1996	Chinese/16–50	213	NR	94	87
Mok et al, 2005 (96)†	Hong Kong	1991–2003	Chinese/ >50	22	NR	66 (from SLE diagnosis)	44 (from SLE diagnosis)
Wong, 1992 (97)	Hong Kong	1985–1989	Chinese	156	NR	97	94
Anstey et al, 1993 (102)	Australia	1984–1991	ATSI	22	91	60	NR
Lin et al, 2013 (45)	Taiwan	2003–2008	NR; all age groups	103	93.66	80.4‡	NR
Lin et al, 2013 (45)	Taiwan	2003–2008	NR; <40 years	103	94.1	81.8	NR
Mu et al, 2018 (47)	China	2007–2015	NR	911	98.2	95.3	93.7
Wang et al, 2018 (54)	China	2009–2010	Chinese	254	98.4	95.5	93.8

* ATSI = Aboriginal and Torres Strait Islander people; NR = not reported; SLE = systemic lupus erythematosus.
† Summarized in ref. 3.
‡ 6-year survival rate was 93.8%.

due to the lack of approval for other medicines in this region (10,24).

Among drugs targeting the immune system, cyclophosphamide demonstrated similar efficacy to MMF for induction therapy for active lupus nephritis and provided a synergistic effect with tacrolimus to promote remission (31,33). However, cyclophosphamide is associated with significantly more adverse events (AEs) compared with MMF, including anemia and thrombocytopenia. The B cell modulator belimumab demonstrated a higher SLE response index rate compared with placebo in a meta-analysis conducted in China (38), an observational study in Japan (34), and a phase III clinical trial in China, Japan, and Korea (41), along with a manageable safety profile. In contrast, the B cell modulator rituximab, which has been studied as an off-label treatment in Australia, Japan, Korea, and Singapore when conventional immunosuppressive therapy fails (17,18,36,39), was associated with a unfavorable safety profile, including malignancy, anaphylaxis, and death. However, it is important to note that these observations were not part of RCTs, and there was no comparator group to determine the occurrence of such events in matched patients with similar SLE severity who were not given rituximab. Azathioprine in conjunction with a glucocorticoid in the maintenance phase demonstrated a similar low rate of renal relapse when compared with tacrolimus in a clinical trial in China, but had a less favorable safety profile, with AEs such as leukopenia, hepatotoxicity, and gastrointestinal discomfort (19,21).

Long-term glucocorticoid use is associated with burdensome adverse effects for SLE patients in the APAC region. In a

Table 3. Summary of economic burden by geographic region of interest*

Geographic region, study design (ref.)	Economic burden
Australia, retrospective review (18)	Butterly et al reported that the cost of off-label use of rituximab was >AUD $210,000, in a review of data from 2005–2008 at a tertiary hospital.
China, observational cost analysis (78)	Zhang et al assessed treatment adherence and disease burden at an outpatient clinic and demonstrated annual direct (CNY ¥₩33,899) and indirect (CNY ¥8,993) costs per patient for SLE patients, which were higher than costs for patients with ankylosing spondylitis or rheumatoid arthritis.
Hong Kong, retrospective cost of illness study (80)	Zhu et al showed that the mean annual total cost of care was USD $13,307 per SLE patient. The direct costs dominated the total costs, and the costs of inpatient care contributed 52% of the direct costs. SLE patients with CVD (USD $25,051), seizures (USD $28,560), and neuropsychiatric SLE (USD $19,174) incurred significantly higher disease costs compared with the general SLE patient population.
Japan, retrospective claims database study (77)	Tanaka et al demonstrated that the mean total direct medical cost from 2010–2012 was USD $27,004. Costs increased with disease severity for patients with mild (USD $5,549), moderate (USD $15,290), and severe (USD $43,322) disease (analysis of variance, $P < 0.0001$).
Korea, cost analysis from a hospital database and medical records (76)	Park et al reported the estimated overall annual direct medical costs of SLE in South Korea as USD $3,305 per patient (2010 currency), of which 60.4% and 39.6% accounted for inpatient and outpatient cost, respectively. The majority of the direct costs were attributable to diagnostic procedures (USD $1,177) and medications (USD $1,269).
Korea, cost analysis using hospital electronic and interview data (63)	Cho et al reported no significant difference in the total, direct, or indirect costs between patients with high or low disease activity (SLEDAI-2K). Patients with disease damage (SDI), however, incurred significantly higher costs than patients without such damage (KRW ₩4.72 million vs. KRW ₩2.80 million [$P = 0.008$]).
Singapore, retrospective cost analysis (27)	Lateef et al assessed the impact of rituximab in patients with refractory SLE. The costs of hospitalization were reported as SGD $5,989 per patient per year before rituximab treatment, and SGD $5,792 per patient/year after rituximab, including drug cost for rituximab (SGD $3,400 per 500 mg vial). Data suggest that rituximab was the dominant therapy, with equivalent cost and fewer hospitalization days (from 17.1 days per year [range 1.9–49 days] to 0 days per year [range 0–14.8 days; $P = 0.027$]).
Taiwan, insurance database analysis (61)	Cheng et al showed that the average annual medical costs and average SLE-related medical costs were NTD $105,059 and NTD $29,770, respectively (61).
Taiwan, population-based cost analysis (62)	Chiu et al showed that the per patient costs for SLE were USD $1,165 for outpatients, USD $4,238 for inpatients, USD $5,094 for SLE with organ damage, and USD $3,167 for SLE without organ damage.
Taiwan, population-based cost analysis of health care utilization (69)	Lai et al reported the inflation in the annual costs of ambulatory medical care utilization from USD $232 to USD $1,134 over 8 years.

* AUD = Australian dollar; CNY = Chinese Yuan; CVD = cardiovascular disease; KRW = Korean won; NTD = New Taiwan dollar; SDI = Systemic Lupus International Collaborating Clinics/ American College of Rheumatology Damage Index; SGD = Singapore dollar; SLE = systemic lupus erythematous SLEDAI-2K = Systemic Lupus Erythematous Disease Activity Index 2000; USD = United States dollar.

Chinese cohort with SLE, glucocorticoid use was associated with comorbidities such as osteoporosis, osteonecrosis, and Cushing's syndrome, and an increased risk of bacterial infection (20). Avascular necrosis (AVN) is one of the most common causes of organ damage in SLE and can cause severe physical disability, and high cumulative glucocorticoid use was one of the most significant risk factors for AVN, along with the use of immunosuppressants. On the contrary, hydroxychloroquine was associated with a reduced risk of incident diabetes mellitus in SLE patients (22), and was found to decrease mortality rates in patients with SLE in Taiwan (24).

In addition to efficacy and safety, adherence to treatment was assessed in 2 studies (24,78). Of 121 patients in a cohort of Chinese patients with SLE, 51.3% of SLE patients were not adherent to treatment. Experiencing treatment-related AEs was negatively associated with drug adherence (odds ratio 2.185 [95% confidence interval 0.925–5.161], $P < 0.05$) (78). Treatment adherence was reported to improve long-term survival in a cohort of SLE patients who were receiving hydroxychloroquine ($P < 0.001$ versus nonadherent patients), demonstrating the importance of patient adherence to therapy for optimal outcomes in SLE (24).

Economic burden of SLE. Ten studies (18,27,61–63,69, 76–78,80) reported the economic burden of SLE in the APAC region (Table 3). These studies highlighted that SLE was associated with a substantial annual per patient health care expenditure that increased with disease severity, and health care utilization was higher in SLE patients with higher disease activity and organ damage (61–63,76–78,80).

Although medical costs varied widely across the included studies, a high economic burden was consistently reported for the treatment of SLE across the region. The main driver that contributed to economic burden was direct medical costs (including hospitalization, medications, and diagnostic materials). The highest annual direct costs were reported for Japan ($27,000 USD) (77), followed by Hong Kong ($13,307 USD) (80), China ($6,919 USD)

Table 4. Summary of humanistic burden by geographic region of interest*

Geographic region, study design (ref.)	Humanistic burden
Australia	
Cross-sectional questionnaire study (66)	Significant discordance between SLE patient and physician health status concerns; the most important concerns were fatigue and functional measures for patients and organ manifestations for physicians
Validation study (57)	SLE patients with disability, low socioeconomic status, or higher disease activity had significantly worse LIT scores than the general population.
Hong Kong	
Epidemiologic review using electronic health care records (46)	Depressive/anxiety symptoms were independently associated with poorer QoL in SLE patients, and patients with more depressive symptoms were more likely to experience work disability than those without.
Cost of illness study (80)	The number and severity of flares (mild/moderate vs. severe), and the number of organs involved (single-organ vs. multiorgan flare) did not influence the domains of HRQoL measured by the SF-36.
China	
Patient questionnaire (68)	Using the Chinese version of the SLEQoL tool fatigue, education level, disease duration, ESR, and disease activity were the predominant influencers of HRQoL in the Chinese Han SLE population.
Case–control follow-up studies (79,82)	In 2 studies that examined the relationship between polymorphisms of the glucocorticoid receptor gene and HRQoL, specific polymorphisms were significantly associated with improvements in HRQoL in Chinese SLE patients treated with glucocorticoids
Patient survey (59)	High prevalence of depressive and anxiety disorders associated with disease activity among SLE patients.
Taiwan	
Prospective, cross-sectional survey (60)	A study that evaluated HRQoL among patients with AS, rheumatoid arthritis (RA), or SLE found that RA patients have a lower HRQoL than AS and SLE patients.
Prospective, cross-sectional study (64)	Self-help via the Braden Self-Help Model played an important role in alleviating patient-perceived disease severity and improving QoL.
Japan	
Patient survey (65)	HRQoL of Japanese female SLE patients was significantly poorer compared with that of age and sex-matched Japanese normal controls in physical and global health perception as well as in social and emotional status.
Prospective cross-sectional study (74)	In corticosteroid-naive SLE patients (n = 11), prevalence of depression was higher compared with control patients without SLE (n = 2) ($P = 0.035$).
Method validation study (67)	Older age was associated with lower QoL scores in a study by Inoue et al, which validated the Japanese LupusPRO for use in SLE patients.
Prospective study (58)	Baba et al assessed the validity of the SF-36 health survey and its association with the SLEDAI-2K and the SDI over a 2-year period and found that HRQoL measured by the SF-36 health survey was reduced in Japanese SLE patients and was associated with disease damage rather than disease activity.
Validation of questionnaire in an observational cohort (75)	In a study that translated the SLAQ (a patient-reported assessment of subjective disease activity in SLE) into Japanese, Okamoto et al reported that the Japanese SLAQ was comparable to the English version.
Korea	
Multicenter cross-sectional study (72)	In the only study identified by the TLR in Korea reporting on the humanistic burden, Moon et al reported lower QoL and poor sleep quality in SLE patients with fibromyalgia as compared to those without fibromyalgia, especially middle-age and female SLE patients.

* AS = ankylosing spondylitis; HRQoL = health-related quality of life; LupusPRO = lupus patient-reported outcomes; LIT = Lupus Impact Tracker; QoL = quality of life; ref. = reference; SF-36 = Short Form 36 health survey; SLAQ = Site-Level Assessment Questionnaire; SLE = systemic lupus erythematosus; SLEDAI-2K = SLE Disease Activity Index 2000; SDI = Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index; TLR = targeted literature review.

(78), Taiwan ($5,403 USD) (61,62), and South Korea ($3,305 USD) (63,76). Inpatient care was the main contributor to direct costs of treating SLE (78,80). In Hong Kong, the presence of comorbidities, such as cardiovascular disease, seizures, and neuropsychiatric SLE, increased annual health care costs by 44–114% compared with the general SLE population, and therefore comprised the overall health care–utilization burden (80). Clinically, the main drivers of economic burden were increased disease activity and damage (Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index) (63,76, 78,80). One study showed a modest reduction in hospital stays and overall hospital expenditure with rituximab treatment, thereby demonstrating that interventions can impact health care utilization and potentially relieve economic burden (27).

Overall, the economic burden of SLE is driven by disease activity, organ damage, and long-term comorbidities and may be exacerbated by the chronic nature of SLE. Therefore, there is consequently an unmet need for novel therapies that can minimize disease activity and damage in SLE patients and target comorbidities to reduce hospitalization and relieve the economic burden of SLE in the APAC region.

Humanistic burden of SLE. In total, 16 studies (57–60, 64–68,71,72,74,75,79,81,82) reported the humanistic burden of SLE in the APAC region (Table 4). Key drivers of QoL burden were disease activity, longer disease duration (57,68), and physical disability and fatigue (66,57). Low economic status and education level, unemployment, and old age also negatively impacted QoL of SLE patients in the geographic regions of interest to this TLR (57,66–68). Symptoms such as fatigue, pain, poor cognition, and inability to perform usual tasks were identified as patient-perceived unmet needs among SLE patients in the APAC region. Importantly, these health status issues are often not addressed by physician-scored disease activity instruments (66). In addition to SLE symptoms, side effects of treatment regimens, which are known to be burdensome (12,13), may also negatively impact QoL.

QoL measures used in the included studies are presented (see Supplementary Appendix C, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24431/abstract). Due to the degree of heterogeneity in the included studies in terms of patient populations, disease stage, current treatments, and tools employed to assess QoL, a direct comparison between all studies is not feasible. However, there were general trends across the studies, including a consistently reported decrease in QoL that was associated with either living with SLE or receiving treatment for SLE.

Unmet needs of SLE in the APAC region. The findings of this TLR highlight important unmet needs for SLE patients in the APAC region. There is a need for novel therapies that are not only more efficacious than what is currently available, but also have a favorable safety profile for SLE patients. There is also an unmet need for treatment options for patients with severe, refractory disease (such as lupus nephritis). Biomarker tests for earlier and/or improved diagnosis of SLE with difficult-to-diagnose manifestations and tests that predict disease prognosis would allow physicians to personalize SLE treatments for individual patients with SLE and highlight another area of unmet need.

Moreover, there is an unmet need for improved HRQoL for SLE patients in the APAC region. The most commonly reported patient-perceived unmet needs (>65% of patients) were tiredness, pain, not being able to do things that the patient used to do, fear of flares, sleeping problems, anxiety/stress, and feeling down/depressed (4). In addition, given the prevalence of comorbidities in patients with SLE and their impact on SLE treatment costs, there is an unmet need for SLE therapies that address the comorbidity burden in SLE. The chronic, progressive nature of SLE and the high levels of unmet need undoubtedly lead to substantial direct and indirect costs worldwide and highlight that innovation in care is required for SLE patients in the APAC region.

DISCUSSION

The results of the current TLR highlight that the APAC population has higher clinical disease severity, significantly higher mean and maximum SLE disease activity, increased susceptibility to renal involvement, and a higher proportion of autoantibody positivity and organ damage accrual (3). Additionally, SLE in the APAC region is associated with early mortality. High economic burden was consistently reported across the geographies of interest, which indicates that SLE in the APAC region is associated with a substantial annual per-patient health care expenditure, which increases with disease severity. The findings also indicate a significant decrease in QoL associated with SLE in the APAC region.

Long-term glucocorticoid treatment is burdensome for patients and can result in low patient satisfaction and poor adherence. Given that adherence negatively impacts treatment outcomes (24), there is an unmet need for therapies that are less burdensome than glucocorticoids and therefore improve adherence in SLE patients in the APAC region.

The Asian population has a higher risk of developing SLE compared with other ethnicities, given the genetic susceptibility to certain types of polymorphisms as well as environmental factors (84). SLRs that included non-Asian SLE patients (from the APAC region as well as geographic regions beyond the scope of this review), highlight key differences in terms of SLE epidemiology and disease burden compared with the Asian population in the APAC region (2–4,85–88) (see Supplementary Appendix D, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24431/abstract). White individuals across the US, Europe, and Canada have lower rates of SLE compared with Asian populations (3,85). SLE incidence is also lower among the White population in Australia compared with Aboriginal and Torres Strait Islander people (3,4). In a cross-ethnicity study (2), SLE incidence rates were highest in the Afro-Caribbean population, followed by Asian and White populations (31.9, 4.1, and 3.4 per 100,000 person-years, respectively) (2,86). Disease burden has been reported to be more severe among Asians compared with non-Asian populations (87,88). Asian patients have higher disease activity, higher rates of hospitalization due to flares, and a higher incidence of renal disease compared with White patients with SLE (83,87,88). This highlights the need for close monitoring of Asian patients with SLE, who may be at risk of more severe disease.

The present review was conducted according to the PRISMA Statement guidelines; however, there were some limitations. Data gaps made generalization and regional comparisons difficult. For example, the same efficacy outcomes were not reported for all clinical studies, and different HRQoL tools were used to measure humanistic burden. Variations between

geographic regions may be due to differences in reporting and/or diagnostic methods, and deriving an average in terms of overall prevalence, severity, mortality, etc. was not possible. Assessment of quality and/or risk of bias in the included studies was not performed, since this was a TLR. The overall scope of this review was also limited by data gaps specific to the APAC region, which are highlighted below. Future research should focus on addressing the evidence gaps that make understanding the burden of disease specific to APAC patients challenging. Research is also needed in the APAC region to drive drug development and clinical studies specific to this region. SLE is a burdensome disease and therefore innovation in care is needed from both patient and societal perspectives.

This TLR has identified a paucity of data relevant to the APAC region, particularly for Australia, Singapore, and Taiwan. In addition, a consistent definition of SLE is lacking across the APAC region. There is limited recent data on the prevalence and incidence of SLE in the APAC region, which is essential to assess the burden and impact of the disease in these regions.

Clinical studies that assess the efficacy and safety of SLE interventions in specific regions and ethnicities that make up the APAC population are lacking. In addition, studies with abatacept, leflunomide, methotrexate, and mizoribine, as well as real-world data with belimumab in the APAC region are lacking. Several clinical efficacy outcomes relevant to SLE were not identified by the present TLR and represent further gaps in knowledge of SLE in the APAC region. Identification of disease biomarkers and diagnostic tests that could stratify patients and predict treatment response may aid development of successful therapies for SLE, however data on SLE biomarkers is lacking for patients in the APAC region.

Data on disease activity or damage was not available for Singapore or Taiwan, and diagnostic or classification criteria and disease severity tools were not cited. A study from the US perspective found that patients who receive early diagnosis experience lower flare rates and less health care utilization compared with those who are diagnosed later (89); yet despite the importance of prompt diagnosis on both clinical and economic outcomes, data on misdiagnosis rates are lacking in the APAC region.

Limited data was identified to describe the total economic burden of SLE in the APAC region, with no studies reporting data from Australia and Singapore. Analysis of the identified evidence was challenging due to considerable variations in study publication date, adopted currency, reference year, and the interventions used. In addition, data on resource use of patients with SLE is lacking. Further, there was limited geographic coverage of data on humanistic burden of SLE in the APAC region, with no studies reporting QoL specific to Singaporean or Taiwanese populations; direct comparison was challenging.

Although SLE is a disease with a high burden of comorbidities, including cancer, cardiovascular, renal, liver, rheumatologic and neurologic diseases, there is a clear lack of evidence on SLE-related comorbidities and the impact of treatment on such comorbidities, and HRQoL of patients. There is consequently a need for further research to address the evidence gaps for SLE in the APAC region.

Overall, this review has summarized the epidemiology and disease burden of SLE in the APAC region, along with the treatment landscape of SLE. The findings clearly indicate an unmet need for innovation in care, as well as a paucity of data specific to patients with SLE in the APAC region. The lack of relevant, comparable, and robust data presents a significant challenge in analyzing the trends across the regions of interest, and more research in this area is clearly warranted.

ACKNOWLEDGMENT

Sigmatic Limited t/a DRG Abacus conducted the literature review and provided medical writing support.

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. Tsai and Yang 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. I-Ching Tsai, Ya-Wen Yang.

Acquisition of data. I-Ching Tsai, Ya-Wen Yang.

Analysis and interpretation of data. Yoshiya Tanaka, Sean O'Neill, Mengtao Li, I-Ching Tsai, Ya-Wen Yang.

ROLE OF THE STUDY SPONSOR

As the sponsor of the study, Janssen Pharmaceuticals oversaw the study design, data collection, data analysis, and writing of the manuscript, and approved the content of the submitted manuscript. All authors, including employees of Janssen, were involved in the decision to submit for publication. Janssen provided funding of the project execution and professional medical writer who assisted the manuscript preparation and submission. Publication of this article was not contingent upon approval by Janssen Pharmaceuticals.

REFERENCES

1. Decision Resources Group. Insights Report, systemic lupus erythematosus, disease landscape &f forecast. 2018. URL: https://decisionresourcesgroup.com/report/716479-biopharma-systemic-lupus-erythematosus-landscape/.
2. Rees F, Doherty M, Grainge MJ, Lanyon P, Zhang W. The worldwide incidence and prevalence of systemic lupus erythematosus: a systematic review of epidemiological studies. Rheumatology (Oxford) 2017;56:1945–61.
3. Jakes RW, Bae SC, Louthrenoo W, Mok CC, Navarra SV, Kwon N. Systematic review of the epidemiology of systemic lupus erythematosus in the Asia-Pacific region: prevalence, incidence, clinical features, and mortality. Arthritis Care Res (Hoboken) 2012;64:159–68.
4. Nikpour M, Bridge JA, Richter S. A systematic review of prevalence, disease characteristics and management of systemic lupus erythematosus in Australia: identifying areas of unmet need. Intern Med J 2014;44:1170–9.

5. La Paglia GM, Leone MC, Lepri G, Vagelli R, Valentini E, Alunno A, et al. One year in review 2017: systemic lupus erythematosus. Clin Exp Rheumatol 2017;35:551–61.
6. Lai JS, Beaumont JL, Jensen SE, Kaiser K, van Brunt DL, Kao AH, et al. An evaluation of health-related quality of life in patients with systemic lupus erythematosus using PROMIS and Neuro-QoL. Clin Rheumatol 2017;36:555–62.
7. Olesińska M, Saletra A. Quality of life in systemic lupus erythematosus and its measurement. Reumatologia 2018;56:45–54.
8. Tanaka Y, Kubo S, Iwata S, Yoshikawa M, Nakayamada S. B cell phenotypes, signaling and their roles in secretion of antibodies in systemic lupus erythematosus. Clin Immunol 2018;186:21–5.
9. Tunnicliffe DJ, Singh-Grewal D, Kim S, Craig JC, Tong A. Diagnosis, monitoring, and treatment of systemic lupus erythematosus: a systematic review of clinical practice guidelines. Arthritis Care Res (Hoboken) 2015;67:1440–52.
10. Yap DY, Ma MK, Mok MM, Tang CS, Chan TM. Long-term data on corticosteroids and mycophenolate mofetil treatment in lupus nephritis. Rheumatology (Oxford) 2013;52:480–6.
11. Tian J, Luo Y, Wu H, Long H, Zhao M, Lu Q. Risk of adverse events from different drugs for SLE: a systematic review and network meta-analysis. Lupus Sci Med 2018;5:e000253.
12. Gladman DD, Urowitz MB, Rahman P, Ibañez D, Tam LS. Accrual of organ damage over time in patients with systemic lupus erythematosus. J Rheumatol 2003;30:1955.
13. Magder LS, Petri M. Incidence of and risk factors for adverse cardiovascular events among patients with systemic lupus erythematosus. Am J Epidemiol 2012;176:708-19.
14. Lateef A, Petri M. Unmet medical needs in systemic lupus erythematosus. Arthritis Res Ther 2012;14 Suppl 4:S4.
15. The Cochrane Collaboration. The Cochrane Handbook for Systematic Reviews of Interventions, version 5.1. March 2011. URL: http://handbook-5-1.cochrane.org/.
16. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA). PRISMA statement [online]. URL: http://www.prisma-statement.org/PRISMAStatement/PRISMAStatement.aspx.
17. Bang SY, Lee CK, Kang YM, Kim HA, Suh CH, Chung WT, et al. Multicenter retrospective analysis of the effectiveness and safety of rituximab in Korean patients with refractory systemic lupus erythematosus. Autoimmune Diseases 2012;2012:6.
18. Butterly SJ, Pillans P, Horn B, Miles R, Sturtevant J. Off-label use of rituximab in a tertiary Queensland hospital. Intern Med J 2010;40:443–52.
19. Chen D, Lian F, Yuan S, Wang Y, Zhan Z, Ye Y, et al. Association of thiopurine methyltransferase status with azathioprine side effects in Chinese patients with systemic lupus erythematosus. Clin Rheumatol 2014;33:499–503.
20. Chen HL, Shen LJ, Hsu PN, Shen CY, Hall SA, Hsiao FY. Cumulative burden of glucocorticoid-related adverse events in patients with systemic lupus erythematosus: findings from a 12-year longitudinal study. J Rheumatol 2018;45:83.
21. Chen W, Liu Q, Chen W, Tang X, Fu P, Liu F, et al. Outcomes of maintenance therapy with tacrolimus versus azathioprine for active lupus nephritis: a multicenter randomized clinical trial. Lupus 2012;21:944–52.
22. Chen YM, Lin CH, Lan TH, Chen HH, Chang SN, Chen YH, et al. Hydroxychloroquine reduces risk of incident diabetes mellitus in lupus patients in a dose-dependent manner: a population-based cohort study. Rheumatology (Oxford) 2015;54:1244–9.
23. Deng J, Huo D, Wu Q, Yang Z, Liao Y. A meta-analysis of randomized controlled trials comparing tacrolimus with intravenous cyclophosphamide in the induction treatment for lupus nephritis. Tohoku J Exp Med 2012;227:281–8.
24. Hsu CY, Lin YS, Cheng TT, Syu YJ, Lin MS, Lin HF, et al. Adherence to hydroxychloroquine improves long-term survival of patients with systemic lupus erythematosus. Rheumatology (Oxford) 2018;57:1743–51.
25. Hu SC, Yen FL, Wang TN, Lin YC, Lin CL, Chen GS. Immunosuppressive medication use and risk of herpes zoster (HZ) in patients with systemic lupus erythematosus (SLE): a nationwide case-control study. J Am Acad Dermatol 2016;75:49–58.
26. Kwon HH, Bang SY, Won S, Park Y, Yi JH, Joo YB, et al. Synergistic effect of cumulative corticosteroid dose and immunosuppressants on avascular necrosis in patients with systemic lupus erythematosus. Lupus 2018;27:1644–51.
27. Lateef A, Lahiri M, Teng GG, Vasoo S. Use of rituximab in the treatment of refractory systemic lupus erythematosus: Singapore experience. Lupus 2010;19:765–70.
28. Li WG, Ye ZZ, Yin ZH, Zhang K. Clinical and immunological characteristics in 552 systemic lupus erythematosus patients in a southern province of China. Int J Rheum Dis 2017;20:68–75.
29. Mok CC, Tse SM, Chan KL, Ho LY. Effect of immunosuppressive therapies on survival of systemic lupus erythematosus: a propensity score analysis of a longitudinal cohort. Lupus 2018;27:722–7.
30. Mok CC, Ying KY, Yim CW, Siu YP, Tong KH, To CH, et al. Tacrolimus versus mycophenolate mofetil for induction therapy of lupus nephritis: a randomised controlled trial and long-term follow-up. Ann Rheum Dis 2016;75:30–6.
31. Onishi A, Sugiyama D, Tsuji G, Nakazawa T, Kogata Y, Tsuda K, et al. Mycophenolate mofetil versus intravenous cyclophosphamide for induction treatment of proliferative lupus nephritis in a Japanese population: a retrospective study. Mod Rheumatol 2013;23:89–96.
32. Park EJ, Jung H, Hwang J, Kim H, Lee J, Ahn JK, et al. Pregnancy outcomes in patients with systemic lupus erythematosus: a retrospective review of 62 pregnancies at a single tertiary center in South Korea. Int J Rheum Dis 2014;17:887–97.
33. Sakai R, Kurasawa T, Nishi E, Kondo T, Okada Y, Shibata A, et al. Efficacy and safety of multitarget therapy with cyclophosphamide and tacrolimus for lupus nephritis: a prospective, single-arm, single-centre, open label pilot study in Japan. Lupus 2018;27:273–82.
34. Tanaka Y, Bass D, Chu M, Egginton S, Ji B, Struemper H, et al. Efficacy and safety of intravenous belimumab in Japanese patients with systemic lupus erythematosus: a subgroup analysis of a phase 3 randomized placebo-controlled trial. Mod Rheumatol 2019;29:452–60.
35. Tanaka Y, Takeuchi T, Akashi N, Takita Y, Kovacs B, Kariyasu S. Efficacy and safety of tabalumab plus standard of care in Japanese patients with active systemic lupus erythematosus: subgroup analyses of the ILLUMINATE-1 study. Mod Rheumatol 2017;27:284–91.
36. Tanaka Y, Takeuchi T, Miyasaka N, Sumida T, Mimori T, Koike T, et al. Efficacy and safety of rituximab in Japanese patients with systemic lupus erythematosus and lupus nephritis who are refractory to conventional therapy. Mod Rheumatol 2016;26:80–6.
37. Tsuru T, Tanaka Y, Kishimoto M, Saito K, Yoshizawa S, Takasaki Y, et al. Safety, pharmacokinetics, and pharmacodynamics of epratuzumab in Japanese patients with moderate-to-severe systemic lupus erythematosus: results from a phase 1/2 randomized study. Mod Rheumatol 2016;26:87–93.
38. Wei LQ, Liang YG, Zhao Y, Liang HT, Qin DC, She MC. Efficacy and safety of belimumab plus standard therapy in patients with systemic lupus erythematosus: a meta-analysis. Clin Ther 2016;38:1134–40.
39. Wongseelashote S, Tayal V, Bourke PF. Off-label use of rituximab in autoimmune disease in the top end of the northern territory, 2008-2016. Intern Med J 2018;48:165–72.
40. Yokogawa N, Eto H, Tanikawa A, Ikeda T, Yamamoto K, Takahashi T, et al. Effects of hydroxychloroquine in patients with cutaneous lupus erythematosus: a multicenter, double-blind,

randomized, parallel-group trial. Arthritis Rheumatol 2017;69: 791–9.
41. Zhang F, Bae SC, Bass D, Chu M, Egginton S, Gordon D, et al. A pivotal phase III, randomised, placebo-controlled study of belimumab in patients with systemic lupus erythematosus located in China, Japan and South Korea. Ann Rheum Dis 2018;77:355–63.
42. Ahn GY, Kim D, Won S, Song ST, Jeong HJ, Sohn IW, et al. Prevalence, risk factors, and impact on mortality of neuropsychiatric lupus: a prospective, single-center study. Lupus 2018;27:1338–47.
43. Cheng Y, Li M, Zhao J, Ye Z, Li C, Li X, et al. Chinese SLE Treatment and Research Group (CSTAR) registry: VIII. Influence of socioeconomic and geographical variables on disease phenotype and activity in Chinese patients with SLE. Int J Rheum Dis 2018;21: 716–24.
44. Ju JH, Yoon SH, Kang KY, Kim IJ, Kwok SK, Park SH, et al. Prevalence of systemic lupus erythematosus in South Korea: an administrative database study. J Epidemiol 2014;24:295–303.
45. Lin WH, Guo CY, Wang WM, Yang DC, Kuo TH, Liu MF, et al. Incidence of progression from newly diagnosed systemic lupus erythematosus to end stage renal disease and all-cause mortality: a nationwide cohort study in Taiwan. Int J Rheum Dis 2013;16: 747–53.
46. Mok CC. Epidemiology and survival of systemic lupus erythematosus in Hong Kong Chinese. Lupus 2011;20:767–71.
47. Mu L, Hao Y, Fan Y, Huang H, Yang X, Xie A, et al. Mortality and prognostic factors in Chinese patients with systemic lupus erythematosus. Lupus 2018;27:1742–52.
48. Ng X, Low AH, Chew LC, Chong YY, Fong KY, Lui NL, et al. Disease patterns of rheumatology outpatients seen in a tertiary hospital serving a multi-ethnic, urban Asian population in Singapore. Int J Rheum Dis 2013;16:273–8.
49. Nossent J, Raymond W, Kang A, Wong D, Ognjenovic M, Chakera A. The current role for clinical and renal histological findings as predictor for outcome in Australian patients with lupus nephritis. Lupus 2018;27:1838–46.
50. Ong C, Nicholls K, Becker G. Ethnicity and lupus nephritis: an Australian single centre study. Intern Med J 2011;41:270–8.
51. Osio-Salido E, Manapat-Reyes H. Epidemiology of systemic lupus erythematosus in Asia. Lupus 2010;19:1365–73.
52. See LC, Kuo CF, Chou IJ, Chiou MJ, Yu KH. Sex- and age-specific incidence of autoimmune rheumatic diseases in the Chinese population: a Taiwan population-based study. Semin Arthritis Rheum 2013; 43:381–6.
53. Shim JS, Sung YK, Joo YB, Lee HS, Bae SC. Prevalence and incidence of systemic lupus erythematosus in South Korea. Rheumatol Int 2014;34:909–17.
54. Wang Z, Li M, Wang Y, Xu D, Wang Q, Zhang S, et al. Long-term mortality and morbidity of patients with systemic lupus erythematosus: a single-center cohort study in China. Lupus 2018;27: 864–9.
55. Yang Y, Thumboo J, Earnest A, Yong SL, Fong KY. The effect of comorbidity on hospital mortality in patients with SLE from an Asian tertiary hospital. Lupus 2014;23:714–20.
56. Yu KH, See LC, Kuo CF, Chou IJ, Chou MJ. Prevalence and incidence in patients with autoimmune rheumatic diseases: a nationwide population-based study in Taiwan. Arthritis Care Res (Hoboken) 2013;65:244–50.
57. Antony A, Kandane-Rathnayake RK, Ko T, Boulos D, Hoi AY, Jolly M, et al. Validation of the Lupus Impact Tracker in an Australian patient cohort. Lupus 2017;26:98–105.
58. Baba S, Katsumata Y, Okamoto Y, Kawaguchi Y, Hanaoka M, Kawasumi H, et al. Reliability of the SF-36 in Japanese patients with systemic lupus erythematosus and its associations with disease activity and damage: a two-consecutive year prospective study. Lupus 2018;27:407–16.
59. Bai R, Liu S, Zhao Y, Cheng Y, Li S, Lai A, et al. Depressive and anxiety disorders in systemic lupus erythematosus patients without major neuropsychiatric manifestations. J Immunol Res 2016. Doi: https://doi.org/10.1155/2016/2829018. E-pub ahead of print.
60. Chen HH, Chen DY, Chen YM, Lai KL. Health-related quality of life and utility: comparison of ankylosing spondylitis, rheumatoid arthritis, and systemic lupus erythematosus patients in Taiwan. Clin Rheumatol 2017;36:133–42.
61. Cheng J, Huang C, Hsu C. Medical care and costs of patients with systemic lupus erythematosus in Taiwan. Value Health 2015;18: A676.
62. Chiu YM, Chuang MT, Lang HC. Medical costs incurred by organ damage caused by active disease, comorbidities and side effect of treatments in systemic lupus erythematosus patients: a Taiwan nationwide population-based study. Rheumatol Int 2016; 36:1507–14.
63. Cho JH, Chang SH, Shin NH, Choi BY, Oh HJ, Yoon MJ, et al. Costs of illness and quality of life in patients with systemic lupus erythematosus in South Korea. Lupus 2014;23:949–57.
64. Chuang TH, Lin KC, Gau ML. Validation of the braden self-help model in women with systemic lupus erythematosus. J Nurs Res 2010;18:206–14.
65. Furukawa M, Kiyohara C, Horiuchi T, Tsukamoto H, Mitoma H, Kimoto Y, et al. Quality of life in Japanese female patients with systemic lupus erythematosus: evaluation using the Short Form 36 Health Survey. Mod Rheumatol 2016;26:240–7.
66. Golder V, Ooi JJY, Antony AS, Ko T, Morton S, Kandane-Rathnayake R, et al. Discordance of patient and physician health status concerns in systemic lupus erythematosus. Lupus 2018;27: 501–6.
67. Inoue M, Shiozawa K, Yoshihara R, Yamane T, Shima Y, Hirano T, et al. The Japanese LupusPRO: a cross-cultural validation of an outcome measure for lupus. Lupus 2017;26:849–56.
68. Jiang HZ, Lin ZG, Li HJ, Du Q, Tian W, Wang SY, et al. The Chinese version of the SLEQOL is a reliable assessment of health-related quality of life in Han Chinese patients with systemic lupus erythematosus. Clin Rheumatol 2018;37:151–60.
69. Lai NS, Tsai TY, Koo M, Huang KY, Tung CH, Lu MC. Patterns of ambulatory medical care utilization and rheumatologist consultation predating the diagnosis of systemic lupus erythematosus: a national population-based study. PLoS One 2014;9:e101485.
70. Leong KP, Chong EY, Kong KO, Chan SP, Thong BY, Lian TY, et al. Discordant assessment of lupus activity between patients and their physicians: the Singapore experience. Lupus 2010;19:100–6.
71. Mok CC, Chan KL, Ho LY. Association of depressive/anxiety symptoms with quality of life and work ability in patients with systemic lupus erythematosus. Clin Exp Rheumatol 2016;34:389–95.
72. Moon SJ, Kang KY, Kwok SK, Ju JH, Hong YS, Park SH, et al. Differences in quality of life determinants according to the presence of fibromyalgia in middle-aged female patients with systemic lupus erythematosus: a multicenter, cross-sectional, single-ethnicity cohort. Int J Rheum Dis 2018;21:1173–84.
73. Moorthy LN, Baldino ME, Kurra V, Puwar D, Llanos A, Peterson MG, et al. Relationship between health-related quality of life, disease activity and disease damage in a prospective international multicenter cohort of childhood onset systemic lupus erythematosus patients. Lupus 2017;26:255–65.
74. Nishimura K, Omori M, Katsumata Y, Sato E, Kawaguchi Y, Harigai M, et al. Psychological distress in corticosteroid-naive patients with systemic lupus erythematosus: a prospective cross-sectional study. Lupus 2016;25:463–71.

75. Okamoto Y, Katsumata Y, Baba S, Kawaguchi Y, Gono T, Hanaoka M, et al. Validation of the Japanese version of the Systemic Lupus Activity Questionnaire that includes physician-based assessments in a large observational cohort. Lupus 2016;25:486–95.
76. Park SY, Joo YB, Shim J, Sung YK, Bae SC. Direct medical costs and their predictors in South Korean patients with systemic lupus erythematosus. Rheumatol Int 2015;35:1809–15.
77. Tanaka Y, Mizukami A, Kobayashi A, Ito C, Matsuki T. Disease severity and economic burden in Japanese patients with systemic lupus erythematosus: a retrospective, observational study. Int J Rheum Dis 2018;21:1609–18.
78. Zhang L, Lu GH, Ye S, Wu B, Shen Y, Li T. Treatment adherence and disease burden of individuals with rheumatic diseases admitted as outpatients to a large rheumatology center in Shanghai, China. Patient Prefer Adherence 2017;11:1591–601.
79. Zhang M, Li SS, Xie QM, Xu JH, Sun XX, Pan FM, et al. Associations of HSP90AA2 gene polymorphisms with disease susceptibility, glucocorticoids efficacy and health-related quality of life in Chinese systemic lupus erythematosus patients. Genes Genomics 2018;40:1069–79.
80. Zhu TY, Tam LS, Lee VW, Lee KK, Li EK. Systemic lupus erythematosus with neuropsychiatric manifestation incurs high disease costs: a cost-of-illness study in Hong Kong. Rheumatology (Oxford) 2009; 48:564–8.
81. Zhu TY, Tam LS, Lee VW, Lee KK, Li EK. Relationship between flare and health-related quality of life in patients with systemic lupus erythematosus. J Rheumatol 2010;37:568–73.
82. Zou YF, Xu JH, Pan FM, Tao JH, Xu SQ, Xiao H, et al. Glucocorticoid receptor genetic polymorphisms is associated with improvement of health-related quality of life in Chinese population with systemic lupus erythematosus. Clin Rheumatol 2015;34:1537–44.
83. Thumboo J, Wee HL. Systemic lupus erythematosus in Asia: is it more common and more severe? APLAR J Rheumatol 2006;9:320–6.
84. Deng Y, Tsao BP. Genetic susceptibility to systemic lupus erythematosus in the genomic era. Nat Rev Rheumatol 2010;6:683–92.
85. Danchenko N, Satia JA, Anthony MS. Epidemiology of systemic lupus erythematosus: a comparison of worldwide disease burden. Lupus 2006;15:308–18.
86. Hopkinson ND, Doherty M, Powell RJ. Clinical features and race-specific incidence/prevalence rates of systemic lupus erythematosus in a geographically complete cohort of patients. Ann Rheum Dis 1994;53:675–80.
87. Golder V, Connelly K, Staples M, Morand E, Hoi A. Association of Asian ethnicity with disease activity in SLE: an observational study from the Monash Lupus Clinic. Lupus 2013;22:1425–30.
88. Boers A, Li Q, Wong M, Miller M, Littlejohn G. Differences in SLE disease activity between patients of Caucasian and South-East Asian/Chinese background in an Australian hospital. APLAR J Rheumatol 2006;9:43–8.
89. Oglesby A, Korves C, Laliberte F, Dennis G, Rao S, Suthoff ED, et al. Impact of early versus late systemic lupus erythematosus diagnosis on clinical and economic outcomes. Appl Health Econ Health Policy 2014;12:179–90.
90. Chen S, Gu YY, Bao CD, Chen SL. An 18-year follow-up study of a lupus cohort in Shanghai. APLAR J Rheumatol 2006;9: 327–30.
91. Kim WU, Min JK, Lee SH, Park SH, Cho CS, Kim HY. Causes of death in Korean patients with systemic lupus erythematosus: a single center retrospective study. Clin Exp Rheumatol 1999;17: 539–45.
92. Chun BC, Bae SC. Mortality and cancer incidence in Korean patients with systemic lupus erythematosus: results from the Hanyang lupus cohort in Seoul, Korea. Lupus 2005;14:635–8.
93. Ichikawa Y, Tsunematsu T, Yokohari R, Tanimoto K, Sakane T, Yoshida H, et al. [Multicenter study of causes of death in systemic lupus erythematosus-a report from the Subcommittee for Development of Therapy, the Research Committee for Autoimmune Diseases Supported by the Ministry of Health and Welfare]. Ryumachi 1985;25:258–64. In Japanese.
94. Pu SJ, Luo SF, Wu YJ, Cheng HS, Ho HH. The clinical features and prognosis of lupus with disease onset at age 65 and older. Lupus 2000;9:96–100.
95. Chang DM, Chang CC, Kuo SY, Chu SJ, Chang ML. The clinical features and prognosis of male lupus in Taiwan. Lupus 1998;7: 462–8.
96. Mok CC, Mak A, Chu WP, To CH, Wong SN. Long-term survival of southern Chinese patients with systemic lupus erythematosus: a prospective study of all age-groups. Medicine (Baltimore) 2005;84: 218–24.
97. Wong KL. Pattern of SLE in Hong Kong Chinese: a cohort study. Scand J Rheumatol 1992;21:289–96.
98. Lee SS, Li CS, Li PC. Clinical profile of Chinese patients with systemic lupus erythematosus. Lupus 1993;2:105–9.
99. Koh ET, Seow A, Leong KH, Chng HH. SLE mortality in an oriental population. Lupus 1997;6:27–31.
100. Bossingham D. Systemic lupus erythematosus in the far north of Queensland. Lupus 2003;12:327–31.
101. Segasothy M, Phillips PA. Systemic lupus erythematosus in Aborigines and Caucasians in central Australia: a comparative study. Lupus 2001;10:439–44.
102. Anstey NM, Bastian I, Dunckley H, Currie BJ. Systemic lupus erythematosus in Australian Aborigines: high prevalence, morbidity and mortality. Aust N Z J Med 1993;23:646–51.
103. Xie SK, Feng SF, Fu H. Long term follow-up of patients with systemic lupus erythematosus. J Dermatol 1998;25:367–73.
104. Iseki K, Miyasato F, Oura T, Uehara H, Nishime K, Fukiyama K. An epidemiologic analysis of end-stage lupus nephritis. Am J Kidney Dis 1994;23:547–54.
105. Kameda S. Epidemiologic study of systemic lupus erythematosus in Fukuoka population. Fukuoka Igaku Zasshi 1988;79:571–8.
106. Funauchi M, Shimadzu H, Tamaki C, Yamagata T, Nozaki Y, Sugiyama M, et al. Survival study by organ disorders in 306 Japanese patients with systemic lupus erythematosus: results from a single center. Rheumatol Int 2007;27:243–9.

Arthritis Care & Research
Vol. 74, No. 2, February 2022, pp 199–207
DOI 10.1002/acr.24461

American College of Rheumatology
Empowering Rheumatology Professionals

Timing of Childhood-Onset Systemic Lupus Erythematosus Diagnosis Relative to Menarche and the Impact on Final Adult Height

Watchareewan Sontichai,[1] Fangming Liao,[2] Daniela Dominguez,[2] Deborah M. Levy,[2] Muna Al Mutairi,[2] Lawrence Ng,[2] Frank Silverio,[2] Earl D. Silverman,[2] Jonathan D. Wasserman,[2] and Linda T. Hiraki[2]

Objective. The aim of this study was to examine the impact of timing of a childhood-onset systemic lupus erythematosus (SLE) diagnosis relative to menarchal status, on final height, accounting for disease-associated factors.

Methods. We conducted a cohort study of female patients age <18 years at childhood-onset SLE diagnosis, followed at a tertiary care pediatric center from July 1982 to March 2016 and restricted to patients with documented age of menarche and final height. We compared final height between patients diagnosed pre- and postmenarche. We tested the association of the timing of childhood-onset SLE diagnosis with final height, adjusted for ethnicity, in linear regression models. We performed subgroup analyses of patients with growth during follow-up, additionally adjusting for average daily corticosteroid dose and disease activity.

Results. Of 401 female childhood-onset SLE patients in the study, 115 patients (29%) were diagnosed premenarche and 286 (71%) postmenarche. Patients diagnosed premenarche were older at menarche compared with patients diagnosed postmenarche (mean ± SD age 13.5 ± 1.4 versus 12.5 ± 1.3 years; $P < 0.001$). The mean ± SD final height for girls diagnosed postmenarche (161.4 ± 6.9 cm) was greater than for those diagnosed premenarche (158.8 ± 7.3 cm; $P = 0.001$). In regression analysis, those diagnosed postmenarche were significantly taller than those diagnosed premenarche, as adjusted for ethnicity and disease severity (mean ± SD $\beta = 2.6 \pm 0.7$ cm; $P = 0.0006$).

Conclusion. In this large cohort study of girls with childhood-onset SLE, patients diagnosed postmenarche achieved a taller final height than those diagnosed premenarche, even after accounting for ethnicity and disease severity.

INTRODUCTION

Adult height is determined by the interaction of multiple factors, including genetics and environmental influences such as nutrition and socioeconomic status (1–3). Chronic medical conditions in childhood can adversely impact growth. Specifically, active, chronic disease can limit food intake, impair nutrient absorption, cause direct nutrient losses, increase metabolic requirements, and impair transport to target tissues (4). Additionally, some treatments such as corticosteroids may contribute to growth retardation or accelerate epiphyseal closure (5).

Systemic lupus erythematosus (SLE) is a chronic inflammatory autoimmune disease that can affect any organ system. The clinical course varies from mild to a severe, life-threatening disease. Up to 20% of patients with SLE are estimated to have been diagnosed during childhood and adolescence (6). Childhood-onset SLE is associated with higher disease activity at presentation and during clinical course, a greater prevalence of renal and neuropsychiatric lupus, and more disease-associated organ damage accrual, compared to adult-onset SLE (6–8). Another difference between childhood-onset SLE and adult-onset SLE are the potential impacts of disease and medications on linear growth and final height, effects seen only in patients with childhood-onset SLE.

Linear growth accelerates markedly in midpuberty, in the context of a pubertal growth spurt. There is wide variation in

Dr. Hiraki's work was supported by a Canadian Institute of Health Research Project Scheme grant, by the Arthritis Society, and by a Stars Award.

[1]Watchareewan Sontichai, MD: Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada, and Chiang Mai University, Chiang Mai, Thailand; [2]Fangming Liao, MSc, Daniela Dominguez, MSc, Deborah M. Levy, MD, MPH, Muna Al Mutairi, MD, Lawrence Ng, BSc, Frank Silverio, BSc, Earl D. Silverman, MD, Jonathan D. Wasserman, MD, PhD, Linda T. Hiraki, MD, FRCPC, ScD: Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada.

No potential conflicts of interest relevant to this article were reported.

Address correspondence to Linda T. Hiraki, MD, FRCPC, ScD, Clinician Scientist, Division of Rheumatology, Child Health Evaluative Sciences, Research Institute, The Hospital for Sick Children, PGCRL, 686 Bay Street, Toronto ON M5G 0A4, Canada. Email: Linda.hiraki@sickkids.ca.

Submitted for publication January 27, 2020; accepted in revised form September 15, 2020.

SIGNIFICANCE & INNOVATIONS

- This is the largest study to date examining the relationship between the timing of childhood-onset systemic lupus erythematosus diagnosis in relation to menarche and the impact on adult height.
- Patients diagnosed postmenarche achieved a taller final height by 2.6 cm than those diagnosed premenarche.
- Patients diagnosed premenarche had a longer growth period vulnerable to impairment by disease and therapy, compared to those diagnosed postmenarche.

timing of the growth spurt. In girls, peak height velocity occurs, on average, 6 months prior to menarche (9). The growth spurt lasts for approximately 2 years until girls attain their final adult height. As a result, the timing of childhood-onset SLE diagnosis may have significant impacts on linear growth and, subsequently, final adult height related to disease activity and medication side-effects. Few studies have described the final height achieved in patients diagnosed with childhood-onset SLE (10,11). No prior study has examined the relationship of menarchal status at childhood-onset SLE diagnosis and final adult height. The aim of this study was to examine the impact of timing of childhood-onset SLE diagnosis relative to menarchal status, on final height, accounting for ethnicity and disease-associated factors.

PATIENTS AND METHODS

Study design. We conducted a cohort study of female patients diagnosed with childhood-onset SLE and followed in our multiethnic Lupus Clinic at The Hospital for Sick Children, Toronto from July 1, 1982 to March 31, 2016. All patients in our study fulfilled ≥4 American College of Rheumatology (ACR) (12) and/or Systemic Lupus International Collaborating Clinics (SLICC) classification criteria for SLE (13) and were diagnosed prior to age 18 years. Clinical and laboratory data were prospectively collected and stored in a central database. We restricted our cohort to include patients with documented date of menarche and those who achieved final height during our study period. We excluded those with documented vertebral collapse confirmed by plain radiography of the vertebral spine (n = 15), as vertebral collapse would affect our main outcome of interest, height. We also excluded patients who received growth hormone therapy (n = 4). The study was approved by the Research Ethics Board at The Hospital for Sick Children (REB# 1000052416).

We extracted the date of childhood-onset SLE diagnosis and self-reported date of menarche from our database, to determine menarchal status at the time of childhood-onset SLE diagnosis. We divided patients into 2 groups: patients diagnosed with childhood-onset SLE before menarche (premenarche group) and those who were diagnosed after menarche (postmenarche group). Among those patients included in our final cohort, we extracted physician and self-rated Tanner stage from our database (14). There is evidence that self-rated Tanner staging is a valid measure of evaluating sexual maturation in epidemiologic studies (15–17). We restricted our data to the first Tanner stage recorded within 365 days before or after childhood-onset SLE diagnosis.

Final height. Barefoot standing height measurements (cm) were recorded with a stadiometer at each clinic visit. In patients who lacked a recorded longitudinal height in the database, we extracted data from hospital medical records. Height measurements were quality checked for potential measurement and data entry errors by reviewing height trajectories for individual patients. An outlier height measurement was defined as a height value ≥5 cm different from earlier or later measures, within 1 year. These outlier values were replaced with the nearest preceding or subsequent height measurement. We defined final height as the height when growth velocity was <1 cm per year for 2 consecutive years or the height at age 18 years, whichever occurred first. Standardization of height was expressed as a Z score using the 2007 World Health Organization (WHO) growth reference (18).

Additional covariates. From the database, we extracted self-reported ethnicity, categorized as European, East and Southeast Asian, African and Caribbean, South Asian, Hispanic and Amerindian, and mixed ethnicity (Canada census categories). Analyses included a category for missing ethnicity. We reviewed ACR and SLICC SLE clinical and laboratory features from diagnosis to the last clinic visit. Lupus nephritis was defined as confirmed by biopsy results and was subclassified based on either the WHO or International Society of Nephrology/Renal Pathology Society 2003 classification criteria (19,20). Laboratory features included the presence of hemolytic anemia, leukopenia (<4,000/mm^3), lymphopenia (<1,500/mm^3), thrombocytopenia (<100,000/mm^3), antinuclear antibody, anti–double-stranded DNA (anti-dsDNA), anti-Sm, anti-RNP, anti-Ro, anti-La, anticardiolipin antibodies, and lupus anticoagulant.

To study the association of steroid treatment, disease activity, and disease damage with final height, we performed a subgroup analysis in patients with demonstrated growth following childhood-onset SLE diagnosis. This subgroup analysis included all those diagnosed premenarche and a subset of patients diagnosed postmenarche with continued growth (≥1 cm of growth per year for 2 consecutive years). We calculated disease duration as the date from SLE diagnosis to the date of achieving final height. The cumulative corticosteroid dose in prednisone equivalents was calculated from the date of diagnosis to the date of achieving final height and reported in mg/kg of body weight (21).

We extracted data on disease activity score measured by the Systemic Lupus Erythematosus Disease Activity Index (SLEDAI) at each clinic visit (22). A cumulative SLEDAI score over time was determined by calculating the area under the curve of serial measurements of the SLEDAI score from diagnosis to date of achieving final height using the trapezoidal rule (23). We calculated the average daily SLEDAI score by calculating the cumulative SLEDAI score over time divided by disease duration. We also assessed irreversible damage using the SLICC/ACR Damage Index (SDI) (24). Parental height was available in a subset of our population. We compared the proportion of patients whose final height was shorter than the predicted midparental height range (25), between the premenarche and postmenarche groups.

Statistical analysis. We compared means ± SDs of final height and final height Z scores between patients in the premenarche and postmenarche groups using Student's *t*-tests. We compared frequencies of clinical and laboratory features between premenarche and postmenarche groups using chi-square and Fisher's exact tests. A Bonferroni correction was applied and set the threshold for significance at 0.005 for 10 variables of clinical and laboratory feature comparisons.

We tested the association of premenarche childhood-onset SLE diagnosis (compared with postmenarche) and final height using univariate and multivariable linear regression models, adjusted for ethnicity categories and the presence of lupus nephritis and neuropsychiatric features, as indicators of disease severity, and anti-dsDNA antibodies, an SLE manifestation more prevalent in the premenarche group compared to the postmenarche group. We did not include age of childhood-onset SLE diagnosis in our multivariable models since it was highly collinear with our main exposure, menarche status at diagnosis. *P* values less than 0.05 were used to indicate statistically significant differences.

To assess the effect of disease activity and steroid exposure on final height, we completed a subgroup analysis of patients with postmenarche-diagnosed childhood-onset SLE, with demonstrated growth following childhood-onset SLE diagnosis, and those diagnosed before menarche. We restricted the analysis to this subpopulation since they had demonstrated growth, and

Table 1. Baseline characteristics of patients based on timing of childhood-onset SLE diagnosis*

Characteristic	Premenarche (n = 115)	Postmenarche (n = 286)	*P*
Age at diagnosis, mean ± SD years	10.7 ± 2.2	15.1 ± 1.6	<0.001
Age at menarche, mean ± SD years	13.5 ± 1.4	12.5 ± 1.4	<0.001
Self-reported ethnicity†			
European	33 (28.7)	81 (28.3)	0.65
East and Southeast Asian	29 (25.2)	77 (26.9)	–
African and Caribbean	21 (18.3)	48 (16.8)	–
South Asian	19 (16.5)	36 (12.6)	–
Hispanic and Amerindian	4 (3.5)	13 (4.5)	–
Mixed ethnicity	7 (6.1)	16 (5.6)	–
SLE clinical features			
Malar rash	93 (80.9)	233 (81.5)	0.89
Discoid rash	6 (5.2)	9 (3.1)	0.32
Photosensitivity	43 (37.4)	85 (29.7)	0.14
Oral or nasal ulcers	56 (48.7)	112 (39.2)	0.08
Arthritis	82 (71.3)	207 (72.4)	0.83
Serositis	23 (20.0)	41 (14.3)	0.16
Neuropsychiatric lupus	29 (25.2)	77 (26.9)	0.73
Lupus nephritis	60 (52.2)	101 (35.3)	0.001
Class II	3 (5.0)	9 (9.1)	0.78
Class III or IV	42 (70.0)	58 (58.6)	0.0009
Class III or IV + V	5 (8.3)	13 (13.1)	0.36
Class V	9 (15.0)	20 (20.2)	0.42
No biopsy	1 (1.7)	1 (1.0)	0.50
SLE laboratory features			
Hematologic involvement	100 (87.0)	237 (82.9)	0.31
Antinuclear antibodies	115 (100.0)	283 (99.0)	0.27
Anti-dsDNA	94 (81.7)	187 (65.4)	0.001
Anti-Sm	57 (49.6)	99 (34.6)	0.005
Anti-RNP	53 (46.1)	121 (42.3)	0.49
Anti-Ro	57 (49.6)	105 (36.7)	0.02
Anti-La	25 (21.7)	46 (16.1)	0.18
Lupus anticoagulant	25 (21.7)	43 (15.0)	0.11
Anticardiolipin	55 (47.8)	111 (38.8)	0.10

* Values are the number (%) unless indicated otherwise. Anti-dsDNA = anti–double-stranded DNA; SLE = systemic lupus erythematosus.
† Missing ethnicity data for 2 patients in premenarche group and 15 patients in postmenarche group.

hence their final height was potentially impacted by measured disease activity and steroid exposure. We compared medians and interquartile ranges (IQRs) of disease duration, average daily steroid dose, and the average daily SLEDAI score between the 2 groups using Kruskal-Wallis tests. The time to first SDI score of ≥1 was calculated according to the Kaplan–Meier survival analysis. Among girls in the postmenarche group, we compared final height, age of diagnosis, age of menarche, and time from menarche to disease in girls with demonstrated growth and those without demonstrated growth. Sensitivity analyses added Tanner stage ratings to our main analyses and tested the association of Tanner stage and final height in the same univariate and multivariable main models. Statistical analyses were performed by using SPSS statistical software and RStudio.

RESULTS

Our cohort included 564 female patients diagnosed with childhood-onset SLE and followed in our center during the study period. Of those patients, 420 (75%) had age of menarche and age of achieving final height documented in the medical record. We excluded 15 patients with vertebral collapse and 4 patients who received growth hormone therapy. A total of 401 patients were included in the study (see Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24461/abstract). There was no difference in baseline demographic or clinical features between those excluded and included in the study.

The mean age of childhood-onset SLE diagnosis in the total cohort was 13.8 ± 2.7 years. The mean ± SD age of menarche was 12.8 ± 1.5 years, with 115 patients (29%) diagnosed with childhood-onset SLE prior to menarche and 286 patients (71%) diagnosed postmenarche. Patients in the premenarche group were older at menarche compared with patients in the postmenarche group (mean ± SD age 13.5 ± 1.4 versus 12.5 ± 1.3 years; $P < 0.001$).

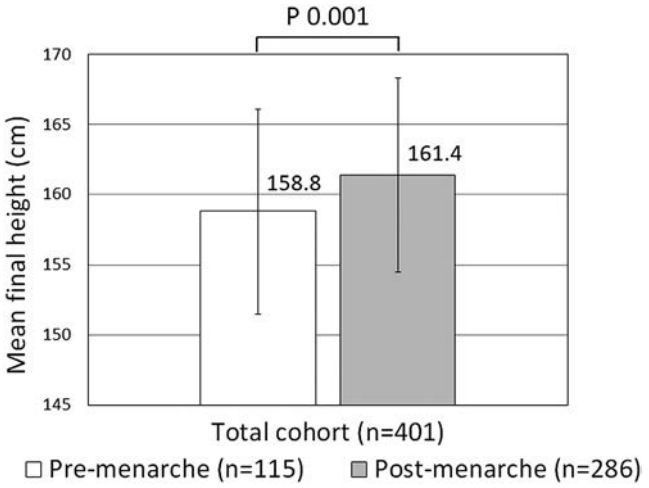

Figure 1. Means of final height of patients based on timing of childhood-onset systemic lupus erythematosus (SLE) diagnosis. Bar graph comparing the mean final height (cm) of patients diagnosed with childhood-onset SLE premenarche and postmenarche. Whiskers represent the SD. *P* value was calculated from an unpaired *t*-test.

There were 17 patients with missing ethnicity data from the record. The remaining cohort was comprised of patients of European (30%), East and Southeast Asian (28%), African and Caribbean (18%), South Asian (14%), Hispanic and Amerindian (4%) and mixed ethnicity (6%), reflecting the ethnic diversity of our patient population. Timing of menarche did not differ significantly among ethnic groups (Table 1).

The prevalence of SLE clinical and laboratory features from diagnosis to last clinic visit is shown in Table 1. Clinical features were not significantly different between premenarche and postmenarche groups, except for lupus nephritis, which was more prevalent in the premenarche group versus those in the postmenarche group (52.2% versus 34.6%; $P = 0.001$). Among those with lupus nephritis, patients in the premenarche group had more proliferative lupus nephritis (class III or IV) than patients in postmenarche group (70.0% versus 58.6%; $P = 0.0009$). The girls with premenarchal-onset SLE were more likely to have anti-dsDNA antibodies than those with postmenarchal disease onset (81.7% versus 65.4%; $P = 0.001$). The presence of autoimmune cytopenias and other extractable nuclear antigens and antiphospholipid antibodies was not significantly different between the groups (Table 1).

The mean ± SD final height of all patients in this cohort was 160.6 ± 7.1 cm (Z score –0.38 ± 1.10). Patients in the postmenarche group attained a significantly greater final height than patients in the premenarche group, with mean ± SD final heights of 161.4 ± 6.9 cm and 158.8 ± 7.3 cm, respectively (Figure 1). The mean ± SD final height Z score in the postmenarche group was greater than the mean Z score in the premenarche group (–0.27 ± 1.07 versus –0.66 ± 1.13; $P = 0.001$). When we stratified by ethnicity, the final heights of patients in the premenarche and postmenarche groups were not significantly different (Figure 2).

Patients diagnosed with childhood-onset SLE postmenarche achieved a final height, on average, 2.6 cm taller than patients diagnosed premenarche in univariate, linear regression models and in multivariable models (2.6 cm [95% confidence interval (95% CI) 1.1, 4.1]; $P < 0.001$) (Table 2), accounting for ethnicity, the presence of nephritis, neuropsychiatric involvement, and anti-dsDNA positivity.

Subgroup analysis restricted to 99 patients with demonstrated growth following childhood-onset SLE diagnosis (35% of the postmenarche group), compared final height to patients diagnosed premenarche. In this subgroup analysis, there was no significant difference in final height between the premenarche and postmenarche groups (mean ± SD premenarche 158.8 ± 7.3 cm, postmenarche 160.5 ± 7.1 cm; $P = 0.09$). When we compared the premenarche group to the postmenarche group with no growth, the difference in final height was most marked (Table 3).

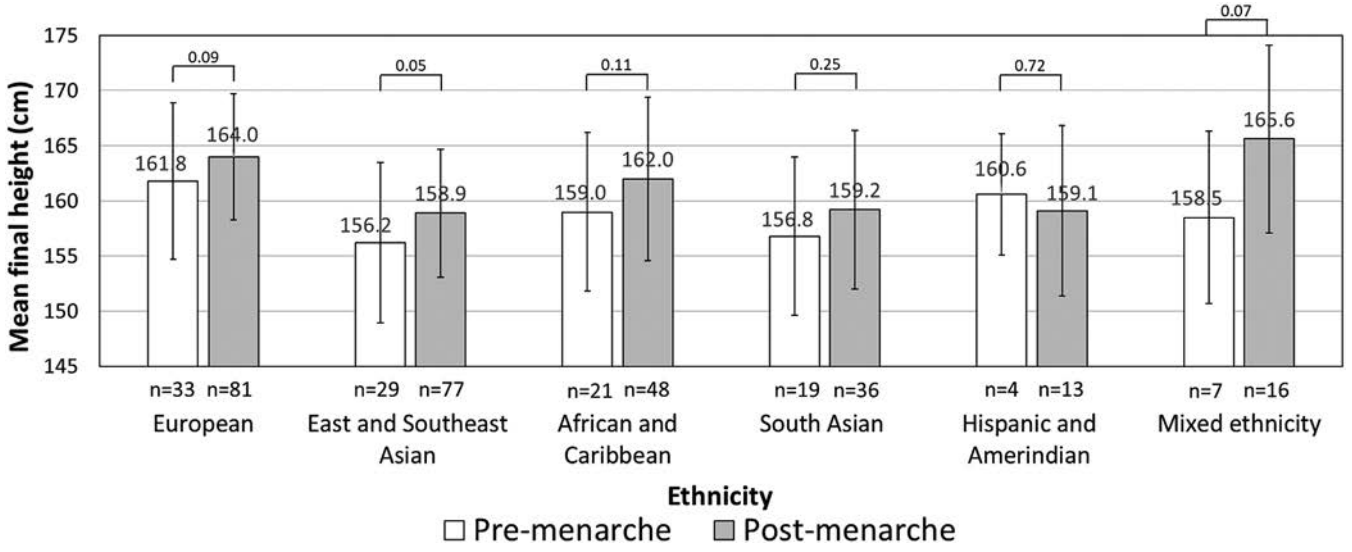

Figure 2. Means of final height of patients based on timing of childhood-onset systemic lupus erythematosus (SLE) diagnosis, categorized by ethnicity. Bar graph compares the mean final height (cm) of patients diagnosed with childhood-onset SLE premenarche and postmenarche, stratified by ethnic group. Whiskers represent the SD. *P* value was calculated from an unpaired *t*-test.

In subgroup analyses examining corticosteroid exposure and disease activity in those with demonstrated growth, the premenarche group had higher cumulative corticosteroid exposure compared to the postmenarche group (median 355 mg/kg [IQR 226, 636] versus 171 mg/kg [IQR 42, 279]; $P < 0.001$) and longer duration of corticosteroid therapy (median 3.9 years [IQR 2.2, 5.7] versus 1.7 years [IQR 0.3, 2.8]; $P < 0.001$). This finding corresponded to longer disease duration before completing growth in the premenarche group compared to the postmenarche group (Table 4). We did not observe a difference in average daily SLEDAI score between the premenarche and postmenarche groups (median 2.6 [IQR 1.3, 3.6] versus 2.8 [IQR 1.6, 3.9]; $P = 0.38$) (Table 4). Kaplan–Meier survival analysis demonstrated no significant difference in time to first sustaining disease damage (SLICC score ≥1) during follow-up between these 2 groups ($P = 0.4$) (see Supplementary Figure 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24461/abstract).

In subgroup analyses, those diagnosed postmenarche who had demonstrated growth were 1.7 cm taller on average than those diagnosed premenarche (95% CI –0.3, 3.6), but this difference did not achieve statistical significance ($P = 0.09$) (Table 4). There was also no difference in multivariable analyses adjusted for ethnicity, lupus nephritis, neuropsychiatric features, daily steroid dose, duration, and disease activity. The following factors were statistically associated with final height in univariate and multivariable analyses: ethnicity, total steroid dose, and duration of steroid therapy (Table 4). Average daily disease activity scores did not significantly differ between patients in the premenarche and postmenarche groups (median 2.6 [IQR 1.3, 3.6] versus 2.8

Table 2. Menarche status at childhood-onset SLE diagnosis and final height in female patients*

	Univariate analysis		Multivariable analysis	
	Final height, cm	*P*	Final height, cm	*P*
Postmenarche	2.6 (1.1, 4.1)	<0.001	2.6 (1.1, 4.1)	<0.001
Lupus nephritis	−1.4 (−2.8, 0.1)	0.06	−0.5 (−1.9, 1.0)	0.54
Neuropsychiatric lupus	−1.3 (−2.9, 0.3)	0.10	−1.5 (−3.0, 0.0)	0.047
Anti-dsDNA positivity	−1.0 (−2.5, 0.6)	0.21	0.2 (−1.3, 1.8)	0.76
Ethnicity†				
East and Southeast Asian	−5.2 (−7.0, −3.4)	<0.001	−5.3 (−7.2, −3.5)	<0.001
African and Caribbean	−2.3 (−4.3, −0.2)	0.03	−2.3 (−4.3, −0.2)	0.03
South Asian	−5.0 (−7.2, −2.8)	<0.001	−4.8 (−7.0, −2.6)	<0.001
Hispanic and Amerindian	−3.9 (−7.4, −0.5)	0.03	−4.2 (−7.6, −0.8)	0.02
Mixed ethnicity	0.1 (−3.0, 3.1)	0.96	−0.1 (−3.1, 3.0)	0.98

* Values are the number (95% confidence interval) unless indicated otherwise. Analyses included n = 401 girls: 115 diagnosed with systemic lupus erythematosus (SLE) premenarche, 99 diagnosed with SLE postmenarche. Anti-dsDNA = anti–double-stranded DNA.
† European referent.

Table 3. Comparison of 3 groups: patients with diagnosis at premenarche, at postmenarche with growth, and at postmenarche with no growth*

	Premenarche (n = 115)	Postmenarche, with growth (n = 99)	Postmenarche, no growth (n = 181)	P†
Final height, mean ± SD cm	158.8 ± 7.3	160.5 ± 7.1	161.9 ± 6.8	0.001
Height Z score, mean ± SD	–0.66 ± 1.13	–0.41 ± 1.10	–0.19 ± 1.04	0.002
Height percentile, mean ± SD	33 ± 30	39 ± 29	44 ± 29	0.005
Age at diagnosis, mean ± SD years	10.7 ± 2.2	13.9 ± 1.4	15.7 ± 1.3	<0.001
Age at menarche, mean ± SD years	13.5 ± 1.4	12.5 ± 1.3	12.5 ± 1.5	<0.001
Time from menarche to diagnosis, years	–2.3 (–3.7, –1.2)	1.2 (0.8, 1.9)	3.2 (2.2, 3.9)	<0.001
Disease duration before achieving final height, years	6.1 (4.8, 7.7)	2.9 (2.0, 3.8)	–	<0.001
Total steroid dose, prednisone equivalent, mg/kg	355 (226, 636)	171 (42, 279)	–	<0.001
Total steroid duration, years	3.9 (2.2, 5.7)	1.7 (0.3, 2.8)	–	<0.001
Average daily SLEDAI score	2.6 (1.3, 3.6)	2.8 (1.6, 3.9)	–	0.38

* Values are the median (interquartile range) unless indicated otherwise. SLEDAI = Systemic Lupus Erythematosus Disease Activity Index.
† *P* values by Kruskal-Wallis test.

[IQR 1.6, 3.9]; $P = 0.38$) (Table 3). In univariate and multivariable models, disease activity scores were not associated with final height.

Within the postmenarche group, girls with demonstrated growth and those without growth had a similar age of menarche (mean ± SD age 12.5 ± 1.4 years). However, girls with demonstrated growth were diagnosed at a younger age (mean ± SD age 13.9 ± 1.5 years versus 15.7 ± 1.3 years; $P < 0.001$), with a shorter time from menarche to childhood-onset SLE diagnosis (mean ± SD 1.2 ± 1.2 years versus 3.1 ± 1.8 years; $P < 0.001$) (Table 3).

We calculated predicted final height based on parent-reported height in 49 patients (19 in the premenarche group, 30 in the postmenarche group). We found that 10% of the girls had a final height lower than the predicted final height, based on midparental height, with no difference by menarche status at diagnosis (16% of premenarche versus 7% postmenarche; $P = 0.36$)

A total of 203 girls had Tanner stage documented within 1 year of childhood-onset SLE diagnosis (mean ± SD 22 ± 66 days), with 70% having Tanner stage documentation contemporaneous with SLE diagnosis. A total of 9% were determined to be at Tanner stage 1, 20% at Tanner 2, 8% at Tanner 3, 32% at Tanner 4, and 32% at Tanner 5. Sensitivity analyses adding Tanner staging to multivariable analyses demonstrated that girls assessed to be at Tanner stage 5 at diagnosis achieved an 8.8 cm greater final height than girls at Tanner 1 at diagnosis (SE 3.4 cm; $P = 0.009$). Girls at Tanner 4 were 7.6 cm taller (SE 3.2 cm; $P = 0.02$), and those at Tanner 3 were 5.0 cm taller (SE 2.4 cm; $P = 0.04$) than girls assessed to be at Tanner stage 1 at diagnosis. There was no significant difference in final height between girls at Tanner 2 versus Tanner stage 1 at diagnosis (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24461/abstract).

Table 4. Subgroup analyses of girls with demonstrated growth: menarchal status at childhood-onset SLE diagnosis and final height, univariate and multivariable linear regression*

	Univariate analysis		Multivariable analysis	
	Final height, cm	P	Final height, cm	P
Postmenarche with growth vs. premenarche	1.7 (–0.3, 3.6)	0.09	–0.6 (–2.7, 1.5)	0.57
Lupus nephritis	–1.6 (–3.6, 0.3)	0.10	1.5 (–0.8, 3.8)	0.20
Neuropsychiatric lupus	–2.9 (–5.3, –0.5)	0.02	–0.7 (–3.2, 1.8)	0.60
Anti-dsDNA positivity	–1.2 (–3.5, 1.2)	0.33	0.5 (–2.1, 3.2)	0.69
Ethnicity†				
East and Southeast Asian	–4.0 (–6.5, –1.4)	0.002	–4.0 (–6.5, –1.6)	0.001
African and Caribbean	–2.4 (–5.5, 0.6)	0.12	–1.7 (–4.7, 1.3)	0.26
South Asian	–4.3 (–7.4, –1.2)	0.007	–3.3 (–6.3, –0.3)	0.03
Hispanic and Amerindian	–5.3 (–10.6, –0.1)	0.048	–3.6 (–8.7, 1.5)	0.17
Mixed ethnicity	–3.0 (–7.3, 1.3)	0.17	–2.2 (–6.3, 1.9)	0.30
Cumulative steroid dose, prednisone equivalents of 10 mg/kg	–0.08 (–0.10, –0.05)	<0.001	–0.08 (–0.13, –0.03)	0.001
Total steroid duration, years	–0.8 (–1.2, –0.5)	<0.001	–0.2 (–0.8, 0.4)	0.52
Average daily SLEDAI score	–0.1 (–0.6, 0.5)	0.75	0.3 (–0.3, 0.9)	0.39

* Values are the number (95% confidence interval) unless indicated otherwise. Subgroup analyses included n = 214 girls: 115 diagnosed with systemic lupus erythematosus (SLE) premenarche, 99 diagnosed with SLE postmenarche. Anti-dsDNA = anti–double-stranded DNA; SLEDAI = Systemic Lupus Erythematosus Disease Activity Index.
† European referent.

DISCUSSION

In this large multiethnic cohort study, we found that the timing of childhood-onset SLE diagnosis as it relates to menarchal status had an impact on final adult height. Female patients diagnosed with childhood-onset SLE after menarche achieved a taller final height compared to those diagnosed before menarche, after we accounted for ethnicity, the presence of lupus nephritis, neuropsychiatric features, and anti-dsDNA antibodies. These findings were likely the result of a longer growth period vulnerable to impairment by disease and therapy among patients diagnosed premenarche compared to postmenarche.

Our findings are consistent with prior studies examining the impact of timing of childhood-onset SLE diagnosis on growth. The Paediatric Rheumatology International Trials Organisation longitudinal study examined 331 male and female patients with active childhood-onset SLE disease who required new or increased doses of corticosteroids and/or immunosuppressants (10). There was a significant reduction in the parent-adjusted height Z score after a 26-month follow-up. Examining the impact of age at disease onset on height demonstrated that females with age of onset less than 12 years had a marked reduction of height Z score with no catch-up growth over this 26 month study period, and the study did not include final adult height and pubertal status in individuals during the study period (10). A cross-sectional study of adults with childhood-onset SLE demonstrated that patients with childhood-onset SLE were, on average, 2.4 cm shorter than their target height. The 17 female patients diagnosed with childhood-onset SLE between age 11 and 13 years had the greatest risk for reduced final height, compared to patients with adult-onset disease; however, the authors of that study did not report on the relationship between age of diagnosis and menarchal status in individuals (11).

Comparison of SLE features in our cohort demonstrated a higher proportion of lupus nephritis in the premenarche group compared to the postmenarche group. Two studies of patients with childhood-onset SLE also demonstrated that lupus nephritis was more frequent in prepubertal patients or patients of younger age at onset, than postpubertal patients with later childhood-onset SLE (26,27). We found that the presence of anti-dsDNA antibodies was more common in patients diagnosed premenarche compared to those diagnosed postmenarche. This finding is in contrast to a prior retrospective study comparing SLE features across age groups that did not show any significant difference in anti-dsDNA antibodies between school-age children (7–11 years) and adolescent groups (12–18 years) (28). Although there was a higher proportion of patients in the premenarche group with lupus nephritis and anti-dsDNA antibodies, we did not find a significant impact of these 2 features on final height in univariate and multivariable regression analyses.

Our study found that patients diagnosed prior to menarche experienced menarche at an older age compared to those diagnosed postmenarche. This finding suggests that in the premenarche group, childhood-onset SLE may have delayed menarche but also provided an increased interval for growth to achieve final adult height. Despite this longer duration of growth, the group diagnosed prior to menarche remained shorter on average than those diagnosed after menarche.

To determine the relative impact of disease activity and corticosteroid exposure on final height, we completed a subgroup analysis on the population with demonstrated continued growth following childhood-onset SLE diagnosis. When we compared patients diagnosed before menarche to those with growth postmenarche (i.e., diagnosis of childhood-onset SLE closer to menarche), we no longer found a significant difference in final height. This finding was also true after accounting for disease activity, which was comparable between the groups, and total prednisone exposure and duration of steroid exposure, both higher in the premenarche group. However, similar final heights in this subgroup analysis meant that we could not parse out attribution for height difference to disease activity versus steroid exposure. The lack of difference in final height likely reflects a greater impact of childhood-onset SLE diagnosis on final height among girls diagnosed more recently postmenarche, compared to girls who completed growth well before childhood-onset SLE diagnosis. Our findings are similar to a previous longitudinal study of 25 patients with childhood-onset SLE followed over 2 years that found no association between the severity of disease activity at disease onset and development of growth failure (29).

Sensitivity analyses adding Tanner staging to our multivariable models highlighted the linear relationship between advanced sexual maturation rating (Tanner staging) at diagnosis, and taller final height. We observed the greatest difference in final height between girls with the most advanced sexual maturation and those who were least advanced. The difference in final height compared to Tanner 1 gradually decreased with decreasing Tanner stage at diagnosis. This finding further supported our observation that more advanced sexual maturation at childhood-onset SLE diagnosis protected patients' final height from the negative impacts of disease.

Our study had some limitations. We did not routinely screen for vertebral collapse in asymptomatic patients. Lateral spine radiographs were only completed in patients with evidence of a bone-mineral density lowest Z score of less than –2 on dual-energy x-ray absorptiometry, and/or when patients were symptomatic. Also, parental height was only available in a small proportion of patients, so we did not include midparental height as a standard comparison in this study.

Our study had a number of strengths. We prospectively collected data on a very large single-center, multiethnic childhood-onset SLE patient cohort, allowing for examination of the association of timing of childhood-onset SLE diagnosis as it relates to menarchal status on final height. We were able to account for important potential confounders of the association between the

timing of menarche and final height, including ethnicity, disease manifestations, steroid dose, disease activity scores, and disease damage.

In summary, female patients diagnosed with childhood-onset SLE postmenarche achieved a taller final height than those diagnosed with childhood-onset SLE premenarche, even after accounting for ethnicity, the presence of lupus nephritis, neuropsychiatric features, and anti-dsDNA antibodies. Our findings suggest that girls diagnosed with childhood-onset SLE prior to menarche require close monitoring of linear growth and pubertal development. Our findings also highlight the negative impact of disease on final height, despite delayed menarche and extended opportunity for growth. However, this difference of 2.6 cm (1 inch) in final height is arguably an acceptable and necessary cost of adequate control of severe disease. Future studies on impacts of the timing of childhood-onset SLE diagnosis and other disease-associated factors on height velocity, specifically aimed at defining the relative impact of disease activity and steroid exposure on growth and final height, are needed to define the period of maximal growth potential before final height is attained.

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. Hiraki 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. Sontichai, Silverman, Wasserman, Hiraki.

Acquisition of data. Sontichai, Dominguez, Levy, Ng, Silverio, Hiraki.

Analysis and interpretation of data. Sontichai, Liao, Al Mutairi, Wasserman, Hiraki.

REFERENCES

1. Jelenkovic A, Sund R, Hur YM, Yokoyama Y, Hjelmborg JV, Moller S, et al. Genetic and environmental influences on height from infancy to early adulthood: an individual-based pooled analysis of 45 twin cohorts. Sci Rep 2016;6:28496.
2. Dubois L, Ohm Kyvik K, Girard M, Tatone-Tokuda F, Perusse D, Hjelmborg J, et al. Genetic and environmental contributions to weight, height, and BMI from birth to 19 years of age: an international study of over 12,000 twin pairs. PLoS One 2012;7:e30153.
3. Perkins JM, Subramanian SV, Davey Smith G, Ozaltin E. Adult height, nutrition, and population health. Nutr Rev 2016;74:149–65.
4. Silventoinen K. Determinants of variation in adult body height. J Biosoc Sci 2003;35:263–85.
5. Simon D, Fernando C, Czernichow P, Prieur AM. Linear growth and final height in patients with systemic juvenile idiopathic arthritis treated with longterm glucocorticoids. J Rheumatol 2002;29:1296–300.
6. Kamphuis S, Silverman ED. Prevalence and burden of pediatric-onset systemic lupus erythematosus. Nat Rev Rheumatol 2010;6:538–46.
7. Aggarwal A, Srivastava P. Childhood onset systemic lupus erythematosus: how is it different from adult SLE? Int J Rheum Dis 2015;18:182–91.
8. Brunner HI, Gladman DD, Ibañez D, Urowitz MD, Silverman ED. Difference in disease features between childhood-onset and adult-onset systemic lupus erythematosus. Arthritis Rheum 2008;58:556–62.
9. Biro FM, Huang B, Crawford PB, Lucky AW, Striegel-Moore R, Barton BA, et al. Pubertal correlates in black and white girls. J Pediatr 2006;148:234–40.
10. Rygg M, Pistorio A, Ravelli A, Maghnie M, Di Iorgi N, Bader-Meunier B, et al. A longitudinal PRINTO study on growth and puberty in juvenile systemic lupus erythematosus. Ann Rheum Dis 2012;71:511–7.
11. Heshin-Bekenstein M, Perl L, Hersh AO, von Scheven E, Yelin E, Trupin L, et al. Final adult height of patients with childhood-onset systemic lupus erythematosus: a cross sectional analysis. Pediatr Rheumatol Online J 2018;16:30.
12. Hochberg MC, for the Diagnostic and Therapeutic Criteria Committee of the American College of Rheumatology. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus [letter]. Arthritis Rheum 1997;40:1725.
13. Petri M, Orbai AM, Alarcon GS, Gordon C, Merrill JT, Fortin PR, 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. Marshall WA, Tanner JM. Variations in pattern of pubertal changes in girls. Arch Dis Child 1969;44:291–303.
15. Rapkin AJ, Tsao JC, Turk N, Anderson M, Zeltzer LK. Relationships among self-rated tanner staging, hormones, and psychosocial factors in healthy female adolescents. J Pediatr Adolesc Gynecol 2006;19:181–7.
16. Rasmussen AR, Wohlfahrt-Veje C, Tefre de Renzy-Martin K, Hagen CP, Tinggaard J, Mouritsen A, et al. Validity of self-assessment of pubertal maturation. Pediatrics 2015;135:86–93.
17. Chavarro JE, Watkins DJ, Afeiche MC, Zhang Z, Sanchez BN, Cantonwine D, et al. Validity of self-assessed sexual maturation against physician assessments and hormone levels. J Pediatr 2017;186:172–8.
18. De Onis M, Onyango AW, Borghi E, Siyam A, Nishida C, Siekmann J. Development of a WHO growth reference for school-aged children and adolescents. Bull World Health Organ 2007;85:660–7.
19. Appel GB, Cohen DJ, Pirani CL, Meltzer JI, Estes D. Long-term follow-up of patients with lupus nephritis: a study based on the classification of the World Health Organization. Am J Med 1987;83:877–85.
20. Weening JJ, D'Agati VD, Schwartz MM, Seshan SV, Alpers CE, Appel GB, et al. The classification of glomerulonephritis in systemic lupus erythematosus revisited. J Am Soc Nephrol 2004;15:241–50.
21. Liu D, Ahmet A, Ward L, Krishnamoorthy P, Mandelcorn ED, Leigh R, et al. A practical guide to the monitoring and management of the complications of systemic corticosteroid therapy. Allergy Asthma Clin Immunol 2013;9:30.
22. Bombardier C, Gladman DD, Urowitz MB, Caron D, Chang CH, et al. Derivation of the SLEDAI: a disease activity index for lupus patients. The Committee on Prognosis Studies in SLE. Arthritis Rheum 1992;35:630–40.
23. Brunner HI, Silverman ED, Bombardier C, Feldman BM. European Consensus Lupus Activity Measurement is sensitive to change in disease activity in childhood-onset systemic lupus erythematosus. Arthritis Rheum 2003;49:335–41.
24. Gladman DD, Urowitz MB, Goldsmith CH, Fortin P, Ginzler E, et al. The reliability of the Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index in patients with systemic lupus erythematosus. Arthritis Rheum 1997;40:809–13.
25. Tanner JM, Goldstein H, Whitehouse RH. Standards for children's height at age 2 to 9 years allowing for height of parents. Arch Dis Child 1970;45:819.

26. Ambrose N, Morgan TA, Galloway J, Ionnoau Y, Beresford MW, Isenberg DA. Differences in disease phenotype and severity in SLE across age groups. Lupus 2016;25:1542–50.
27. Tarr T, Derfalvi B, Gyori N, Szanto A, Siminszky Z, Malik A, et al. Similarities and differences between pediatric and adult patients with systemic lupus erythematosus. Lupus 2015;24:796–803.
28. Zhu J, Wu F, Huang X. Age-related differences in the clinical characteristics of systemic lupus erythematosus in children. Rheumatol Int 2013;33:111–5.
29. Abdalla E, Jeyaseelan L, Ullah I, Abdwani R. Growth pattern in children with systemic lupus erythematosus. Oman Med J 2017;32: 284–90.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 2, February 2022, pp 208–218
DOI 10.1002/acr.24464

Prevalence, Predictors, and Prognostic Benefits of Remission Achievement in Patients With Systemic Lupus Erythematosus: A Systematic Review

Ziyi Yang,[1] Cheng Cheng,[1] Ziqian Wang,[1] Yanhong Wang,[2] Jiuliang Zhao,[1] Qian Wang,[1] Xinping Tian,[1] Evelyn Hsieh,[3] Mengtao Li,[1] and Xiaofeng Zeng[1]

Objective. To systematically review and evaluate the prevalence, potential predictors, and prognostic benefits of remission achievement in patients with systemic lupus erythematosus (SLE).

Methods. Studies reporting on the prevalence, predictors, and prognostic benefits of remission in adult patients with SLE were searched and selected from PubMed and Embase databases. Studies were reviewed for relevance and quality. Two reviewers independently assessed the studies and extracted data.

Results. Data from 41 studies including 17,270 patients were included and analyzed. Although no consensus has been achieved on the definition of remission, clinical disease activity, serologic activity, duration, and treatment are agreed to be critical components of defining remission status. In most studies published in the recent 5 years, 42.4–88% of patients achieved and maintained the remission status for 1 year, and 21.1–70% did so for at least 5 years. Factors associated with remission included older age at diagnosis, lower baseline disease activity, and absence of major organ involvement, while positive serologic results were shown to be negatively associated with remission. Remission (especially prolonged remission) when achieved, demonstrated an association with lower accrual of damage and better quality of life among patients with SLE.

Conclusion. Remission is an achievable and desirable target for SLE patients and proven to be associated with prognostic benefits. Further development and assessment of a clear remission definition, a risk stratification model, as well as a full algorithm with frequency of monitoring time points for treatment adjustment and drug withdrawal are required.

INTRODUCTION

Over the past 50 years, the therapeutic strategy for some of the most common chronic diseases has evolved from a symptom-based to a target-based approach, such as the management of hypertension, diabetes mellitus, and cardiovascular disease. The term "treat-to-target" in diseases requiring long-term management refers to initiating treatment steps to achieve a goal that is verified to significantly benefit the disease prognosis. As for the rheumatic diseases, the treat-to-target approach was first recommended in the management of rheumatoid arthritis (RA) (1). Recently, the concept of treat-to-target has been introduced in the spondyloarthritides as well.

Systemic lupus erythematosus (SLE) is a chronic autoimmune disease affecting multiple organ systems characterized by a fluctuating disease course. This relapsing–remitting pattern makes the treatment of SLE typically long term or even lifelong. This has led to the SLE treat-to-target recommendations published in 2014 (2). In 2016, an international expert panel (the Definition of Remission in SLE [DORIS] project) stated that remission

Supported by the Chinese National Key Technology Research and Development Program, Ministry of Science and Technology (grants 2021YFC2501301-6 and 2017YFC0907601-5), the Beijing Municipal Science and Technology Commission (grant Z201100005520022,23, 25-27), and the Chinese Academy of Medical Sciences (Innovation Fund for Medical Sciences grant 2021-I2M-1-005). Dr. Hsieh's work was supported by the NIH (Fogarty International Center grant K01TW009995).

[1]Ziyi Yang, MD, Cheng Cheng, MD, Ziqian Wang, MD, Jiuliang Zhao, MD, Qian Wang, MD, Xinping Tian, MD, Mengtao Li, MD, Xiaofeng Zeng, MD: Peking Union Medical College Hospital, Peking Union Medical College and Chinese Academy of Medical Sciences, National Clinical Research Center for Dermatologic and Immunologic Diseases, Ministry of Science and Technology, Key Laboratory of Rheumatology and Clinical Immunology, Ministry of Education, Beijing, China; [2]Yanhong Wang, PhD: Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China; [3]Evelyn Hsieh, PhD, MPH, MD: Yale University School of Medicine, New Haven, Connecticut.

Drs. Yang and Cheng contributed equally to this work.

No potential conflicts of interest relevant to this article were reported.

Address correspondence to Mengtao Li, MD, or Xiaofeng Zeng, MD, Peking Union Medical College Hospital, No. 1 Shuaifuyuan, Beijing 100730, China. Email: mengtao.li@cstar.org.cn or zengxfpumc@163.com.

Submitted for publication May 12, 2020; accepted in revised form September 22, 2020.

SIGNIFICANCE & INNOVATIONS

- Remission is an achievable and desirable therapeutic target in treat-to-target approaches in treating patients with systemic lupus erythematosus (SLE) and is associated with benefits in decreasing accrual of damage and increasing quality of life.
- Even though SLE remission has been brought into focus over the last decade, the definition of remission remains elusive. This review compares the prevalence, potential predictors, and prognostic benefits of remission in the setting of various definitions.
- This review highlights the need for well-designed studies assessing all definitions to identify the one with superior prognostic value and for developing a full algorithm with a detailed follow-up and drug adjustment regimen.

could be considered as a treatment target for patients with SLE. Although remission is not the same as cure, it is a state that, if sustained, is associated with a lower likelihood of adverse outcomes (3).

Even though numerous studies focusing on SLE remission have been published over the last decade, the definition of remission remains elusive. The prevalence of remission achievement varies among cohorts and regions. Similar remission-associated factors and prognostic benefits of remission have been found by studies, but further confirmation is still required by strictly designed clinical trials. In this review, the definitions adopted by the SLE remission–related studies from the 1980s to the present, as well as the prevalence, possible predictors, and prognostic benefits of remission of disease, are summarized, and we aim to evaluate the existing evidence and identify the outstanding knowledge gaps.

MATERIALS AND METHODS

Search strategy. This systematic review was written in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (4). Two electronic databases, PubMed and Embase, were used in the search of published literature from inception to February 18, 2020, independently by 2 investigators (ZY and CC). The search was restricted to English-language publications. Original articles were included. Reference lists from original articles and reviews were manually scanned to identify any other eligible studies. The search keywords in PubMed were the following: "lupus erythematosus, systemic," "remission," "treat to target," "clinically quiescent," "prevalence," "frequency," "predictor(s)," "outcome(s)," "prognosis," "quality of life," "damage," and "disease progression." The search keywords in Embase were the following: "lupus erythematosus, systemic," "remission," "treat to target," "clinically quiescent," "prevalence," "frequency," "predictive validity," "predictor(s)," "outcome(s)," "prognosis," and "disease exacerbation." The exact search strategies are provided in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24464/abstract.

Eligibility criteria and study selection. ZY and CC independently screened the titles and abstracts of references, identified articles for full-text review using the inclusion criteria, and assessed the methodologic quality of the studies. Any discrepancies were resolved through consensus. EH and ML participated in resolving disagreements.

Observational studies (case–control or cohort studies) and clinical trials were included, whereas case reports, literature reviews, and editorials were excluded. We considered publications involving adults with SLE, addressing the prevalence, predictors, and effects on disease outcomes of remission in SLE patients. Studies were excluded if they focused on pediatric patients or other autoimmune diseases. When multiple reports were published on the same study, the report with more complete information was extracted.

Data extraction. The 2 investigators independently extracted data from each study using a systematic data extraction form (available on request) developed for this specific purpose, including sample size, ethnicity, remission types, and definition, prevalence, predictors, and prognostic outcome of each remission category. After extracting data independently from each study, discrepancies were resolved through consensus.

RESULTS

Literature search. The literature search resulted in a total of 2,992 articles after duplicates were removed. A total of 2,899 articles were excluded after title and abstract screening, and the full-text manuscripts of the 93 remaining articles were reviewed. Ultimately, 41 studies were selected for inclusion in the systematic review. Figure 1 illustrates the PRISMA flow diagram detailing the process of abstract screening, article selection, and inclusion.

Study characteristics. The main characteristics of the 41 selected studies including 17,270 patients are summarized in Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24464/abstract. For the included studies, the geographic distribution is as follows: 17 from Europe, 15 from North America, 5 from South America, and 4 from Asia. Most cohorts were single centered, and no randomized controlled trials were identified. Included studies were published between 1979 and 2019, and 29 of the 41 included studies had been carried out in the past 5 years.

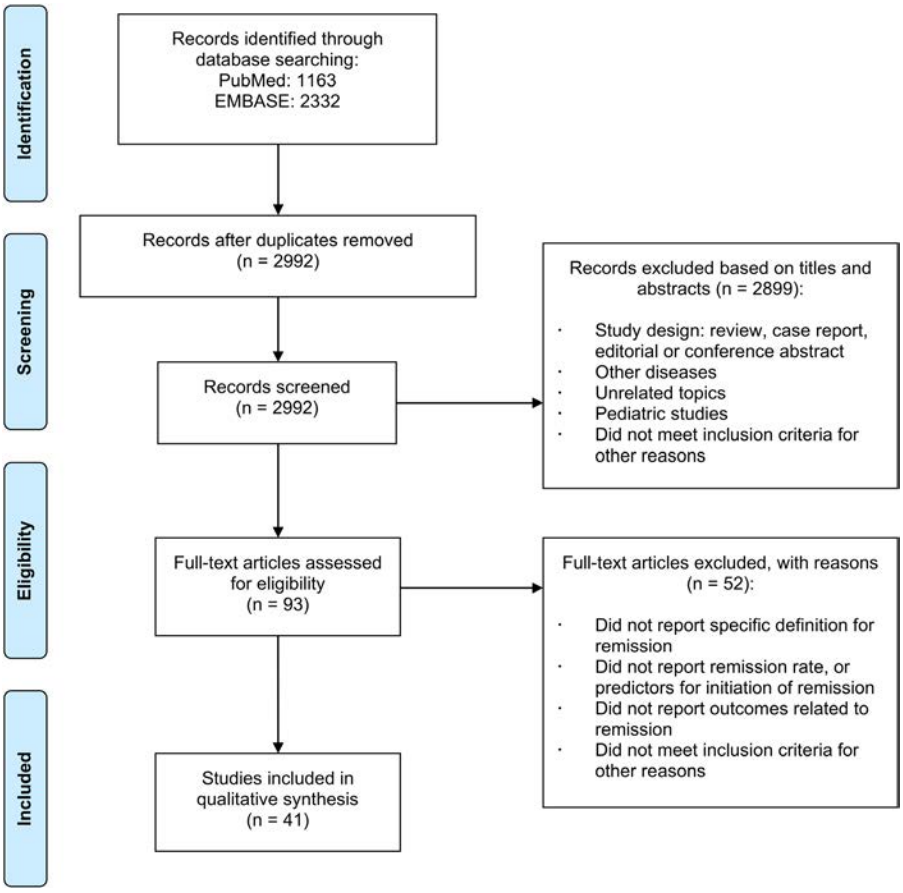

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram detailing the process of abstract screening, article selection, and inclusion.

Definition of remission. The concept of remission was first described in 1979 as an absence of clinical manifestations of disease (5–7) or not currently experiencing a symptomatic flare (8). Subsequently developed disease activity assessment indices (Systemic Lupus Erythematosus Disease Activity Index 2000 [SLEDAI-2K] [9–11], British Isles Lupus Assessment Group [BILAG] [12], Systemic Lupus Activity Measure [SLAM] [13,14], European Consensus Lupus Activity Measure [ECLAM] [15], and physician global assessment of disease activity [PhGA] [16]) were adopted by research studies when defining remission. In 2015, Zen et al (11) proposed the concepts of complete remission, clinical remission without steroids, and clinical remission with low-dose steroids. In 2016, the DORIS project established guiding principles for the definitions of remission and agreed on 4 domains critical to defining remission in SLE: clinical disease activity, serologic activity, state duration, and treatment (3). The majority of subsequent studies have adopted this definition format. The classification and definition of remission tends to be unified, with some differences existing in the assessment index, treatment regimen, and remission duration. Evolution of this concept is summarized in Table 1, with additional details shown in Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24464/abstract. What should be noticed is that before remission was proposed, some studies (17–23) used the concept of serologically active clinically quiescent (SACQ) disease, long quiescence (LQ), and clinically quiescent disease (CQD) to describe disease status. A clinical SLEDAI-2K score of 0 with positive anti–double-stranded DNA (anti-dsDNA) antibody titers and/or hypocomplementemia were widely adopted as a definition in these studies, which was similar to the clinical remission definition. The detailed characteristics of these studies are shown in Supplementary Tables 2 and 3, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24464/abstract.

Prevalence of remission. Remission prevalence varied among studies, as summarized in Table 2. Before the definition format was proposed by the DORIS project in 2016, the prevalence of remission ranged from 2.5% to 37.4% (5,7,9–12,24,25). After DORIS established guiding principles for the definitions of remission, the prevalence of remission was further subdivided into subgroups. Among the studies evaluating all 4 types or 3 types of remission, prevalence was highest based on clinical remission on treatment, ranging from 15.6% to

Table 1. Definitions of remission across included studies*

Author, year (ref.) and classification of remission	Clinical activity	SA	Duration, years	Treatment permissible		
				Low-dose GC	IS, BI	AM
Drenkard et al 1996 (25)						
ND	Lack of clinical disease activity	Yes	≥1	No	No	No
Formiga et al, 1999 (24)						
ND	Lack of clinical disease activity	Yes	≥1	No	No	No
Steiman et al, 2014 (10)						
ND	SLEDAI score = 0	Yes	≥5	No	No	Yes
Zen et al, 2015 (11)						
Prolonged complete remission	SLEDAI-2K score = 0	No	≥5	No	No	Yes
Prolonged clinical remission without treatment	Clinical SLEDAI-2K score = 0	Yes	≥5	No	Yes	Yes
Prolonged clinical remission with treatment	Clinical SLEDAI-2K score = 0	Yes	≥5	Yes	Yes	Yes
Medina-Quinoes et al, 2016 (12)						
Complete remission	BILAG score C, D, or E	No	≥3	No	No	Yes
Clinical remission	BILAG score C, D, or E	Yes	≥3	No	No	Yes
Serological remission	BILAG score A or B	No	≥3	Yes	Yes	Yes
Mok et al, 2017 (16)						
Complete remission without treatment	SLEDAI score = 0; PhGA score <0.5	No	ND	No	No	Yes
Complete remission with treatment	SLEDAI score = 0; PhGA score <0.5	No	ND	Yes	Yes	Yes
Clinical remission without treatment	Clinical SLEDAI score = 0; PhGA score <0.5	Yes	ND	No	No	Yes
Clinical remission with treatment	Clinical SLEDAI score = 0; PhGA score <0.5	Yes	ND	Yes	Yes	Yes
Romo-Tena et al, 2018 (27)						
Complete remission	SLEDAI-2K score = 0	No	≥1	No	No	Yes
Clinical remission without treatment	Clinical SLEDAI-2K score = 0	Yes	≥1	No	Yes	Yes
Clinical remission with treatment	Clinical SLEDAI-2K score = 0	Yes	≥1	Yes	Yes	Yes

* The term 'clinical' for the Systemic Lupus Erythematosus Disease Activity Index (SLEDAI) refers to symptoms, signs, and routine laboratory testing, disregarding only the points that can be given for the presence of anti-DNA antibodies and/or low complement. The term 'serological activity' (SA) in systemic lupus erythematosus generally refers to the presence of anti-DNA antibodies and/or hypocomplementemia. Low-dose glucocorticoids (GC) refers to no more than 5 mg/day of prednisone (or equivalent) (ref. 3). AM = antimalarial; BI = biological immunomodulator; BILAG = British Isles Lupus Assessment Group; IS = immunosuppressant; ND = not described; PhGA = physician global assessment of disease activity; SLEDAI-2K = SLEDAI 2000.

49.6% (11,16,26–29). Remission type with the lowest prevalence was clinical remission off treatment in 4 studies (16,26,27,30) and complete remission off treatment in 4 studies (11,28,29,31).

Before the concept of remission was widely accepted, the prevalence of SACQ disease was analyzed in some studies, ranging from 6.1% to 9% (6,21,22). Similar criteria included LQ and CQDs, with prevalence ranging from 16% to 34% (19,23). With remission criteria developing, the SLEDAI became the most frequently used index in remission definition, with prevalence mostly ranging from 42.4% to 88.1% (for ≥1 year) (27,28,31–33) and 28% to 70% (for ≥5 years) (11,26,28,31,32). In studies employing both the SLEDAI and PhGA in remission definition, prevalence ranged from 45.5% to 90.4% (for a single visit) (16,29,30,34) and 21.1% to 35.5% (for ≥5 years) (16,29,30). Those using the BILAG, ECLAM, or SLAM (12,13,15,35) were ~35%.

Currently, the duration used to define a prolonged remission is not consistent across studies, ranging from 1 to 10 years. Generally, the longer the duration, the lower the prevalence. In studies published after 2015, the prevalence of prolonged remission for 1 year ranged from 42.4% to 88% in 6 studies (16,27,28,30,32,33), and that for 5 years ranged from 21.1% to 70% in 9 studies (11,16,26,28–32,36), while only 10.1% of patients achieved remission for 10 years in the studies by Tselios et al (18,37,38).

Predictors of presence of remission. The majority of the studies showed that multiple variables had a significant association with remission, as shown in Table 3. Although factors varied across studies, consistencies existed. In terms of demographic characteristics, 3 studies agreed that patients with older age at disease onset or diagnosis were more likely to achieve remission (10,12,38). Two studies examined the relationship between ethnicity and remission and reported that African American ethnicity (39) or non-White ethnicity (12) were negatively associated with remission. For baseline disease activity and organ involvement, 6 studies showed that higher disease activity prior to remission was negatively associated with remission (9,10,14,24,27,31). Five studies evaluated the association between organ/system involvement and remission (11,14,15,31,39), in which nephritis, hematologic activity, mucocutaneous and neurologic involvement, arthritis, and vasculitis were negatively related to remission. Analysis of the association between laboratory test results with remission revealed that hypocomplementemia, positive baseline anti-dsDNA, and anti-SSB (Ro) were negatively associated with remission (9,15,39).

Table 2. Prevalence of remission (studies with ≥70 patients)*

Author, year (ref.)	No.	Ethnicity	Duration	Remission	Clinical remission with treatment	Clinical remission without treatment	Complete remission with treatment	Complete remission without treatment
Gladman et al, 1979 (6)	180	ND	ND	SACQ: 7.8%	ND	7.8%	ND	ND
Tozman et al, 1982 (5)	160	ND	ND	2.5%	ND	ND	ND	2.5%
Heller and Schur, 1985 (7)	305	ND	ND	4.3%	ND	ND	4.3%	ND
Drenkard et al, 1996 (25)	667	ND	≥1 year	23.4%	ND	ND	23.4%	ND
Formiga et al, 1999 (24)	100	ND	≥1 year	24%	ND	ND	24%	ND
Barr et al, 1999 (23)	204	African American 58%; White 42%	≥1 year	LQ: PhGA criteria 16%; M-SLEDAI criteria 25%	ND	ND	ND	ND
Urowitz et al, 2005 (9)	703	ND	≥1 year; ≥5 years	6.5% (≥1 year); 1.7% (≥5 years)	ND	6.5% (≥1 year); 1.7% (≥5 years)	ND	ND
Ng et al, 2006 (22)	290	White, African Caribbean, other	At least 2 occasions for 6 months	SACQ: 9%	ND	ND	ND	ND
Steiman et al, 2010 (21)	924	White 73.2%; Black 9.2%; Chinese 9.4%; other 6.8%	≥2 years	SACQ: 6.1%	ND	ND	ND	ND
Zen et al, 2012 (19)	165	White	≥3 years	Persistent CQD: 34%	ND	ND	ND	ND
Kasitanon et al, 2014 (17)	95	ND	ND; ≥5 years	Clinical quiescence: 69.5% (ND); prolonged complete remission: 2.1% (≥5 years)	ND	ND	ND	ND; 2.1% (≥5 years)
Steiman et al, 2014 (10)	1,613	White, African American, Asian	Prolonged: ≥5 years	4.5%	2.1%	2.4%	ND	ND
Zen et al, 2015 (11)	224	White	Prolonged: 5 years	37.4%	15.6%	14.7%	ND	7.1%
Medina-Quinoes et al, 2016 (12)†	532	White, African Caribbean, Indian, Pakistani, Bangladeshi, Asian	≥3 years	35.4%	ND	8.5%	ND	14.5%

(Continued)

Table 2. (*Cont'd*)

Author, year (ref.)	No.	Ethnicity	Duration	Remission	Clinical remission with treatment	Clinical remission without treatment	Complete remission with treatment	Complete remission without treatment
Zen et al, 2017 (28)	293	ND	≥1 year; ≥5 years	88.1% (≥1 year); 38.6% (≥5 years)	33.0% (≥1 year); 12.3% (≥5 years)	32.1% (≥1 year); 17.1% (≥5 years)	ND	11.9% (≥1 year); 9.2% (≥5 years)
Ugarte-Gil et al, 2017 (41)	1,350	White, African American	ND	20.5%	ND	ND	16.8%	3.7%
Tsang-A-Sjoe et al, 2017 (36)	117	White	Prolonged: 5 years	37.6%	ND	ND	ND	6.0%
Mok et al, 2017 (16)	769; 751; 613‡	Chinese	At last visit (n = 769); ≥1 year (n = 751); ≥5 years (n = 613)	70.1% (n = 769); 60.1% (n = 751); 31.5% (n = 613)	28.0% (n = 769); 21.6% (n = 751); 7.0% (n = 613)	5.7% (n = 769); 5.3% (n = 751); 3.4% (n = 613)	22.2% (n = 769); 19.8% (n = 751); 10.1% (n = 613)	14.2% (n = 769); 13.3% (n = 751); 10.9% (n = 613)
Tani C et al, 2018 (29)	115	White	Baseline; 5 years	90.4% (baseline); 35.5% (5 years)	49.6% (baseline); 21.7% (5 years)	10.4% (baseline); 5.2% (5 years)	28.7% (baseline); 7.8% (5 years)	1.7% (baseline); 0.8% (5 years)
Sebastiani et al, 2018 (15)	185	White 92.4%	ND	Clinical remission 35.3%	ND	ND	ND	ND
Romo-Tena et al, 2018 (27)	124	Mexican	1 year	52.4%	20.2%	12.9%	ND	19.4%
Petri et al, 2018 (43)	1,356	White, African American	ND	40% (of 77,105 patient months)	27%	13%	ND	ND
Tselios et al, 2018 (18)	267	White, Black, Chinese, other	Prolonged: ≥10 years	10.1%	10.1%	ND	ND	ND
Mathian et al, 2019 (30)	407	ND	ND; ≥1 year; >5 years	62.4% (ND); 45.9% (≥1 year); 21.1% (>5 years)	15.2% (ND); ND (≥1 year); ND (>5 years)	11.5% (ND); ND (≥1 year); ND (>5 years)	14.5% (ND); ND (≥1 year); ND (>5 years)	21.1% (ND); ND (≥1 year); ND (>5 years)
Tsang et al, 2019 (34)	154	White 69.5%	Baseline	45.5%	18.2% remission with therapy	27.3% remission without therapy	ND	ND
Poomsalood et al, 2019 (33)	237	ND	Prolonged: ≥1 year	42.4%	22.9%	19.5%	ND	ND
Fasano et al, 2019 (26)	294	ND	Prolonged: 5 years	44.5%	19.7%	10.2%	ND	14.6%
Ugarte-Gil et al, 2019 (14)	902	White, Mestizo, African Latin American, other	ND	21.7%	ND	ND	21.7%	ND
Ruiz-Irastorza et al, 2019 (31)§	CC: 92; BC: 81	94% White (CC); 89% White (BC)	Year 1 (CC); prolonged: 5 years (CC); year 1 (BC); prolonged: 5 years (BC)	84% (year 1; CC); 70% (prolonged: 5 years; CC); 43% (year 1; BC); 28% (prolonged: 5 years; BC)	84% (year 1; CC); 70% (prolonged: 5 years; CC); 43% (year 1; BC); 28% (prolonged: 5 years; BC)	39% (year 1; CC); 25% (prolonged: 5 years; CC); 25% (year 1; BC); 16% (prolonged: 5 years; BC)	35% (year 1; CC); 21% (prolonged: 5 years; CC); 15% (year 1; BC); 10% (prolonged: 5 years; BC)	16% (year 1; CC); 9% (prolonged: 5 years; CC); 9% (year 1; BC); 5% (prolonged: 5 years; BC)
Tselios et al, 2019 (37,38)	267	White, Black, Chinese, other	Prolonged: ≥10 years	10.1%	10.1%	ND	ND	ND

(Continued)

Table 2. (*Cont'd*)

Author, year (ref.)	No.	Ethnicity	Duration	Remission	Clinical remission with treatment	Clinical remission without treatment	Complete remission with treatment	Complete remission without treatment
Margiotta et al, 2019 (32)	136	ND	≥1 year; prolonged: ≥5 years	82% (≥1 year); 39% (prolonged: ≥5 years)	82% (≥1 year); 39% (prolonged: ≥5 years)	ND (≥1 year); ND (prolonged: ≥5 years)	ND (≥1 year); ND (prolonged: ≥5 years)	ND (≥1 year); ND (prolonged: ≥5 years)
Alarcon et al, 2019 (13)	558	White, African descent, Hispanic	ND	1.83% (of 3,879 visits)	ND	ND	ND	ND
Ugarte-Gil et al, 2019 (35)	483	White 28%; African 37%; Hispanic 35%	ND	1.1% (of 2,004 patient/years)	ND	ND	ND	ND

* BC = Bordeaux Lupus Cohort; CC = Cruces Lupus Cohort; CQD = clinically quiescent disease; LQ = long quiescence; M-SLEDAI = modified Systemic Lupus Erythematosus Disease Activity Index; ND = not described; PhGA = physician global assessment of disease activity; SACQ = serologically active clinically quiescent.
† In this study, 12.4% of patients were in serologic remission, which has been contained in the total remission rate.
‡ N = 751 and n = 613 indicate the number of patients in the cohort with the same disease duration.
§ In this paper, clinical remission with treatment was the least stringent type of remission; all patients fulfilling criteria for the other 3 types of remission were also classified as being in clinical remission with treatment. The nomenclature of remission status in some studies (11,12,27,31,36,41) may be different from what we adopt in this table, but the definition content is the same.

Table 3. Factors associated with systemic lupus erythematosus remission in included studies*

Author, year (ref.)	Positively associated factors	Negatively associated factors
Ruiz-Irastorza et al, 2019 (31)	NR	Baseline SLEDAI score and neurologic involvement, baseline nephritis
Ugarte-Gil et al, 2019 (14)	Absence of mucocutaneous, renal, and hematologic involvement; use of immunosuppressive drugs before the baseline visit; lower SLEDAI score at cohort entry	NR
Sebastiani et al, 2018 (15)	NR	Oral ulcers, arthritis, low C4, anti-SSB (Ro) antibodies, therapy with mycophenolate
Tselios et al, 2019 (38)	Higher mean prednisone dose at enrollment; older age at disease onset	NR
Novelli et al, 2019 (49)	NR	CD44v3 on CD8+ T cells and CD44v6 on CD4+ and CD8+ T cells
Wilhelm et al, 2017 (39)	NR	African American ethnicity, baseline anti–double-stranded DNA, low C3, low C4 and hematologic activity
Romo-Tena et al, 2018 (27)	NR	SLEDAI-2K score at the third month of follow-up and total number of disease flares
Urowitz et al, 2005 (9)	Lower overall disease activity (SLEDAI score, adjusted mean), lower prevalence of anti-DNA antibodies, lower use of steroids and antimalarials	NR
Steiman et al, 2014 (10)	NR	Younger at diagnosis, more disease activity prior to remission
Medina-Quinoes et al, 2016 (12)	Diagnosed at older age, longer disease duration	Non-White patients
Formiga et al, 1999 (24)	NR	Higher SLEDAI score (longer to achieve remission)
Zen et al, 2015 (11)	NR	Vasculitis, glomerulonephritis, and hematologic manifestations over disease course

* NR = not reported; SLEDAI = Systemic Lupus Erythematosus Disease Activity Index; SLEDAI-2K = SLEDAI 2000.

Prognostic benefits of remission. It should be noted that most studies have demonstrated that achieving remission states is associated with less damage accrual, reduction in risk of developing flares, easier glucocorticoid withdrawal, better quality of life, and reduced mortality (Table 4). The Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index was used in some studies to assess organ damage (40). Some studies (16,29) emphasized the effect of the percentage of time spent in remission on damage accrual. The duration varied from 1 to 5 years, and the percentage of remission in all visits ranged from very small to at least 50%. A continuous period of remission was more beneficial than the sum of short periods of remission interspersed with flares (28). Optimizing the health-related quality of life (HRQoL) of patients was assessed by the Medical Outcomes Study Short Form 36 (SF-36) and the Lupus Patient-Reported Outcome questionnaires in some studies, which found that the percentage of time that lupus patients stay in remission was associated with a better quality of life (35). However, a few studies revealed that no significant difference was found in accrual of damage (27) or mortality (41) between remitted patients and the unremitted ones.

DISCUSSION

Remission has been selected as a target for SLE due to its attainability, clinical value, and feasibility. The prevalence of remission in various definitions was generally higher in studies published in the recent 5 years when compared with older studies, as shown in Table 2. This may be attributed to a better understanding of SLE and advances in therapies in the recent decade. Additionally, increased awareness of remission as a potential target is probably another reason.

In recent studies, more than one-half of patients were able to achieve at least 1 remission state during their follow-up visits (16, 27–32), which makes remission an achievable goal. Moreover, its clinical value of predicting less accrual of organ damage, reduced risk of death, lower chance of flare, and better life quality has been demonstrated in recent studies. Remission is a useful outcome measure in clinical trials, and failure to achieve remission serves as an indication for treatment modification or intensification. Additionally, the indices and laboratory results used to determine remission are commonly part of routine clinical order sets, thus enhancing its feasibility. Although the concept of Lupus Low Disease Activity State (LLDAS) has been regarded as a less stringent, alternative target for remission (42), a longer period of time in LLDAS was required to achieve a comparable reduction in organ damage (43). Remission is still a more valid state with lower risks of damage, even when only a small percentage of time (<25%) was maintained (43).

The definition of remission tended to be unified into 4 categories (see Supplementary Table 4, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/

Table 4. Prognostic outcomes related to systemic lupus erythematosus remission in included studies

Outcome related to remission	Reference
Lower risk of damage accrual	11, 16, 36, 41, 13, 43, 20, and 37
Lower risk of cardiovascular event	37 and 26
Lower risk of flares	37 and 50
Reduced risk of death	25
Better quality of life	32, 16, 33, 34, 51, 35, 37, and 52
Complete remission before glucocorticoid reduction was associated with increased likelihood of successful glucocorticoid withdrawal	53
No significant difference in damage accrual and disease pattern	27

acr.24464/abstract) after the DORIS project (3), which is still the mainstream construction recommended. Yet, controversies exist regarding the following: 1) the index used for defining disease activity; 2) whether serologic activity and low-dose glucocorticoids usage are allowed; and 3) the appropriate duration required to define remission.

The SLEDAI/SLEDAI-2K was the index most commonly applied in the definition of remission, and PhGA was employed to compensate for deficits in the SLEDAI (where hemolytic anemia, myelitis, and gastrointestinal activities are not considered) (3). Disagreement on whether patients' perspectives should be considered in the definition of remission exists, and no study has employed this directly in defining remission despite the recommendation for shared patient–physician decision-making in determining treatment plans (42). Therefore, the SLEDAI/SLEDAI-2K is still the major scale used for the measurement of disease activity in remission.

Due to the heterogeneity of SLE clinically, the prognostic capabilities of different definitions of remission vary between studies. Most studies allowing serologic activity and therapy with low-dose glucocorticoids have demonstrated an association between remission with lower damage accrual and better life quality, while some studies stated that patients in remission without therapy obtained even less damage accrual (11) and higher HRQoL (34) than patients in remission with therapy. When serologic activity and low-dose glucocorticoids were permitted, prevalence was higher (39). Recently, an analysis using clinical remission as response measurement to evaluate the belimumab treatment has been published. In their cohort, 39.6% of patients attained clinical remission without glucocorticoids, and 23.2% achieved clinical remission with glucocorticoids. The authors concluded that belimumab might be more efficacious in inducing low disease activity and clinical remission in SLE patients with limited or no organ damage accrued prior to treatment initiation. Moreover, that study has proven that remission can be used as a measure for the efficacy of treatment (44). Hence, it is proposed that SLE remission should not be regarded as a single ultimate goal, but as consecutive steps, with clinical remission with treatment serving as a first-step goal, after which appropriate drug tapering occurs, with the goal of eventual remission without treatment (45).

The frequency of follow-up visits in the studies considered varied, ranging from 3 months to 1 year. Whether a treat-to-target algorithm with a standardized follow-up interval and treatment adjustment suggestion could result in a superior prognosis compared to the conventional follow-up regimens, which are based on physician experience, needs to be further investigated. In RA, frequent monitoring (every 1–3 months) has been recommended by the European Alliance of Associations for Rheumatology for patients with active disease (46). We suggest that, in SLE, the interval between visits should be determined according to baseline disease activity, stratification of risk factors for nonremission, and the level of remission achieved. A shorter interval (2–3 months) should be considered for patients with high disease activity or those with negative predictors for remission, while a longer interval (4–5 months) may be suitable for patients with lower disease activity and lower risks for nonremission. Patients whose illness is in clinical remission with treatment may require more frequent visits than those already reaching complete remission without treatment.

To date, few studies have investigated the predictors for remission using multivariable analysis, while there is relative consensus in several studies regarding factors negatively associated with remission, such as higher baseline disease activity, major organ/system involvement, and active serologic results. Patients with negative predictors were less likely to achieve remission. Thus, the development of a risk stratification model and a test of its validity are essential in future studies, to enable early identification of patients less likely to achieve remission and individualize therapeutic regimens to attain better outcomes. In RA, higher prevalence of remission is achieved more often when intensive therapy is initiated earlier (47). Similar results have been shown in SLE, where immunosuppressant use or higher mean dose of prednisone use during the early phase of the disease are positively associated with remission. Controversies exist (9,15), however, and to date, only a few clinical trials have evaluated the impact of treatment strategy on rates of remission. Recently, research comparing different remission definitions was conducted by Saccon et al (48). They indicated that a clinical SLEDAI score of 0 is an easy-to-achieve definition with the best performance in predicting damage progression compared to other definitions, including PhGA and low-dose prednisone. More studies involving other ethnicities, outcome variables, and defining items are still required for a universal definition (48). More research is needed to assess the effects of treatment on achievement of remission and the appropriate adjustment of therapy if remission is not achieved within a certain time window.

Once remission with therapy is achieved, the optimal duration of remission before drug tapering or withdrawal needs to be discussed. Generally, an increasing duration in remission/LDAS was associated with a better SF-36 score or less damage accrual

(13,35). A consecutive remission duration of 5 years was shown in most studies to correlate with better outcomes. A minimum of 2 consecutive years of remission was shown to be necessary for White patients in another study (28). Whether remission for 1 year is sufficient remains uncertain (27,33). What needs to be recognized are the dual effects of glucocorticoids in terms of both controlling disease activity and causing organ damage. Therefore, both extending the duration of remission and minimizing glucocorticoid dosages should be balanced in the long-term disease control process. Additionally, feasibility should be considered because a decreasing prevalence of remission was observed when longer time was required. This study proposes that a minimum of 1 year is a prerequisite in the definition of remission, and the time point after which the damage caused by glucocorticoids to maintain remission status exceeds the benefits of prolonged remission itself should be clarified further.

In conclusion, although consensus regarding the definition of SLE remission and classification has not been fully established, agreement exists that clinical disease activity, serologic activity, duration, and treatment are critical components of defining remission status. Our findings show that remission can be regarded as an achievable and desirable target for SLE patients. Patients' demographic data, initial disease activity, clinical manifestations, serologic results, as well as treatment regimens, are all found to play key roles in the achievement of remission. Remission, especially prolonged remission, has been associated with a better prognosis in terms of damage accrual and HRQoL in SLE patients. Due to the heterogeneity of SLE clinically, however, very few recommendations have been proposed regarding the prognostic capabilities of different definitions of remission that would be broadly applicable. Establishing risk stratification models to select the patients who require intensive therapeutic regimens to achieve remission is critical. Development of a full algorithm with frequency of monitoring, time points for evaluation of remission and adjustment of treatment if remission is not achieved, and the timing when glucocorticoids can be further tapered or withdrawn after achieving target will be important parts of the research agenda in the coming years.

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. Li and Zeng 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. Yang, Cheng, Zhao, Q. Wang, Tian, Hsieh, Li, Zeng.

Acquisition of data. Yang, Cheng.

Analysis and interpretation of data. Z. Wang, Y. Wang.

REFERENCES

1. Smolen JS, Breedveld FC, Burmester GR, Bykerk V, Dougados M, Emery P, et al. Treating rheumatoid arthritis to target: 2014 update of the recommendations of an international task force. Ann Rheum Dis 2016;75:3–15.
2. Van Vollenhoven RF, Mosca M, Bertsias G, Isenberg D, Kuhn A, Lerstrom K, et al. Treat-to-target in systemic lupus erythematosus: recommendations from an international task force. Ann Rheum Dis 2014;73:958–67.
3. Van Vollenhoven R, Voskuyl A, Bertsias G, Aranow C, Aringer M, Arnaud L, et al. A framework for remission in SLE: consensus findings from a large international task force on definitions of remission in SLE (DORIS). Ann Rheum Dis 2017;76:554–61.
4. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 2009;339:b2700.
5. Tozman EC, Urowitz MB, Gladman DD. Prolonged complete remission in previously severe SLE. Ann Rheum Dis 1982;41:39–40.
6. Gladman DD, Urowitz MB, Keystone EC. Serologically active clinically quiescent systemic lupus erythematosus: a discordance between clinical and serologic features. Am J Med 1979;66:210–5.
7. Heller CA, Schur PH. Serological and clinical remission in systemic lupus erythematosus. J Rheumatol 1985;12:916–8.
8. Schneider M. Response and remission criteria for clinical trials in lupus-what can we learn from other diseases? Lupus 1999;8:627–31.
9. Urowitz MB, Feletar M, Bruce IN, Ibanez D, Gladman DD. Prolonged remission in systemic lupus erythematosus. J Rheumatol 2005;32:1467–72.
10. Steiman AJ, Urowitz MB, Ibanez D, Papneja A, Gladman DD. Prolonged clinical remission in patients with systemic lupus erythematosus. J Rheumatol 2014;41:1808–16.
11. Zen M, Iaccarino L, Gatto M, Bettio S, Nalotto L, Ghirardello A, et al. Prolonged remission in Caucasian patients with SLE: prevalence and outcomes. Ann Rheum Dis 2015;74:2117–22.
12. Medina-Quiñones CV, Ramos-Merino L, Ruiz-Sada P, Isenberg D. Analysis of complete remission in systemic lupus erythematosus patients over a 32-year period. Arthritis Care Res (Hoboken) 2016;68:981–7.
13. Alarcon GS, Ugarte-Gil MF, Pons-Estel G, Vila LM, Reveille JD, McGwin G Jr. Remission and low disease activity state (LDAS) are protective of intermediate and long-term outcomes in SLE patients. Results from LUMINA (LXXVIII), a multiethnic, multicenter US cohort. Lupus 2019;28:423–6.
14. Ugarte-Gil MF, Wojdyla D, Pons-Estel GJ, Quintana R, Gómez-Puerta JA, Catoggio LJ, et al. Predictors of remission and low disease activity state in systemic lupus erythematosus: data from a multiethnic, multinational Latin American cohort. J Rheumatol 2019;46:1299–308.
15. Sebastiani GD, Prevete I, Iuliano A, Piga M, Iannone F, Coladonato L, et al. Early Lupus Project: one-year follow-up of an Italian cohort of patients with systemic lupus erythematosus of recent onset. Lupus 2018;27:1479–88.
16. Mok CC, Ho LY, Tse SM, Chan KL. Prevalence of remission and its effect on damage and quality of life in Chinese patients with systemic lupus erythematosus. Ann Rheum Dis 2017;76:1420–5.
17. Kasitanon N, Intaniwet T, Wangkaew S, Pantana S, Sukitawut W, Louthrenoo W. The clinically quiescent phase in early-diagnosed SLE patients: inception cohort study. Rheumatology (Oxford) 2014;54:868–75.
18. Tselios K, Gladman DD, Touma Z, Su J, Anderson N, Urowitz MB. Monophasic disease course in systemic lupus erythematosus. J Rheumatol 2018;45:1131–5.
19. Zen M, Bassi N, Nalotto L, Canova M, Bettio S, Gatto M, et al. Disease activity patterns in a monocentric cohort of SLE patients: a seven-year follow-up study. Clin Exp Rheumatol 2012;30:856–63.
20. Steiman AJ, Gladman DD, Ibañez D, Urowitz MB. Outcomes in patients with systemic lupus erythematosus with and without a

prolonged serologically active clinically quiescent period. Arthritis Care Res (Hoboken) 2012;64:511–8.
21. Steiman AJ, Gladman DD, Ibañez D, Urowitz MB. Prolonged serologically active clinically quiescent systemic lupus erythematosus: frequency and outcome. J Rheumatol 2010;37:1822–7.
22. Ng KP, Manson JJ, Rahman A, Isenberg DA. Association of antinucleosome antibodies with disease flare in serologically active clinically quiescent patients with systemic lupus erythematosus. Arthritis Rheum 2006;55:900–4.
23. Barr SG, Zonana-Nacach A, Magder LS, Petri M. Patterns of disease activity in systemic lupus erythematosus. Arthritis Rheum 1999;42: 2682–8.
24. Formiga F, Moga I, Pac M, Mitjavila F, Rivera A, Pujol R. High disease activity at baseline does not prevent a remission in patients with systemic lupus erythematosus. Rheumatology (Oxford) 1999;38:724–7.
25. Drenkard C, Villa AR, Garcia-Padilla C, Pérez-Vázquez ME, Alarcón-Segovia D. Remission of systematic lupus erythematosus. Medicine (Baltimore) 1996;75:88–98.
26. Fasano S, Margiotta DP, Pierro L, Navarini L, Riccardi A, Afeltra A, et al. Prolonged remission is associated with a reduced risk of cardiovascular disease in patients with systemic lupus erythematosus: a GIRRCS (Gruppo Italiano di Ricerca in Reumatologia Clinica e Sperimentale) study. Clin Rheumatol 2019;38:457–63.
27. Romo-Tena J, Reyna-de la Garza R, Bartnicki-Navarrete I, Alcocer-Varela J, Gomez-Martin D. Factors associated with remission in patients with systemic lupus erythematosus: new insights into a desirable state. Clin Rheumatol 2018;37:3033–42.
28. Zen M, Iaccarino L, Gatto M, Bettio S, Saccon F, Ghirardello A, et al. The effect of different durations of remission on damage accrual: results from a prospective monocentric cohort of Caucasian patients. Ann Rheum Dis 2017;76:562–5.
29. Tani C, Vagelli R, Stagnaro C, Carli L, Mosca M. Remission and low disease activity in systemic lupus erythematosus: an achievable goal even with fewer steroids? Real-life data from a monocentric cohort. Lupus Sci Med 2018;5:e000234.
30. Mathian A, Mouries-Martin S, Dorgham K, Devilliers H, Yssel H, Garrido Castillo L, et al. Ultrasensitive serum interferon-α quantification during SLE remission identifies patients at risk for relapse. Ann Rheum Dis 2019;78:1669–76.
31. Ruiz-Irastorza G, Ruiz-Estevez B, Lazaro E, Ruiz-Arruza I, Duffau P, Martin-Cascon M, et al. Prolonged remission in SLE is possible by using reduced doses of prednisone: an observational study from the Lupus-Cruces and Lupus-Bordeaux inception cohorts. Autoimmun Rev 2019;18:102359.
32. Margiotta DP, Fasano S, Basta F, Pierro L, Riccardi A, Navarini L, et al. The association between duration of remission, fatigue, depression and health-related quality of life in Italian patients with systemic lupus erythematosus. Lupus 2019;28:1705–11.
33. Poomsalood N, Narongroeknawin P, Chaiamnuay S, Asavatanabodee P, Pakchotanon R. Prolonged clinical remission and low disease activity statuses are associated with better quality of life in systemic lupus erythematosus. Lupus 2019;28:1189–96.
34. Tsang AS, Bultink IE, Heslinga M, van Tuyl LH, van Vollenhoven RF, Voskuyl AE. The relationship between remission and health-related quality of life in a cohort of SLE patients. Rheumatology (Oxford) 2019;58:628–35.
35. Ugarte-Gil MF, Pons-Estel GJ, Vila LM, McGwin G, Alarcón GS. Time in remission and low disease activity state (LDAS) are associated with a better quality of life in patients with systemic lupus erythematosus: results from LUMINA (LXXIX), a multiethnic, multicentre US cohort. RMD Open 2019;5.
36. Tsang-A-Sjoe MW, Bultink IE, Heslinga M, Voskuyl AE. Both prolonged remission and Lupus Low Disease Activity State are associated with reduced damage accrual in systemic lupus erythematosus. Rheumatology (Oxford) 2017;56:121–8.
37. Tselios K, Gladman DD, Touma Z, Su J, Anderson N, Urowitz MB. Disease course patterns in systemic lupus erythematosus. Lupus 2019;28:114–22.
38. Tselios K, Gladman DD, Touma Z, Su J, Anderson N, Urowitz MB. Clinical remission and low disease activity outcomes over 10 years in systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2019;71:822–8.
39. Wilhelm TR, Magder LS, Petri M. Remission in systemic lupus erythematosus: durable remission is rare. Ann Rheum Dis 2017;76: 547–53.
40. Gladman D, Ginzler E, Goldsmith C, Fortin P, Liang M, Urowitz M, et al. The development and initial validation of the Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index for Systemic Lupus Erythematosus. Arthritis Rheum 1996;39:363–9.
41. Ugarte-Gil MF, Wojdyla D, Pons-Estel GJ, Catoggio LJ, Drenkard C, Sarano J, et al. Remission and Low Disease Activity Status (LDAS) protect lupus patients from damage occurrence: data from a multiethnic, multinational Latin American Lupus Cohort (GLADEL). Ann Rheum Dis 2017;76:2071–4.
42. Fanouriakis A, Kostopoulou M, Alunno A, Aringer M, Bajema I, Boletis JN, et al. 2019 update of the EULAR recommendations for the management of systemic lupus erythematosus. Ann Rheum Dis 2019;78:736–45.
43. Petri M, Magder LS. Comparison of remission and Lupus Low Disease Activity State in damage prevention in a United States systemic lupus erythematosus cohort. Arthritis Rheumatol 2018;70:1790–5.
44. Parodis I, Johansson P, Gomez A, Soukka S, Emamikia S, Chatzidionysiou K. Predictors of low disease activity and clinical remission following belimumab treatment in systemic lupus erythematosus. Rheumatology (Oxford) 2019;58:2170–6.
45. Gatto M, Zen M, Iaccarino L, Doria A. New therapeutic strategies in systemic lupus erythematosus management. Nat Rev Rheumatol 2019;15:30–48.
46. Smolen JS, Landewe RB, Bijlsma JW, Burmester GR, Dougados M, Kerschbaumer A, 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.
47. Ajeganova S, Huizinga T. Sustained remission in rheumatoid arthritis: latest evidence and clinical considerations. Ther Adv Musculoskelet Dis 2017;9:249–62.
48. Saccon F, Zen M, Gatto M, Margiotta DP, Afeltra A, Ceccarelli F, et al. Remission in systemic lupus erythematosus: testing different definitions in a large multicentre cohort. Ann Rheum Dis 2020;79:943–50.
49. Novelli L, Barbati C, Ceccarelli F, Perricone C, Spinelli FR, Alessandri C, et al. CD44v3 and CD44v6 isoforms on T cells are able to discriminate different disease activity degrees and phenotypes in systemic lupus erythematosus patients. Lupus 2019;28:621–8.
50. Zen M, Saccon F, Gatto M, Montesso G, Larosa M, Benvenuti F, et al. Prevalence and predictors of flare after immunosuppressant discontinuation in patients with systemic lupus erythematosus in remission. Rheumatology (Oxford) 2020;59:1591–8.
51. Ugarte-Gil MF, Gamboa-Cárdenas RV, Reátegui-Sokolova C, Medina-Chinchón M, Zevallos F, Elera-Fitzcarrald C, et al. Better health-related quality of life in systemic lupus erythematosus predicted by low disease activity state/remission: data from the Peruvian Almenara Lupus Cohort. Arthritis Care Res (Hoboken) 2020;72:1159–62.
52. Goswami RP, Chatterjee R, Ghosh P, Sircar G, Ghosh A. Quality of life among female patients with systemic lupus erythematosus in remission. Rheumatol Int 2019;39:1351–8.
53. Tani C, Elefante E, Signorini V, Zucchi D, Lorenzoni V, Carli L, et al. Glucocorticoid withdrawal in systemic lupus erythematosus: are remission and low disease activity reliable starting points for stopping treatment? A real-life experience. RMD Open 2019;5.

Arthritis Care & Research
Vol. 74, No. 2, February 2022, pp 219–228
DOI 10.1002/acr.24444

American College of Rheumatology
Empowering Rheumatology Professionals

Significant Gains in Rheumatoid Arthritis Quality Measures Among RISE Registry Practices

Zara Izadi,[1] Gabriela Schmajuk,[2] Milena Gianfrancesco,[1] Meera Subash,[1] Michael Evans,[1] Laura Trupin,[1] and Jinoos Yazdany[1]

Objective. Using the American College of Rheumatology Rheumatology Informatics System for Effectiveness (RISE) registry, our objective was to examine performance on rheumatoid arthritis (RA) quality measures and to assess the association between practice characteristics and changes in performance over time among participating practices.

Methods. We analyzed data from practices enrolled in RISE between January 1, 2015 and December 31, 2017. Eight quality measures in the areas of RA disease management, cardiovascular risk reduction, and patient safety were examined. Variability in performance was evaluated at the practice level. Multivariate linear models were used to predict change in measure performance by year and to determine the effect of practice characteristics on change in performance over time.

Results. Data from 59,986 patients from 54 practices were examined. The mean ± SD age was 62 ± 14 years, 77% were female, 69% were Caucasian, and most patients were seen in a single-specialty group practice (46%). The average performance on measures related to RA treatments was consistently high (>90%) across the study period. Measures related to RA functional status and disease activity assessment had the greatest improvements over time (8.4% and 13.0% increase per year, respectively; $P < 0.001$). Single-specialty group practices had the fastest rates of improvement over time across all measures.

Conclusion. Among practices participating in RISE between 2015 and 2017, performance on most RA quality measures improved. Single-specialty group practices saw the fastest rates of improvement over time. Identification of workflow patterns leading to dramatic improvements in quality of care will help guide process redesign to address gaps in priority areas, such as tuberculosis screening and blood pressure control.

INTRODUCTION

The Rheumatology Informatics System for Effectiveness (RISE) is an electronic health record (EHR)–enabled registry developed by the American College of Rheumatology (ACR) to facilitate quality improvement among rheumatology practices nationally. RISE passively extracts EHR data from individual practices, aggregates and analyzes these data centrally, and feeds this information back to clinicians as actionable data using a web-based quality dashboard. By providing robust health IT infrastructure, the registry aims to decrease the burden of data collection on practices and streamline participation in federal quality reporting programs such as the Merit-Based Incentive Payment System. An additional benefit of the dashboard is to facilitate local rapid-cycle quality improvement by providing continuous performance feedback and benchmarking (1).

Previous studies have shown gaps in rheumatoid arthritis (RA) care in many different settings, including timely initiation and maintenance of RA treatments (2,3), patient safety (4), and cardiovascular risk reduction (5). However, these studies have been

The views expressed herein represent those of the authors and not necessarily those of the American College of Rheumatology, of the Agency for Healthcare Research and Quality, or of the National Institute of Arthritis and Musculoskeletal and Skin Diseases.

Supported by the American College of Rheumatology (Rheumatology Informatics System for Effectiveness registry), the Agency for Healthcare Research and Quality (grants R18-HS-025638 and R01-HS-024412), and the NIH (National Institute of Arthritis and Musculoskeletal and Skin Diseases grants K24-AR-074534 and P30-AR-070155). Dr. Schmajuk's work was supported by the Russell/Engleman Medical Research Center for Arthritis.

[1]Zara Izadi, PharmD, MAS, Milena Gianfrancesco, MPH, PhD, Meera Subash, MD, Michael Evans, MS, Laura Trupin, MPH, Jinoos Yazdany, MD, MPH: University of California, San Francisco; [2]Gabriela Schmajuk, MD, MSc: San Francisco Veterans Affairs Medical Center and University of California, San Francisco.

No potential conflicts of interest relevant to this article were reported.

Address correspondence to Zara Izadi, PharmD, MAS, Department of Epidemiology and Biostatistics, Mission Hall, Global Health and Clinical Sciences Building, 550 16th Street, 2nd Floor, Box #0560, San Francisco, CA 94158-2549. Email: zara.izadi@ucsf.edu.

Submitted for publication March 13, 2020; accepted in revised form September 8, 2020.

SIGNIFICANCE & INNOVATIONS

- Among practices participating in the Rheumatology Informatics System for Effectiveness (RISE) registry from 2015 to 2017, performance on most measures for individuals with rheumatoid arthritis improved.
- There were significant variations in performance over time between practices, suggesting that future work to identify workflow patterns leading to high performance or to dramatic improvements in quality are warranted.
- Performance on quality measures across RISE practices provides a useful benchmark for rheumatologists seeking to improve quality in their practices.

limited to single institutions or regions, or to administrative data from a single insurance carrier, thus lacking generalizability. In recent years, several new performance measures for RA have been endorsed that are now operationalized as EHR-enabled measures (e-measures). New e-measures such as disease activity and functional status assessments require changes in workflow, which can make implementation difficult (6,7). Practices with long-established workflows may be more equipped to capture specific data elements than practices with more recent changes to clinical workflows. Existing studies did not report on new measures, nor on whether participation in the RISE registry, with access to a dashboard that facilitates quality improvement, were associated with improvements in performance. In addition, although EHR-derived performance on quality measures has been previously reported (8), due to lack of interoperability, direct comparisons of performance on measures between different EHR systems has not been possible to date among rheumatology practices.

In this study we aimed to examine performance on 8 quality measures most relevant to the care of patients with RA and to identify practice characteristics associated with high performance or substantial improvements in performance over a 3-year period. We examined performance on measures for the subset of patients with RA because RA was the initial focus for quality improvement for the registry.

PATIENTS AND METHODS

Study population and timelines. Data were derived from the ACR's RISE registry. RISE is a national EHR-enabled registry that passively collects data on all patients seen by participating practices, reducing the selection bias present in single-insurer claims databases (9). As of December 2017, RISE held validated data from 1,257 providers in 236 practices, representing ~36% of the US clinical rheumatology workforce. We analyzed data collected on all patients with a prevalent diagnosis of RA between January 1, 2015 and December 31, 2017. For each quality measure, the measurement period was defined as the 12 months (24 months for tobacco-use screening and cessation) preceding the last date of each quarter during which a visit occurred (e.g., if a patient with RA had an RA-coded visit on December 15, 2015, the measurement period was defined from January 1, 2015 to December 31, 2015).

RISE is a dynamic registry with practices able to enter and leave over time. To allow for consistent longitudinal analysis, we only included practices that were enrolled in RISE during the entire study period and had ≥30 RA patients at every quarter to reduce the variation in performance due to small sample sizes. Overall, 54 practices were included; 6 were excluded because they did not contribute data during the entire period, and 4 were excluded because they did not have ≥30 RA patients at every quarter. We included patients who had at least 1 clinical face-to-face encounter in each quarter of the year; thus patients were not included in the denominator during quarters when they did not have any clinical face-to-face encounters. An RA diagnosis was defined as having 2 International Classification of Diseases, Ninth Revision (ICD-9) or International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10) codes for RA (714.0 and M06.9, respectively) ≥30 days apart. For each patient, we only included quality measures that were recorded at or after the first clinic visit associated with an ICD-9/ICD-10 code for RA.

Quality measures. As of December 2017, the RISE registry calculated patient-level performance on 24 quality measures (complete list available in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24444/abstract). We examined quality measures in the areas of RA management, including disease activity assessment, functional status assessment, and disease-modifying antirheumatic drug (DMARD) prescribing. Cardiovascular risk reduction and patient safety measures, specifically tobacco-use screening and cessation counseling, blood pressure control, tuberculosis (TB) screening prior to biologic drug start, and use of high-risk medication in the elderly were also examined. Performance on each measure was defined as detailed in Table 1. The use of high-risk medication in the elderly measures are reported to the Merit-Based Incentive Payment System as inverse measures, with lower percentages indicating better performance. To pool performance across quality measures in this study, performance on these (inverted) measures was inverted, such that higher percentages indicated better performance, e.g., a performance of 1% on the inverted measure became 99% in the modified measure. We selected these 8 measures because they are relevant to the care of RA patients, are endorsed by the National Quality Forum, and have been implemented in the RISE registry since January 2015.

Table 1. RISE registry quality measures included in this study*

Measure ID	NQF no.	CMS no.	Measure title	Measure definition	NQS domain	Subspecialty	Measure type
ACR 01	2523	NA	Disease activity measurement for patients with rheumatoid arthritis	Percentage of patients age ≥18 years with a diagnosis of RA whose disease activity is assessed using a standardized measurement tool at 50% or more encounters for RA with the same clinician during the measurement period	Effective clinical care	Rheumatoid arthritis measures	Process
ACR 02	2524	NA	Functional status assessment for patients with rheumatoid arthritis	Percentage of patients age ≥18 years with a diagnosis of RA whose functional status is assessed using a standardized measurement tool at least once during the measurement period	Effective clinical care	Rheumatoid arthritis measures	Process
ACR 03	0054	NA	Disease-modifying antirheumatic drug therapy for active rheumatoid arthritis	Percentage of patients age ≥18 years with active RA who are treated with a DMARD during the measurement period	Effective clinical care	Rheumatoid arthritis measures	Process
ACR 04	NA	NA	Tuberculosis test prior to first course biologic therapy	Percentage of patients age ≥18 years with a diagnosis of RA who are newly prescribed a biologic therapy during the measurement period and whose medical record indicates TB testing in the 12 months preceding the biologic prescription	Patient safety	TB measures	Process
PQRS 226	0028	138v4	Preventive care and screening: tobacco use: screening and cessation intervention	Percentage of patients age ≥18 years who were screened for tobacco use 1 or more times within 24 months AND who received cessation counseling intervention if identified as a tobacco user	Community and population health	–	Process, cross-cutting
PQRS 236	0018	165v4	Controlling high blood pressure	Percentage of patients ages 18–85 years who had a diagnosis of hypertension and whose blood pressure was adequately controlled (<140/90 mm Hg) during the measurement period	Effective clinical care	Hypertension measure	Intermediate outcome, cross-cutting
PQRS 238	0022	156v4	Use of high-risk medications in the elderly†	Percentage of patients age ≥66 years who were ordered high-risk medications. Two rates are reported: 1) percentage of patients who were ordered at least 1 high risk medication; 2) percentage of patients who were ordered at least 2 different high risk medications. Inverse measure: lower count indicates better performance	Patient safety	–	Process

* The measurement period for all measures is 12 months, unless stated otherwise. Practice-level performance was calculated at every quarter. Patients included in the denominator at every quarter must have had at least 1 visit during that quarter. ACR = American College of Rheumatology; CMS = Centers for Medicare and Medicaid Services; DMARD = disease-modifying antirheumatic drug; NA = not applicable; NQF = National Quality Forum; NQS = National Quality Strategy; PQRS = Physician Quality Reporting System; RA = rheumatoid arthritis; RISE = Rheumatology Informatics System for Effectiveness; TB = tuberculosis.
† List of medications defined as high-risk available at: https://qpp.cms.gov/docs/QPP_quality_measure_specifications/CQM-Measures/2019_Measure_238_MIPSCQM.pdf.

Other variables. We examined both sociodemographic characteristics of patients as well as practice characteristics using RISE data extracted through June 30, 2018. Practice characteristics included the number of providers, practice type (multispecialty group practice, single-specialty group practice, solo practitioner, and health system), practice EHR software

Table 2. Characteristics of patients and practices in the RISE registry*

Characteristic	Patient-level	Practice-level
Patients (n = 59,986)†		
Age, mean ± SD years	62.3 ± 13.9	–
Female	46,117 (76.9)	–
Race/ethnicity		
Non-Hispanic Caucasian	41,126 (68.6)	–
Black or African American	5,409 (9.0)	–
Hispanic or Latino	2,650 (4.4)	–
Asian	893 (1.5)	–
Other	6,102 (10.2)	–
Missing	3,806 (6.3)	–
Insurance status		
Private	48,402 (80.7)	–
Medicare ≥65 years	3,303 (5.5)	–
Medicaid	2,106 (3.5)	–
Medicare <65 years	1,318 (2.2)	–
None	8 (0)	–
Missing	4,849 (8.1)	–
Practices (patient-level: n = 59,986; practice-level: n = 54)‡		
Practice size		
1–4 providers	23,023 (38.4)	32 (59.3)
5–9 providers	19,805 (33.0)	15 (27.8)
10–20 providers	17,158 (28.6)	7 (13.0)
Practice type		
Single-specialty group	45,018 (75.1)	35 (64.8)
Solo practitioner	4,837 (8.1)	10 (18.5)
Multispecialty group	9,613 (16.0)	8 (14.8)
Health system	518 (0.9)	1 (1.9)
EHR software		
NextGen	40,969 (68.3)	34 (63.0)
eClinicalWorks	12,065 (20.1)	8 (14.8)
GE Centricity	2,753 (4.6)	4 (7.4)
Allscripts	2,631 (4.4)	3 (5.6)
Amazing Charts	886 (1.5)	2 (3.7)
Other§	682 (1.1)	3 (5.6)
US geographic region		
South	38,879 (64.8)	30 (55.6)
West	5,824 (9.7)	12 (22.2)
Northeast	6,076 (10.1)	7 (13.0)
Midwest	9,207 (15.4)	5 (9.3)

* Values are the number (%) unless indicated otherwise. EHR = electronic health record; RISE = Rheumatology Informatics System for Effectiveness.
† Dynamic cohort.
‡ Fixed cohort (practices that remained in RISE from January 2015 to December 2017).
§ Other included Greenway/Primesuite, eMD-Plus, and UniCharts.

(NextGen, eClinicalWorks, GE Centricity, Allscripts, Amazing Charts, Greenway/Primesuite, eMD-Plus, and UniCharts), and US geographical region (South, West, Northeast, and Midwest). Practice-level sociodemographic information was calculated from patients' eligible for each quality measure during each quarter and included mean age, proportion female, proportion non-Caucasian, and proportion with public insurance (Medicare or Medicaid).

Statistical analysis. Practice-level performance on quality measures, defined as the percentage of eligible patients in a practice receiving recommended care, was the primary outcome. For each measure, we reported the median performance and performance at the 99th percentile in each year (2015, 2016, and 2017). We assessed changes in performance over time by calculating the change in performance on each measure across practices and comparing within-practice changes in performance across geographic regions. Intraclass correlation coefficients (ICCs) were calculated to determine how much of the variability in each quality measure was explained by between-practice variability. Statistical process control charts were used to determine whether changes in performance represented common-cause variation or improvement.

We used bivariate hierarchical linear mixed-effects models to predict change in practice-level measure performance per year, accounting for repeated measurement of practices over time

Table 3. Variability in practice-level performance and change in practice-level performance across measures*

Quality measure	2015		2016		2017		Performance 2015–2017, percentage change (min, max)
	RA patients eligible for measure, no.	Practice-level performance	RA patients eligible for measure, no.	Practice-level performance	RA patients eligible for measure, no.	Practice-level performance	
Disease activity assessment	36,355	57, 100	32,344	68, 100	29,803	78, 100	–97, 98
Functional status assessment	36,355	56, 100	32,344	77, 100	29,803	85, 100	–59, 99
DMARD prescribing	36,355	95, 100	32,609	96, 99	29,803	95, 100	–21, 59
TB screening	4,680	67, 100	2,427	71, 100	1,660	67, 100	–72, 74
Tobacco-use screening and cessation	36,989	91, 100	33,331	91, 100	30,497	92, 100	–32, 85
Blood pressure control	4,889	63, 100	6,512	63, 93	6,924	58, 100	–58, 46
One high-risk medication in elderly	15,910	97, 100	15,170	97, 100	15,207	97, 100	–12, 82
Two high-risk medications in elderly	15,910	100, 100	15,170	100, 100	15,207	100, 100	–1, 1

* Values are the 50th and 99th percentile unless indicated otherwise. Performance on quality measures was defined as the percentage of eligible patients within a practice receiving recommended care. DMARD = disease-modifying antirheumatic drug; max = maximum; min = minimum; RA = rheumatoid arthritis; TB = tuberculosis.

(10). To determine the independent effect of practice characteristics on practice-level measure performance, we used multivariate models. All multivariate models were adjusted for time (as a continuous predictor) and included the number of providers in the practice, practice type, EHR software, and US region as predictors. We adjusted for US region to account for residual confounding that may remain due to geographical variations in practice characteristics that were not available or the underlying prevalence of exposures or disease; for example, providers may be more likely to screen for TB in regions that have a high prevalence of this condition. Patient sociodemographic factors (mean age, proportion female, proportion non-Caucasian, and proportion with public insurance) varied across practices. To account for these differences, the demographics of patients eligible for each quality measure at each practice were indirectly standardized to the overall RISE patient population eligible for the respective quality measure using standardized ratio weights (11,12). Finally, we evaluated whether the change in measure performance over time was modified by any of the practice characteristics by fitting an interaction term between time and each of the covariates in separate multivariate models that included all practice predictors and accounted for repeated measurement of practices over time. All models were checked for linearity of continuous predictors using component plus residual plots and normality of the residuals using residual versus predictor and QQ plots. All analyses were performed in Stata software, version 16.0. The study procedures were approved by the University of California, San Francisco, Institutional Review Board.

RESULTS

Data from 59,986 patients from 54 practices were examined. The mean ± SD age was 62 ± 14 years, 77% were female, 69% were Caucasian, and 81% had private insurance. The most common practice structure was a single-specialty group practice (65%), followed by solo practitioner (18.5%) and multispecialty group practice (15%). NextGen was the most commonly used EHR brand (63%), followed by eClinicalWorks (15%) and GE Centricity (7%). Most practices were in the South region of the US (56%), followed by West (22%) and Northeast regions (13%) (Table 2).

We found a large amount of variability in performance on these measures over the 3-year study period, observing high-performing practices and practices with substantial improvements over time (Table 3). Performance on disease activity assessment, functional status assessment, TB screening, and blood pressure control had the highest variability across practices, with percentage point changes over the 3 years, ranging from –97% to 99%; as a result, these quality measures were selected for further multivariate analysis. Median performance on tobacco-use screening and cessation was high (>80%) and saw the greatest improvements during the earlier parts of the study period. The average performance on DMARD prescribing, use of 1 high risk medication, and use of ≥2 high risk medications in the elderly was also consistently high (>90%) (Figure 1) (control charts for all measures available in the supplementary material; see Supplementary Tables 1 and 2 and Supplementary Figures 1–8, available on the *Arthritis Care & Research* website at

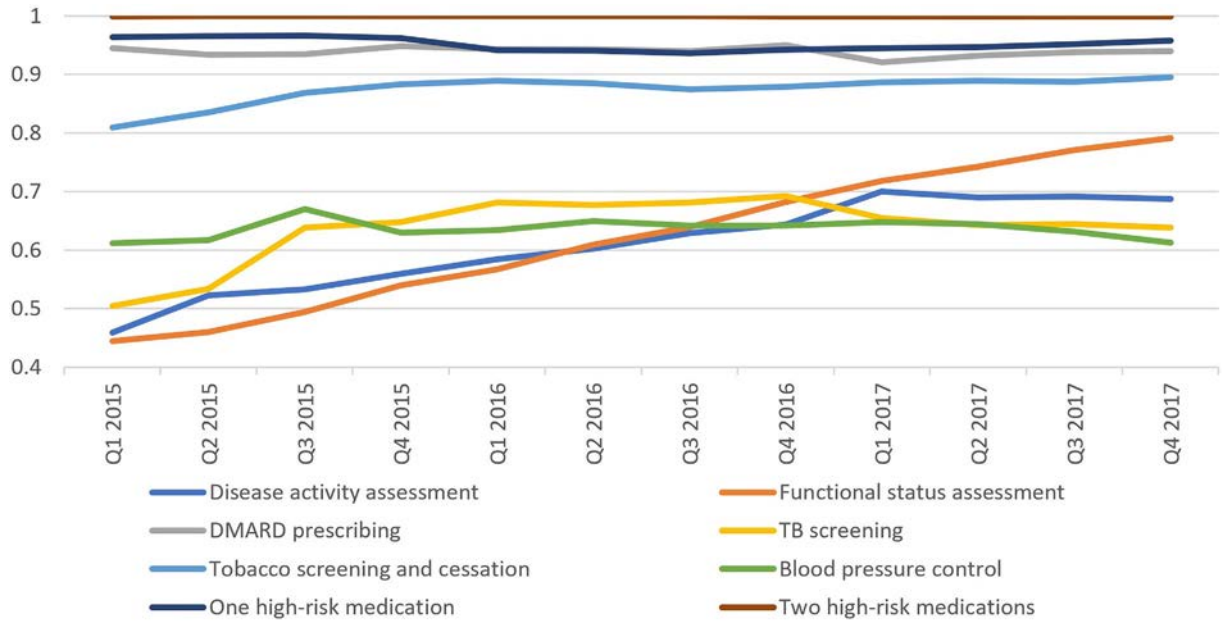

Figure 1. Proportion of patients with rheumatoid arthritis meeting the quality measures from 2015 to 2017 (y-axis). DMARD = disease-modifying antirheumatic drug; Q = quarter; TB = tuberculosis.

http://onlinelibrary.wiley.com/doi/10.1002/acr.24444/abstract). We assessed change in performance over time on disease activity assessment, functional status assessment, TB screening, and blood pressure control across US regions. Most practices in the Northeast region saw no change or decreases in average performance on these 4 measures, while practices in the West had improvements in performance (Figure 2). We also observed few practices within the South with substantial improvements in performance over time.

Between-practice variability explained about one-half of the variation in performance on quality measures across the study period (ICCs ranged from 41% for tobacco-use screening and cessation to 58% for TB screening), indicating important within-practice changes in performance over time. Results from bivariate hierarchical linear mixed-effects models (predicting change in performance as a function of time) from January 2015 to December 2017 showed a significant improvement in performance on functional status assessment (13.9% per year [95% confidence interval (95% CI) 11.8, 16]; $P < 0.001$) and disease activity assessment (8.4% [95% CI 6.2, 10.5]; $P < 0.001$).There were smaller improvements in performance on TB screening (4.3% [95% CI 2.8, 5.7]; $P < 0.001$) and tobacco-use screening and cessation (2.9% [95% CI 1.8, 4]; $P < 0.001$). While improvements in blood pressure control (1.6% [95% CI 0.2, 3]; $P = 0.022$) and DMARD prescribing (1% [95% CI 0.3, 1.6]; $P = 0.004$) were statistically significant, they can most reliably be explained by expected common-cause variation (see Supplementary Figures 3 and 6, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24444/abstract). Changes in performance on the use of high-risk medication in the elderly over time were negligible and not statistically significant.

Multivariate analyses showed that at any time point, larger practices with 10–20 providers performed better than small practices with 1–4 providers on all 4 measures (disease activity assessment, functional status assessment, TB screening, and blood pressure control), with differences reaching statistical significance for functional status assessment (75.5% versus 45.0%; $P = 0.001$) and blood pressure control (71.5% versus 59.7%; $P = 0.001$) (see Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24444/abstract). Performance on disease activity assessment and functional status assessment was significantly higher in health systems compared to single-specialty group practices. Northeast region practices had better performance than those in the South on TB screening and blood pressure control but demonstrated worse performance on disease activity assessment.

For each of the 4 quality measures, we also determined the effect of practice characteristics on change in measure-performance per year. Single-specialty group practices had significantly higher rates of improvement per year than health systems across all 4 measures (P for interaction ≤0.010) (Table 4). Single-specialty group practices also had higher gains in performance than multispecialty group practices across measures, although differences did not reach statistical significance. The EHR software eClinicalWorks was associated with faster improvements in functional status assessment than NextGen. NextGen was associated with faster improvements in disease activity assessment than both Allscripts and Amazing Charts. The West was associated with significantly faster improvements in TB screening than the South.

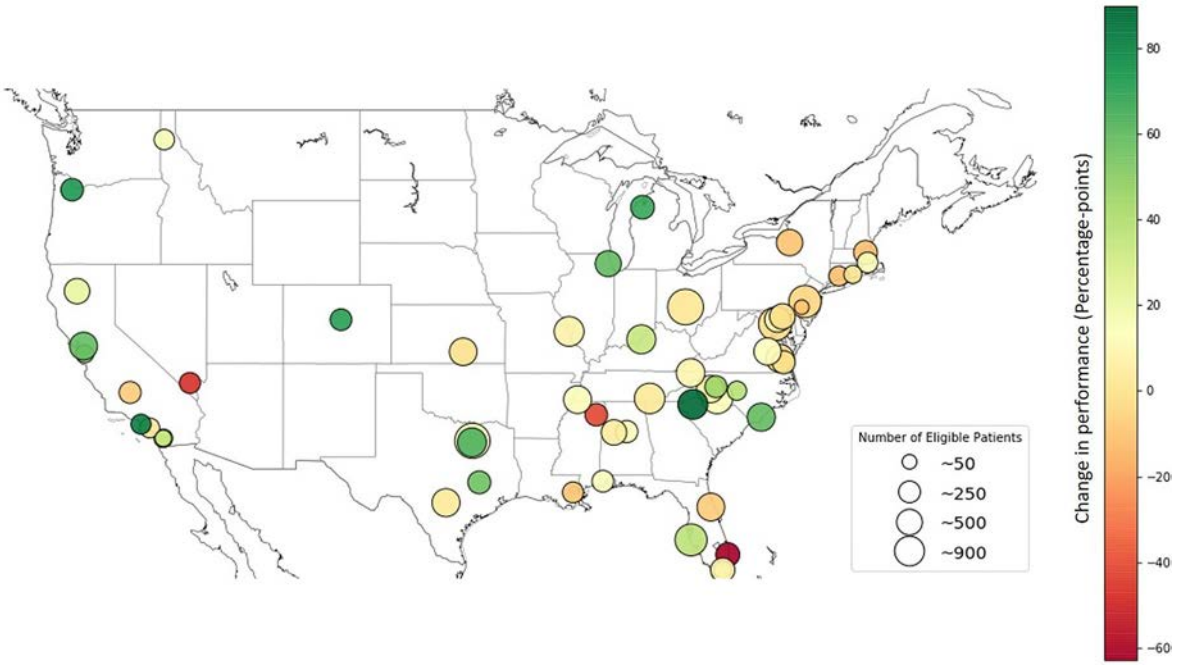

Figure 2. Within-practice change in average measure performance from 2015 to 2017 across the US regions. Measures included disease activity assessment, functional status assessment, blood pressure control, and tuberculosis screening. Cirles represent individual practices; circle size represents practice size.

DISCUSSION

Practices participating in the RISE registry had significant improvements over time in performance on multiple quality measures, including RA disease activity assessment, functional status assessment, TB screening, and tobacco-use screening and cessation. The greatest improvements were in the assessment of functional status and disease activity. We observed considerable variability in performance across practices and regions. Larger practices had better performance on measures compared with small practices, while single-specialty group practices had significantly faster rates of improvement over time compared with multispecialty group practices and health systems.

Two nationally endorsed RA-specific quality measures (disease activity assessment and functional status assessment) are the first examples of e-measures that collect outcomes, including patient-reported outcomes, across the registry. We are encouraged to observe steady and significant improvements in performance on these measures. Measurement of these outcomes using validated tools facilitates a treat-to-target approach and is a key part of high-quality rheumatology care. Additionally, collection of these measures allows for tracking of outcomes, benchmarking across rheumatology practices, and creation of a learning health system in which information about outcomes and performance is fed back to providers to continuously improve quality of care (13). The average performance on DMARD prescribing was consistently high (>90%) from January 2015 to December 2017. This rate is consistent with results from an earlier analysis of this measure using EHR data (94%) (8). Since performance on this measure has been shown to be optimized, this measure has been retired from the Merit-Based Incentive Payment System program.

Regarding measures related to cardiovascular disease prevention, performance on the blood pressure control measure in this study was suboptimal and comparable with previous reports (5). Suboptimal performance on other cardiovascular risk reduction measures has also been reported among RA patients in the US and Canada, including hyperlipidemia and diabetes mellitus screening (14–16). Protocols for controlling hypertension are proven to be effective in primary care settings (17). Such protocols have not been extensively studied in specialty settings such as rheumatology clinics, but they hold potential. A recent study on the implementation of a rheumatology–primary care partnership protocol for the management of high blood pressure showed that timely primary care follow-ups for patients with in-network primary care reduced rheumatology visits by patients with high blood pressure, indicating reduced population-level rates of high BP (18). Performance on tobacco-use screening and cessation in this study was significantly higher than indicated from a previous study using manually abstracted data at an academic rheumatology practice (smoking status was documented at 39% of visit notes with smokers, and smoking cessation counseling was documented in 10%) (19). Incompleteness of manual chart review abstraction and incomplete documentation in notes could explain part of the discrepancy. Notably, tobacco-use screening and cessation were part of the Meaningful Use program implemented

across specialties in 2014, which may explain the steady increase in performance during the earlier parts of the study period.

Among patient safety measures, while performance on TB screening has improved since early 2015, the evident lower performance on this measure indicates both a gap in quality and the fact that reliably capturing TB screening in practice requires further work to ensure accurate data capture from the EHR. Low performance rates on the TB screening measure have been reported previously, even in studies that used extensive chart reviews to examine performance on this measure, and look-back periods that were longer than 12 months to define incident users (4). Therefore, the low performance on the TB measure in RISE probably represents a meaningful gap in care. However, since there is currently no clear evidence to guide appropriate look-back periods for TB testing, possibly the low performance at least partially reflects clinical controversy about which patients need updated TB screening.

We observed better performance on quality measures among larger practices at all points in time, but single-specialty group practices saw the fastest rates of improvement over time. Larger practices, including health systems and multispecialty group practices, likely have more resources to invest in quality improvement activities and infrastructure, such as workflows that facilitate the documentation of disease activity and functional status assessments. Another explanation might be the availability of structured fields within more mature EHR systems for documentation of disease activity and patient-reported outcomes. Single-specialty group practices saw greater gains in performance over time, possibly because RISE provides IT infrastructure for quality improvement, and also because those practices had a lower performance at the start of the study period and hence more room for improvement. Practice and regional variability data from RISE facilitate identification of targets for quality improvement and education initiatives.

We were interested in exploring whether EHR software was associated with performance, but we found that commonly used vendors such as NextGen, eClinicalWorks, and GE Centricity were comparable across most quality measures. This finding suggests that practices can achieve high performance regardless of software, and also that current software does not seem to support high-quality performance. In addition, practices that join RISE may be selected to have clinical workflows to electronically capture required information and the necessary health information technology support staff to build and test the quality measures locally. Notably, measures selected in this study were part of different incentive programs over time; for example, tobacco cessation counseling was a meaningful-use measure and therefore less likely to have significant variation between EHR vendors. In contrast, rheumatology-specific measures, such as functional status or disease activity assessments are less likely to be supported uniformly by EHRs and therefore may be more prone to variability. Notably, EPIC, one of the largest market-share holders among EHR vendors, was not used by practices participating in RISE during the

Table 4. The effect of practice characteristics on change in measure performance per year*

	Disease activity		Functional status		Tuberculosis screening		Blood pressure control	
	Change in performance	*P*	Change in performance	*P*	Change in performance	*P*	Change in performance	*P*
Number of providers								
1–4	5.3 (–0.9, 11.5)	Ref.	14 (8.2, 19.8)	Ref.	5.2 (0.2, 10.3)	Ref.	1.7 (–2.4, 5.8)	Ref.
5–9	13.4 (3.3, 23.6)	0.178	12.4 (–0.1, 24.9)	0.823	3.3 (–1.9, 8.4)	0.586	0.8 (–1.8, 3.3)	0.709
10–20	10.1 (0.2, 20)	0.420	18.2 (7.6, 28.8)	0.490	2.1 (–5.3, 9.6)	0.493	–0.2 (–2.3, 1.9)	0.415
Practice type								
Single-specialty group	13 (6.8, 19.1)	Ref.	17.7 (10.9, 24.5)	Ref.	4.7 (0.6, 8.9)	Ref.	1.7 (–1.7, 5)	Ref.
Solo practitioner	0.8 (–9.9, 11.4)	0.052	8.2 (–0.9, 17.3)	0.097	9.2 (0.7, 17.7)	0.348	2.3 (–2.4, 6.9)	0.835
Multispecialty group	0 (–4.7, 4.6)	0.001	7.4 (–1.8, 16.7)	0.078	–0.2 (–6.9, 6.5)	0.217	–1.5 (–7.8, 4.9)	0.386
Health system	–15.4 (–35.1, 4.3)	<0.001	–0.2 (–15.7, 15.4)	<0.001	–17.1 (–26, –8.1)	<0.001	–2.8 (–7, 1.4)	0.010
EHR brand								
NextGen	9.2 (3.9, 14.5)	Ref.	12.5 (5.8, 19.2)	Ref.	3.9 (–0.3, 8.2)	Ref.	1.5 (–1.8, 4.8)	Ref.
eClinicalWorks	12.6 (1.1, 24.2)	0.587	25.3 (14.8, 35.9)	0.044	7.3 (–0.5, 15.2)	0.446	1 (–4.4, 6.5)	0.881
GE Centricity	13.7 (–11.5, 38.9)	0.726	17.6 (–2.6, 37.7)	0.633	–1.4 (–14.6, 11.9)	0.446	1.4 (–2.1, 4.8)	0.953
Allscripts	–3.2 (–9.6, 3.2)	0.004	–2.2 (–6.4, 2)	<0.001	–0.4 (–9.1, 8.3)	0.372	2 (–1.8, 5.7)	0.858
Amazing Charts	–20.2 (–39.2, –1.2)	0.004	10.4 (–6.9, 27.7)	0.820	–1.2 (–8.5, 6)	0.225	–13.3 (–24.7, –1.8)	0.016
Other	7.4 (–25.5, 40.3)	0.916	16.6 (7.3, 25.8)	0.479	20.2 (8.6, 31.8)	0.011	8.6 (–3.6, 20.8)	0.265
Regions								
South	8 (1.9, 14.1)	Ref.	12.5 (5.7, 19.4)	Ref.	1.4 (–2.4, 5.2)	Ref.	–0.3 (–3.7, 3.1)	Ref.
West	10.9 (–2.7, 24.5)	0.699	21.4 (10.8, 32)	0.163	10 (3, 17)	0.034	5.4 (–0.3, 11)	0.090
Northeast	2.3 (–8.1, 12.7)	0.348	4.9 (–7, 16.7)	0.267	2.7 (–9.3, 14.6)	0.845	–1.4 (–5.4, 2.5)	0.662
Midwest	9.3 (–0.1, 18.8)	0.818	16.2 (2.3, 30.1)	0.637	10 (–4.4, 24.5)	0.254	3.5 (–0.6, 7.5)	0.157

* Values are the percentage (95% confidence interval) unless indicated otherwise. Marginal means were estimated using weighted multivariate regression models accounting for repeated measurement of practices over time. The model additionally incorporated year as a continuous variable. *P* values indicate statistical significance for interaction. EHR = electronic health record; Ref. = reference.

study period due to a combination of factors. Academic centers are the main users of EPIC and faced either institutional or vendor-related barriers in contributing data to the RISE registry.

Our study has important strengths. This study is the first to report change in performance on quality measures over time across a large patient population with diverse geographical coverage across the nation and in RISE practices using different EHR systems. In addition, we used statistical methods that account for variability in sociodemographics across practices and produce reliable estimates generalizable to the overall sociodemographic populations represented within RISE. Limitations of this study include lack of a control group. Without a comparison group of practices who did not join the RISE registry, how much of the improvements in performance on these measures over time are attributable to participation in the registry itself is unclear. These data may underestimate the care provided to patients because documentation within an EHR may be inconsistent, and nonstructured information is difficult to query systematically. To enable a meaningful longitudinal analysis, we included practices that were early and sustained users of the RISE registry; however, these practices were more likely to serve privately insured patients compared to all practices currently participating in the registry (1). Finally, since ICD codes were used to identify RA diagnosis in many denominators, performance on these measures may have been underestimated in our study, as some patients may not have truly had RA. However, since these codes were assigned by rheumatologists, the positive predictive value may be higher than in other studies where codes could be assigned by any provider (20).

With the capacity of RISE to facilitate rapid-cycle quality improvements for participating practices and the emerging payment reforms put into place by Medicare Access and the Children's Health Insurance Program Reauthorization Act of 2015, there is an urgent need to develop new measures to define value in rheumatology. The measures assessed in this study were process measures that assess the actions taken in the course of health care. In 2016, the ACR began development of a new outcome measure to assess the effects of these actions on health status using clinical data from the RISE registry (21). Understanding the scientific validity, feasibility, usefulness, and intended and unintended consequences of quality measures also continues to be an important strategic goal of RISE. As more practices join RISE, larger studies will be powered to facilitate further subgroup analyses that identify target areas for quality improvement. Key questions regarding the role of sociodemographic factors, health care access, and patient satisfaction remain and can serve as the focus of future research. Furthermore, given the variability in performance across RISE practices, further qualitative research is needed to better understand facilitators and barriers to improvement on these measures.

In summary, this article provides a systematic benchmarking of the ACR quality measures using data from 54 practices participating in the RISE registry. Results from this study indicate excellent performance on DMARD prescribing and steady improvements in documentation of disease activity and functional status over a 3-year period between 2015 to 2017. Blood pressure control and TB screening measures may deserve the most attention in performance improvement initiatives, although notable improvements on these measures were observed among some practices. Identification of workflow patterns leading to high performance or dramatic improvements in quality of care will help guide strategies to address gaps in priority areas.

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. Izadi 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. Izadi, Schmajuk, Yazdany.

Acquisition of data. Izadi, Schmajuk, Yazdany.

Analysis and interpretation of data. Izadi, Schmajuk, Gianfrancesco, Subash, Evans, Trupin, Yazdany.

REFERENCES

1. Yazdany J, Bansback N, Clowse M, Collier D, Law K, Liao KP, et al. Rheumatology Informatics System for Effectiveness: a national informatics-enabled registry for quality improvement. Arthritis Care Res (Hoboken) 2016;68:1866–73.
2. Schmajuk G, Trivedi AN, Solomon DH, Yelin E, Trupin L, Chakravarty EF, et al. Receipt of disease-modifying antirheumatic drugs among patients with rheumatoid arthritis in Medicare managed care plans. JAMA 2011;305:480–6.
3. Shafrin J, Ganguli A, Gonzalez YS, Shim JJ, Seabury SA. Geographic variation in the quality and cost of care for patients with rheumatoid arthritis. J Manag Care Spec Pharm 2016;22:1472–81.
4. Patterson S, Schmajuk G, Evans M, Aggarwal I, Izadi Z, Gianfrancesco M, et al. Gaps in ambulatory patient safety for immunosuppressive specialty medications. Jt Comm J Qual Patient Saf 2019;45:348–57.
5. Bartels CM, Johnson H, Alcaraz Voelker K, Ogdie A, McBride P, Jacobs EA, et al. Frequency and predictors of communication about high blood pressure in rheumatoid arthritis visits. J Clin Rheumatol 2018;24:210–7.
6. Abernethy AP, Gippetti J, Parulkar R, Revol C. Use of electronic health record data for quality reporting. J Oncol Pract 2017;13:530–4.
7. Desai SP, Leatherwood C, Forman M, Ko E, Stevens E, Iversen M, et al. Treat-to-target approach in rheumatoid arthritis: a quality improvement trial. Arthritis Care Res (Hoboken) 2021;73:207–14.
8. Adhikesavan LG, Newman ED, Diehl MP, Wood GC, Bili A. American College of Rheumatology quality indicators for rheumatoid arthritis: benchmarking, variability, and opportunities to improve quality of care using the electronic health record. Arthritis Rheum 2008;59:1705–12.
9. Yazdany J, Robbins M, Schmajuk G, Desai S, Lacaille D, Neogi T, et al. Development of the American College of Rheumatology's rheumatoid arthritis electronic clinical quality measures. Arthritis Care Res (Hoboken) 2016;68:1579–90.
10. Ryan TP. Statistical methods for quality improvement. 3rd ed. Hoboken (NJ): Wiley; 2011.
11. Kurth T, Walker AM, Glynn RJ, Chan KA, Gaziano JM, Berger K, et al. Results of multivariable logistic regression, propensity matching,

propensity adjustment, and propensity-based weighting under conditions of nonuniform effect. Am J Epidemiol 2006;163:262–70.

12. Sturmer T, Rothman KJ, Glynn RJ. Insights into different results from different causal contrasts in the presence of effect-measure modification. Pharmacoepidemiol Drug Saf 2006;15:698–709.
13. Schmajuk G, Yazdany J. Leveraging the electronic health record to improve quality and safety in rheumatology. Rheumatol Int 2017;37:1603–10.
14. Schmidt TJ, Avina-Zubieta JA, Sayre EC, Abrahamowicz M, Esdaile JM, Lacaille D. Cardiovascular disease prevention in rheumatoid arthritis: compliance with diabetes screening guidelines. J Rheumatol 2018;45:1367–74.
15. Schmidt TJ, Avina-Zubieta JA, Sayre EC, Abrahamowicz M, Esdaile JM, Lacaille D. Quality of care for cardiovascular disease prevention in rheumatoid arthritis: compliance with hyperlipidemia screening guidelines. Rheumatology (Oxford) 2018;57:1789–94.
16. Navarro-Millán I, Yang S, Chen L, Yun H, Jagpal A, Bartels CM, et al. Screening of hyperlipidemia among patients with rheumatoid arthritis in the United States. Arthritis Care Res (Hoboken) 2019;71:1593–9.
17. Jaffe MG, Lee GA, Young JD, Sidney S, Go AS. Improved blood pressure control associated with a large-scale hypertension program. JAMA 2013;310:699–705.
18. Bartels CM, Ramly E, Johnson HM, Lauver DR, Panyard DJ, Li Z, et al. Connecting rheumatology patients to primary care for high blood pressure: specialty clinic protocol improves follow-up and population blood pressures. Arthritis Care Res (Hoboken) 2019;71:461–70.
19. Vreede AP, Johnson HM, Piper M, Panyard DJ, Wong JC, Bartels CM. Rheumatologists modestly more likely to counsel smokers in visits without rheumatoid arthritis control: an observational study. J Clin Rheumatol 2017;23:273–7.
20. Ng B, Aslam F, Petersen NJ, Yu HJ, Suarez-Almazor ME. Identification of rheumatoid arthritis patients using an administrative database: a Veterans Affairs study. Arthritis Care Res (Hoboken) 2012;64:1490–6.
21. Suter LG, Barber CE, Herrin J, Leong A, Losina E, Miller A, et al. American College of Rheumatology white paper on performance outcome measures in rheumatology. Arthritis Care Res (Hoboken) 2016;68:1390–401.

Arthritis Care & Research
Vol. 74, No. 2, February 2022, pp 229–235
DOI 10.1002/acr.24469

American College of Rheumatology
Empowering Rheumatology Professionals

Weight Fluctuation and the Risk of Cardiovascular Events in Patients With Rheumatoid Arthritis

Joshua F. Baker,[1] George Reed,[2] and Joel Kremer[3]

Objective. Fluctuations in weight have been linked to cardiovascular (CV) outcomes in the general population. The present study was undertaken to evaluate whether weight fluctuation was independently predictive of CV events in patients with rheumatoid arthritis (RA).

Methods. We studied patients with RA from the Corrona registry. Weight change was categorized as loss of ≥10%, loss of 5–10%, stable, gain of 5–10%, and gain of ≥10%. We also categorized patients by quintile of variability in weight in prior observation periods. Cox proportional hazards models explored independent associations between time-varying weight change and weight variability and risk of CV events before and after adjusting for CV risk factors, RA disease features, and disability.

Results. Among 31,381 participants, those who lost or gained 10% of their weight had greater disease activity and worse physical function, and they were more likely to smoke, have diabetes mellitus, receive corticosteroids, and be disabled. In adjusted models, a greater risk of CV events was observed in those who experienced 10% weight loss (hazard ratio [HR] 1.18 [95% confidence interval (95% CI) 1.03–1.36], $P = 0.02$) or weight gain (HR 1.20 [95% CI 1.04–1.38], $P = 0.01$). The association between weight change and CV events was stronger among participants with body mass index <25 kg/m^2 for 10% weight loss (HR 1.34 [95% CI 1.08–1.66], $P = 0.001$] and 10% weight gain (HR 1.74 [95% CI 1.41–2.24], $P < 0.001$). Patients with greater variability in weight had a higher risk of CV events.

Conclusion. Recent changes and high variability in weight predict CV events in RA, particularly among thin patients. Further study is necessary to determine if weight fluctuation has adverse cardiometabolic consequences that are independent of other risk factors.

INTRODUCTION

Weight cycling, or weight fluctuation, is described as high variability in an individual's weight over time (alternating weight loss and gain). In several large studies in the general population, weight fluctuation has been associated with a greater risk of adverse outcomes such as cardiovascular (CV) disease and death (1). For example, among patients with coronary artery disease, those with the greatest variability in weight over time have been shown to have a 2-fold greater risk of CV events and death compared to those with the most stable weight (2,3).

After periods of starvation, mice that are refed gain a high percentage of fat and demonstrate a persistently slow metabolic rate (4). Mice that experience multiple weight cycling events also have greater internal fat deposition even when compared to mice fed a high-fat diet (5). This change in metabolism and increase in internal fat related to weight fluctuation is hypothesized to lead to higher risk of metabolic syndrome and supports a causative role in adverse health outcomes. However, clinical studies have not consistently shown a higher risk of diabetes mellitus among patients who have experienced weight fluctuation (6). Weight fluctuation may also reflect, in some cases, metabolic stress related to poor health.

Patients with rheumatoid arthritis (RA) are known to experience weight changes related to their disease, treatments, and comorbidities (7–9). Weight loss is an important predictor of disability and death in this population (8,10–12). These fluctuations in weight may reflect catabolic processes that are themselves

The contents herein do not represent the views of the Department of Veterans Affairs or the US government.

Dr. Baker's work was supported by the Department of Veterans Affairs (Clinical Science Research and Development Career Development award IK2CX000955 and Merit award I01CX001703).

[1]Joshua F. Baker, MD, MSCE: Philadelphia VA Medical Center and University of Pennsylvania, Philadelphia; [2]George Reed, PhD: Corrona Research Foundation, Albany Medical College, and the Center for Rheumatology, Albany, New York; [3]Joel Kremer, MD: University of Massachusetts Medical School, Worcester.

Dr. Baker has received consulting fees from Bristol Myers Squibb, Gilead, the Corrona registry, and Burns-White (less than 10,000 each). No other disclosures relevant to this article were reported.

Address correspondence to Joshua F. Baker, MD, MSCE, 5th Floor White Building, 3600 Spruce Street, Philadelphia, PA 19104. Email: bakerjo@uphs.upenn.edu.

Submitted for publication April 21, 2020; accepted in revised form September 24, 2020.

SIGNIFICANCE & INNOVATIONS

- Weight changes and weight fluctuation are observed in rheumatoid arthritis in association with disease activity, smoking, disability, and comorbidity.
- Fluctuations in weight are independently associated with cardiovascular events, particularly among thin patients.
- It remains unclear whether fluctuations in weight are predictive of cardiovascular disease through direct detrimental effects or by acting as a marker of poor metabolic health.

predictive of adverse CV outcomes. In cancer, changes in weight may be most predictive among those who are already thin and possibly experiencing disease-related cachexia (13). Thus, variability in weight can be predictive of long-term adverse events either through a link to other underlying health problems or possibly through the adverse consequences of the weight fluctuation itself. No previous studies have evaluated weight changes and the prediction of CV events in patients with RA.

Variability in weight is hypothesized to represent an important predictor that might help in risk stratification in this population. We aimed to determine factors associated with changes in weight over time among patients with RA. We further aimed to determine if recent weight loss and weight gain were associated independently with a greater risk of CV events among patients with RA. We also evaluated a previously defined measure of weight variability and its association with CV events. We hypothesized that more dramatic changes in weight over time would be associated with greater subsequent risk, particularly among normal and underweight patients with RA.

MATERIALS AND METHODS

Study setting. We utilized data from the Corrona database. The Corrona registry was initiated in 2001 and is the largest independent database in North America, collecting data from both rheumatologists and patients at the time of a clinical encounter every 3–6 months. We included all participants enrolled up to February 2017. Data are collected using structured case report forms that include medication use, RA disease activity, function, comorbid illnesses, and acute events, such as CV events, infections, and cancer. We included participants with a diagnosis of RA who had at least 3 clinical visits. Registry activities are approved by the Corrona central Institutional Review Board. All patients signed informed consent prior to participation.

Weight change and variability categorization. The percent change in weight between observations was determined for each participant. The rate of weight change was categorized as the percent change in weight standardized to a 1-year change. For example, a 5% change over 6 months would be considered to represent a rate of change of 10% per year. In these data, the median (interquartile range [IQR]) for the interval between visits was 5.1 months (3.6–6.7 months). The rate of weight change was categorized as ≤–10%, –10 to –5%, –4.99 to 4.99%, 5–10%, and ≥10% change per year. Similar categories of percent weight change have been previously defined, and a >5% weight change over 6–12 months has been considered to represent a clinically important change (14–16). We also summarized absolute weight changes over all prior observation intervals among observations with >2 prior observations and categorized the variability by quintile, as previously described (2,3).

Cardiovascular outcomes. At each registry visit, physicians report whether adverse events have occurred between visits. All physician-reported CV events prompt administration of a second questionnaire to the site to confirm the CV event and to obtain additional details and verify that it was indeed an incident event. We also request additional information to adjudicate a proportion of events including medical records from the treating acute care hospital. All medical records were reviewed by an adjudication committee, as described elsewhere (17). The positive predictive value of the questionnaires was 85%. The components of the previously defined composite CV outcome utilized in the primary analyses are shown in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24469/abstract (17).

Covariable assessments. We evaluated factors that have been incorporated in previous risk assessment tools or that have been previously associated with weight changes in patients with RA (7,8,17). These factors, captured at the time of each individual visit on structured case report forms, included calendar year, demographics, diabetes mellitus, hyperlipidemia, existing CV disease, hypertension, current smoking, alcohol use, and reported exercise. We assessed disease activity using the Clinical Disease Activity Index (CDAI). Physical function was assessed with the modified Health Assessment Questionnaire (M-HAQ). We also incorporated a measure of disease duration and current prednisone use (stratified by dose), methotrexate use, and tumor necrosis factor inhibitor use.

Statistical analysis. Observations missing key data were excluded from the analysis. Visit 2 characteristics for the study population were described across weight change categories (defined based on weight changes occurring between visit 1 and visit 2) and tested for significant differences using analysis of variance or Kruskal-Wallis tests for nonparametric data. Separate mixed effects logistic regression models evaluated predictors of 10% weight loss and 10% weight gain by the subsequent visit

(outcome lagged 1 visit from exposure visit) over all observations in the data set. Models were informed by univariate associations and by hypothesized predictors of weight change. Prehypothesized predictors included age, sex, disease activity, disability, smoking, and prednisone use. With the exception of age, race, disease duration, history of CV disease, and history of diabetes mellitus, we utilized time-varying exposures to model the dynamic interaction more accurately between exposures and the outcome over time.

Baseline (visit 2) characteristics among patients with RA were evaluated by category of weight change since visit 1. Visit 2 was considered baseline because it was the first visit where the exposure of interest could be determined. Weight change between each visit and the prior visit was determined and categorized as described. We quantified overall weight variability based on the SD of the absolute change in weight occurring over all prior observations in the data set, as previously described by Bangalore et al (3). Similar to this prior publication, this exposure was categorized into quintiles.

Cox proportional hazards models evaluated the independent risk associated with recent weight change or weight variability and CV events in sequential multivariable models adjusting for demographics and time-varying covariables including CV risk factors, disease activity and duration, exercise, and physical function and disability. Model specification included tiered models incorporating a priori confounders. The association between changes in CDAI score over the same interval was also explored as a predictor of weight change. A significant change in CDAI score was defined as an absolute change greater than the minimum clinically important difference (MCID) (18).

Prehypothesized interactions between current weight and weight change or variability category were evaluated by testing the significance of multiplicative interaction terms in fully adjusted models and the results of stratified models presented separately. Sensitivity analyses were also performed that evaluated the risk of weight change focused on the following: 1) individuals with no previous history of CV disease; 2) the development of major adverse cardiovascular events (MACE) only (CV death, myocardial infarction, stroke); 3) analyses among observations where the distance between the current and previous visit was >2 months; and 4) adjustment for the total number of visits in the last year. Similar results were also observed when excluding reported body mass index (BMI) values that were identical to the prior observation (data not shown). All analyses were performed using Stata, version 16.0.

Table 1. Characteristics of study participants at visit 2 among patients categorized by weight change from visit 1*

	Percent change in weight (visit 1 to visit 2)					
Characteristic	≤–10% (n = 3,139)	>–10% to ≤–5% (n = 3,317)	>–5% to <5% (n = 16,791)	≥5% to <10% (n = 3,961)	≥10% (n = 3,972)	*P*
Age, mean ± SD years	58.2 ± 13.9	59.3 ± 13.1	58.6 ± 13	57.4 ± 13.3	56.8 ± 13.7	<0.0001
Female	2,498 (79.6)	2,513 (75.8)	12,722 (75.8)	3,087 (77.9)	3,105 (78.2)	<0.0001
RA duration, mean ± SD per yr	9 ± 9.8	9.5 ± 10.2	9.2 ± 9.7	8.7 ± 9.7	8.2 ± 9.4	<0.0001
BMI, mean ± SD kg/m^2	30.5 ± 7.7	29.6 ± 6.9	29.3 ± 6.9	29.1 ± 7	28.4 ± 6.8	<0.0001
Weight, mean ± SD pounds	183 ± 49	179 ± 46	178 ± 45	176 ± 45	172 ± 43	<0.0001
M-HAQ score, mean ± SD	0.4 ± 0.5	0.4 ± 0.4	0.3 ± 0.4	0.3 ± 0.4	0.4 ± 0.5	<0.0001
CDAI score, mean ± SD	13 ± 12.7	11.7 ± 11.9	11.1 ± 11.3	11.8 ± 11.6	13.1 ± 12.2	<0.0001
Disabled	472	380 (11.5)	1,907 (11.4)	449 (11.3)	537 (13.5)	<0.0001
CV disease	297 (9.5)	324 (9.8)	1,433 (8.5)	363 (9.2)	372 (9.4)	0.0793
Hyperlipidemia	662 (21.1)	751 (22.6)	3,630 (21.6)	837 (21.1)	811 (20.4)	0.1944
Hypertension	1,039 (33.1)	1,083 (32.6)	5,280 (31.4)	1,197 (30.2)	1,191 (30)	0.0118
Diabetes mellitus	301 (9.6)	280 (8.4)	1,342 (8)	338 (8.5)	352 (8.9)	0.0314
Statin use	621 (19.8)	664 (20)	3,361 (20)	776 (19.6)	758 (19.1)	0.7396
TNF use	1,223 (39)	1,296 (39.1)	6,662 (39.7)	1,622 (40.9)	1,601 (40.3)	0.3582
MTX use	1,967 (62.7)	2,082 (62.8)	10,423 (62.1)	2,497 (63)	2,575 (64.8)	0.0302
Prednisone	784 (25)	742 (22.4)	3,611 (21.5)	914 (23.1)	1,126 (28.3)	<0.0001
Any alcohol	1,209 (38.5)	1,361 (41)	7,363 (43.9)	1,795 (45.3)	1,638 (41.2)	<0.0001
Any exercise	2,065 (65.8)	2,175 (65.6)	11,174 (66.5)	2,590 (65.4)	2,507 (63.1)	0.0017
Smoking						
Never	1,697 (54.1)	1,758 (53)	9,544 (56.8)	2,249 (56.8)	2,155 (54.3)	
Former	899 (28.6)	1,010 (30.4)	4,826 (28.7)	1,131 (28.6)	1,145 (28.8)	<0.0001
Current	543 (17.3)	549 (16.6)	2,421 (14.4)	581 (14.7)	672 (16.9)	

* Values are the number (%) unless indicated otherwise. BMI = body mass index; CDAI = Clinical Disease Activity Index; CV = cardiovascular; M-HAQ = modified Health Assessment Questionnaire; MTX = methotrexate; RA = rheumatoid arthritis; TNF = tumor necrosis factor.

Table 2. Factors independently associated with 10% loss of body weight (model 1) and 10% gain in body weight (model 2) by the next visit in adjusted logistic regression models*

	10% weight loss		10% weight gain	
	OR (95% CI)	*P*	OR (95% CI)	*P*
Race (versus White)				
Black	0.99 (0.93–1.05)	0.71	1.13 (1.06–1.20)	<0.001
Asian	0.96 (0.85–1.09)	0.54	0.72 (0.63–0.81)	<0.001
BMI category (versus <20 kg/m^2)				
20–25 kg/m^2	1.26 (1.11–1.43)	<0.001	0.59 (0.54–0.65)	<0.001
25–30 kg/m^2	1.59 (1.40–1.81)	<0.001	0.47 (0.42–0.51)	<0.001
30–35 kg/m^2	1.80 (1.58–2.06)	<0.001	0.39 (0.35–0.44)	<0.001
>35 kg/m^2	2.02 (1.77–2.30)	<0.001	0.32 (0.29–0.36)	<0.001
Alcohol use	0.92 (0.90–0.95)	<0.001	0.90 (0.88–0.93)	<0.001
Exercise	0.97 (0.94–0.99)	0.02	0.98 (0.95–1.01)	0.10
CDAI category (versus remission)				
Low	1.04 (1.00–1.08)	0.04	1.04 (1.00–1.08)	0.052
Moderate	1.08 (1.03–1.12)	0.001	1.18 (1.13–1.23)	<0.001
High	1.18 (1.12–1.24)	<0.001	1.37 (1.31–1.44)	<0.001
M-HAQ score (versus 0)				
0–0.125	1.00 (0.95–1.05)	0.97	1.06 (1.01–1.11)	0.01
0.125–0.5	1.08 (1.04–1.12)	<0.001	1.10 (1.06–1.14)	<0.001
0.5–1.0	1.18 (1.13–1.23)	<0.001	1.21 (1.16–1.27)	<0.001
>1.0	1.28 (1.21–1.36)	<0.001	1.28 (1.20–1.35)	<0.001
Disabled	1.09 (1.04–1.14)	<0.001	1.09 (1.04–1.14)	<0.001
Prednisone use (versus none)				
1–4 mg	1.17 (1.11–1.23)	<0.001	1.00 (0.95–1.06)	0.964
5–9 mg	1.10 (1.05–1.15)	<0.001	1.15 (1.10–1.19)	<0.001
≥10 mg	1.12 (1.05–1.19)	0.001	1.53 (1.43–1.61)	<0.001
Methotrexate use	0.99 (0.96–1.01)	0.37	1.05 (1.02–1.08)	0.001
Biologic/tsDMARD use	1.02 (0.99–1.04)	0.30	1.01 (0.98–1.03)	0.72
Smoking (versus never)				
Current	1.23 (1.19–1.29)	<0.001	1.11 (1.07–1.16)	<0.001
Former	1.11 (1.07–1.14)	<0.001	1.12 (1.08–1.15)	<0.001
History of cardiovascular disease	1.02 (0.96–1.07)	0.54	1.08 (1.02–1.14)	0.01
Diabetes mellitus	0.97 (0.92–1.02)	0.26	1.14 (1.08–1.20)	<0.001
Disease duration (per day)	1.002 (1.000–1.003)	<0.001	0.994 (0.992–0.996)	<0.001

* Also adjusted for age, sex, and age × sex interaction. 95% CI = 95% confidence interval; BMI = body mass index; CDAI = Clinical Disease Activity Index; M-HAQ = modified Health Assessment Questionnaire; OR = odds ratio; tsDMARD = targeted synthetic disease-modifying antirheumatic drug.

RESULTS

Identification of factors associated with weight fluctuations. Among 48,535 total patients in the registry, 31,180 had at least 2 follow-up visits and were not missing data for key variables. At visit 2, 16,791 patients (54%) had stable weight since the baseline visit. However, 3,139 (10%) lost weight at a per-year rate of >10%, and 3,972 (13%) gained at a rate of ≥10% per year. Table 1 illustrates the baseline characteristics of patients stratified by changes in weight between visit 1 and visit 2. There were a number of significant differences in these baseline characteristics. Notably, those who lost or gained at a rate of ≥10% of their weight per year had significantly higher disability and disease activity, were more likely to smoke and use prednisone, and were less likely to drink alcohol. Those who gained weight tended to be younger and have shorter disease duration.

Independent predictors in models of weight loss and weight gain over all observations in the registry are shown in Table 2 and Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24469/abstract. Overall, older patients were less likely to gain weight, and women were more likely to gain at early ages

Table 3. Sequential Cox proportional hazards models evaluating the risk of weight change from the prior visit, adjusting for demographics, cardiovascular risk factors, disease characteristics, and disability*

	Model 1 HR (95% CI)	Model 2 HR (95% CI)
Lost ≥10%	1.34 (1.17–1.53)†	1.18 (1.03–1.36)‡
Lost 5–10%	1.09 (0.95–1.25)	1.03 (0.90–1.18)
No change	1 (ref.)	1 (ref.)
Gained 5–10%	1.26 (1.11–1.43)†	1.22 (1.08–1.39)§
Gained ≥10%	1.31 (1.14–1.51)†	1.20 (1.03–1.38)‡

* Model 1 is adjusted for age, sex, race, current body mass index, and calendar year. Model 2 is adjusted as model 1, plus adjustment for diabetes mellitus, hypertension, hyperlipidemia, cardiovascular disease, statin use, smoking, alcohol, reported exercise, disease duration, Clinical Disease Activity Index score, methotrexate use, tumor necrosis factor inhibitor use, prednisone dose, Health Assessment Questionnaire score, and work disability. 95% CI = 95% confidence interval; HR = hazard ratio; ref. = reference.
† $P < 0.001$. ‡ $P < 0.05$. § $P < 0.01$.

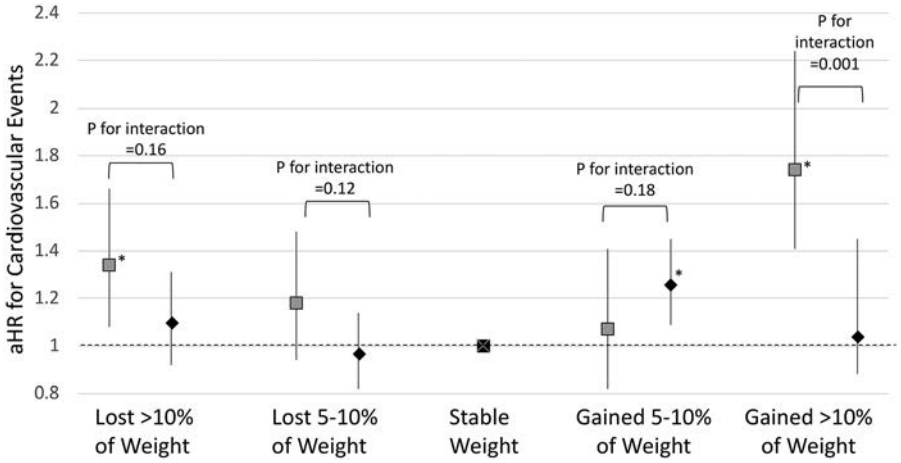

Figure 1. Adjusted hazard ratios (aHR) for the association between weight loss and weight gain from the prior visit and cardiovascular events stratified by current body mass index (BMI) category (overweight/obese [≥25 kg/m^2; diamonds] versus underweight/normal [<25 kg/m^2; shaded squares]). Adjusted for age, sex, race, current BMI, calendar year, diabetes mellitus, hypertension, hyperlipidemia, cardiovascular disease, statin use, smoking, alcohol, reported exercise, disease duration, Clinical Disease Activity Index score, methotrexate use, tumor necrosis factor inhibitor use, prednisone dose, Health Assessment Questionnaire score, and work disability. * = $P < 0.05$ compared to stable weight.

compared to men. Black patients were more likely to experience weight gain compared to White patients, and Asian patients were less likely to lose or gain weight. There were some common predictors of both 10% weight loss and 10% wait gain, shown in Table 2. These included higher CDAI score, prednisone use (≥5 mg), higher disability scores, disability from work, current or former smoking, and abstaining from alcohol use. For example, patients with high CDAI scores had significantly higher risk of both weight loss (odds ratio [OR] 1.18 [95% confidence interval (95% CI) 1.12–1.24], $P < 0.001$) and weight gain (OR 1.36 [95% CI 1.30–1.43], $P < 0.001$). A diagnosis of diabetes mellitus was associated with a greater odds of weight gain. Exercise was associated with a lower odds of weight loss. An improvement in disease activity more than the MCID was associated with a lower likelihood of weight gain (OR 0.95 [95% CI 0.92–0.98], $P = 0.002$) but was not associated with weight loss (OR 1.02 [95% CI 0.99–1.05], $P = 0.32$) over the same time interval after adjustment (full models not shown).

Weight loss and weight gain and associations with CV disease risk. A total of 31,180 participants were included in analyses focused on prediction of CV events. The median follow-up time was 3.4 years (IQR 1.5–6.1 years), and the median time to event was 2.9 years (IQR 1.2–5.6 years). In models adjusting for age, sex, race, BMI, and calendar year, weight gain and weight loss were each associated with a greater risk of CV events (Table 3). The strength of this association was partially attenuated but still significant after the adjustment for CV risk factors, disease characteristics, physical function, and disability. For example, a 10% loss of weight since the prior visit was associated with an 18% greater risk of CV events (hazard ratio [HR] 1.18 [95% CI 1.032–1.36], $P = 0.02$).

In analyses stratified by BMI above and below 25 kg/m^2, the effect of 10% weight loss on the risk of CV events was numerically stronger for thinner participants (HR 1.34 [95% CI 1.08–1.66], $P = 0.001$ [P for interaction = 0.16]) (see Figure 1 and Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24469/abstract). The effect of 10% weight gain was significantly more pronounced among thin participants (HR 1.74 [95% CI 1.41–2.24], $P < 0.001$ [P for interaction = 0.001]). Other factors associated with a greater risk of CV events included older age, male sex, higher disease activity, higher reported M-HAQ score, active smoking, and prednisone use (full model in Supplementary Table 4, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24469/abstract).

Weight variability and associations with CV disease risk. Mean ± SD weight variability increased from quintile 1 to quintile 5 (mean 0.5, 1.2, 1.7, 2.4, 5.4, respectively). Higher quintiles were associated with higher CV risk. In the overall population, higher quintiles were associated with a higher risk compared to the lowest quintile after adjustment (quintile 5 versus quintile 1: HR 1.23 [95% CI 1.06–1.58], $P = 0.005$) (Figure 2). Among underweight and normal-weight patients, there was a significantly greater risk of CV events among those with the greatest weight variability (quintile 5 versus quintile 1: HR 1.42 [95% CI 1.08–1.85], $P = 0.01$). Associations between weight variability and CV events were statistically similar among those above and below 25 kg/m^2 (P for interaction = 0.19).

Sensitivity analyses. Results were similar when limiting analyses to participants without a prior history of CV disease

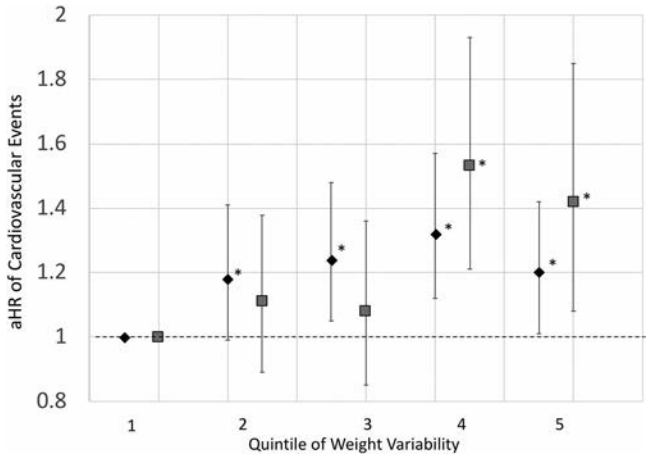

Figure 2. Association between variability in weight over all prior observations (per quintile) and subsequent cardiovascular events. Analysis shown for patients with body mass index (BMI) ≥25 kg/m^2 (diamonds) and for those with BMI <25 kg/m^2 (squares). Adjusted for age, sex, race, current BMI, calendar year, diabetes mellitus, hypertension, hyperlipidemia, cardiovascular disease, statin use, smoking, alcohol, reported exercise, disease duration, Clinical Disease Activity Index score, methotrexate use, tumor necrosis factor inhibitor use, prednisone dose, Health Assessment Questionnaire score, and work disability. * = $P < 0.05$; aHR = adjusted hazard ratio.

(not shown), when adjusting for the number of clinical visits over the prior year, and when excluding participants with <2 months between visits (to reduce the overestimation of change related to small changes over short periods; not shown). When limiting the analysis to MACE events, a significant association was observed only for 10% weight gain among patients with BMI <25 kg/m^2 (HR 2.08 [95% CI 1.40–3.09], $P < 0.001$) (full model can be found in Supplementary Table 5, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24469/abstract).

DISCUSSION

These data support the importance of weight fluctuations in the prediction of CV events in patients with RA. The predictive value of recent weight fluctuation appears to be observed among RA patients who have normal or low BMI, suggesting that fluctuating weight in this group should raise concern. The immediate implication of these results is that risk stratification in RA may benefit from attention to recent weight changes and overall variability of weight over time, particularly among patients with normal to low BMI.

This study identified a number of factors associated with clinically significant weight changes among patients with RA. These data shed additional light on the factors implicated in weight changes over time and causes of weight fluctuation among patients with RA. In this study, weight loss was more likely among older participants, smokers, those with greater disability and comorbidity, those who were more sedentary, those who abstained from alcohol, and those with greater disease activity and use of prednisone. Weight gain was also more likely among those with higher disease activity and greater disability as well as among younger patients, patients with diabetes mellitus, smokers, and those who abstained from alcohol. Overall, these observations emphasize how factors such as RA disease activity, disability, aging, and behavioral factors are likely to contribute significantly to shifts in weight over time. In other words, multiple measures of poor overall health are likely to contribute to weight fluctuations over time, suggesting that these factors likely affect metabolism and/or diet. Because of the relationship between poor health and weight fluctuations, it may be difficult to make inferences about the presence of any causal adverse impact of weight fluctuation itself, as has been proposed elsewhere (2,3). However, cardiometabolic changes have been described in mice that experience weight fluctuation (5), suggesting that the process of weight fluctuation may itself lead to adverse outcomes through adverse effects on metabolism and promotion of cardiometabolic disease (2,3).

The association between recent weight change and CV outcomes was observed primarily among individuals with low and normal BMI. This distinction is important because thin individuals may be overlooked when considering CV risk and are in the greatest need of better risk prediction tools. In other words, substantial weight fluctuation in a thin patient with RA might help to promote a reevaluation of CV risk. The relevance of weight fluctuation has also been shown to be more important among thin patients in studies in other populations, including those with cancer (4,13,19).

It was outside of the scope of the current study to evaluate the incorporation of weight fluctuation in clinical tools to predict CV disease (17). However, the strength of the association in this population is similar to other risk factors observed that are typically utilized to predict long-term risks. Weight fluctuation can be easily measured and visualized in the electronic medical record. The quantification of weight fluctuation could be algorithmicized and provide additional data to clinical providers. Further study may help to quantify the added value of incorporating such information into risk stratification algorithms.

As with many observational studies, the nature of these data does not allow distinguishing intentional and unintentional weight change. Weight fluctuation in this population may reflect intentional changes in diet and exercise or may instead represent unintentional changes as the result of catabolic processes related to the underlying disease or comorbid conditions. It seems more likely that these changes are largely unintentional in nature based on prior studies showing that a greater proportion of weight loss is unintentional among older individuals (20). Unintentional weight loss in other populations has been described to have important implications for long-term health (20–23).

This study demonstrates that fluctuations in weight are predictive of CV events in patients with RA and may aid in risk stratification among normal or underweight patients. However, the current

observational study is unable to determine whether weight fluctuation is a causal mediator or whether it simply represents a marker of adverse processes related to severe disease or comorbidity. While we adjusted for a number of important variables, unmeasured confounding may still be present. There are therefore insufficient data to support the initiation of specific interventions with the aim of preventing weight fluctuation. This study focused on CV disease, a common cause of morbidity and mortality in this population; however, future study should also evaluate the prediction of other important outcomes such as malignancy, disability, as well as CV and overall mortality. Strengths of this study include the large cohort of patients with confirmed RA, the adjudicated definition of CV events, the long-term follow-up, and the detailed collection of important covariables commonly used in the prediction of CV disease risk in addition to disease-specific variables.

In conclusion, weight fluctuation is independently associated with higher CV risk among patients with RA (particularly among patients with low or normal BMI). This association between weight fluctuation and metabolic processes may be partially related to its association with aging, active disease, prednisone use, smoking, disability, and comorbidity. The observation of weight fluctuation in a thin patient with RA should prompt a reevaluation of CV risk.

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. Baker 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. Baker, Kremer.
Acquisition of data. Baker, Reed, Kremer.
Analysis and interpretation of data. Baker, Reed, Kremer.

REFERENCES

1. Oh TJ, Moon JH, Choi SH, Lim S, Park KS, Cho NH, et al. Body-weight fluctuation and incident diabetes mellitus, cardiovascular disease, and mortality: a 16-year prospective cohort study. J Clin Endocrinol Metab 2019;104:639–46.
2. Bangalore S, Fayyad R, DeMicco DA, Colhoun HM, Waters DD. Body weight variability and cardiovascular outcomes in patients with type 2 diabetes mellitus. Circ Cardiovasc Qual Outcomes 2018;11:e004724.
3. Bangalore S, Messerli FH, Waters DD. Body-weight fluctuations and outcomes in coronary disease. N Engl J Med 2017;377:95–6.
4. Montani JP, Schutz Y, Dulloo AG. Dieting and weight cycling as risk factors for cardiometabolic diseases: who is really at risk? Obes Rev 2015;16 Suppl 1:7–18.
5. Schofield SE, Parkinson JR, Henley AB, Sahuri-Arisoylu M, Sanchez-Canon GJ, Bell JD. Metabolic dysfunction following weight cycling in male mice. Int J Obes (Lond) 2017;41:402–11.
6. Mackie GM, Samocha-Bonet D, Tam CS. Does weight cycling promote obesity and metabolic risk factors? Obes Res Clin Pract 2017; 11:131–9.
7. Baker JF, Sauer BC, Cannon GW, Teng CC, Michaud K, Ibrahim S, et al. Changes in body mass related to the initiation of disease-modifying therapies in rheumatoid arthritis. Arthritis Rheumatol 2016;68: 1818–27.
8. Baker JF, Cannon GW, Ibrahim S, Haroldsen C, Caplan L, Mikuls TR. Predictors of longterm changes in body mass index in rheumatoid arthritis. J Rheumatol 2015;42:920–7.
9. Jurgens MS, Jacobs JW, Geenen R, Bossema ER, Bakker MF, Bijlsma JW, et al. Increase of body mass index in a tight controlled methotrexate-based strategy with prednisone in early rheumatoid arthritis: side effect of the prednisone or better control of disease activity? Arthritis Care Res (Hoboken) 2013;65:88–93.
10. Baker JF, Toedter G, Baker DG, von Feldt JM. Treatment-related changes in weight and adipokine levels and associations with radiographic progression in rheumatoid arthritis [abstract]. Arthritis Rheum 2011;63 Suppl:S144–5.
11. Baker JF, Billig E, Michaud K, Ibrahim S, Caplan L, Cannon GW, et al. Weight loss, the obesity paradox, and the risk of death in rheumatoid arthritis. Arthritis Rheumatol 2015;67:1711–7.
12. Baker JF, England BR, Mikuls TR, Sayles H, Cannon GW, Sauer BC, et al. Obesity, weight loss, and progression of disability in rheumatoid arthritis. Arthritis Care Res (Hoboken) 2018;70:1740–7.
13. Vagnildhaug OM, Blum D, Wilcock A, Fayers P, Strasser F, Baracos VE, et al. The applicability of a weight loss grading system in cancer cachexia: a longitudinal analysis. J Cachexia Sarcopenia Muscle 2017;8:789–97.
14. Hsiao PY, Mitchell DC, Wood GC, Jensen GL, Still CD, Hartman TJ. The association of dietary patterns and weight change in rural older adults 75 years and older. J Nutr Gerontol Geriatr 2014;33:357–75.
15. Wong CJ. Involuntary weight loss. Med Clin North Am 2014;98: 625–43.
16. Williamson DA, Bray GA, Ryan DH. Is 5% weight loss a satisfactory criterion to define clinically significant weight loss? Obesity (Silver Spring) 2015;23:2319–20.
17. Solomon DH, Greenberg J, Curtis JR, Liu M, Farkouh ME, Tsao P, et al. Derivation and internal validation of an expanded cardiovascular risk prediction score for rheumatoid arthritis: a Consortium of Rheumatology Researchers of North America Registry study. Arthritis Rheumatol 2015;67:1995–2003.
18. Curtis JR, Yang S, Chen L, Pope JE, Keystone EC, Haraoui B, et al. Determining the minimally important difference in the Clinical Disease Activity Index for improvement and worsening in early rheumatoid arthritis patients. Arthritis Care Res (Hoboken) 2015;67:1345–53.
19. Dulloo AG, Jacquet J, Montani JP. Pathways from weight fluctuations to metabolic diseases: focus on maladaptive thermogenesis during catch-up fat. Int J Obes Relat Metab Disord 2002;26 Suppl 2: S46–57.
20. Wannamethee SG, Shaper AG, Lennon L. Reasons for intentional weight loss, unintentional weight loss, and mortality in older men. Arch Intern Med 2005;165:1035–40.
21. Strandberg TE, Stenholm S, Strandberg AY, Salomaa VV, Pitkala KH, Tilvis RS. The "obesity paradox," frailty, disability, and mortality in older men: a prospective, longitudinal cohort study. Am J Epidemiol 2013;178:1452–60.
22. Campbell KL, MacLaughlin HL. Unintentional weight loss is an independent predictor of mortality in a hemodialysis population. J Ren Nutr 2010;20:414–8.
23. Gaddey HL, Holder K. Unintentional weight loss in older adults. Am Fam Physician 2014;89:718–22.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 2, February 2022, pp 236–242
DOI 10.1002/acr.24448

BRIEF REPORT

Sunlight Exposure, Sun-Protective Behavior, and Anti–Citrullinated Protein Antibody Positivity: A General Population-Based Study in Quebec, Canada

Naizhuo Zhao,[1] Audrey Smargiassi,[2] Ines Colmegna,[1] Marie Hudson,[3] Marvin Fritzler,[4] and Sasha Bernatsky[1]

Objective. To examine associations between sunlight exposure and anti–citrullinated protein antibodies (ACPAs) using general population data in Quebec, Canada.

Methods. A random sample of 7,600 individuals (including 786 subjects who were ACPA positive and 201 self-reported rheumatoid arthritis [RA] cases) from the CARTaGENE cohort was studied cross-sectionally. All subjects were nested in 4 census metropolitan areas, and mixed-effects logistic regression models were used to calculate odds ratios (ORs) and 95% confidence intervals (95% CIs) for ACPA positivity related to sunlight exposure, adjusting for sun-block use, industrial fine particulate matter ($PM_{2.5}$) exposures, smoking, age, sex, French Canadian ancestry, and family income. We also performed sensitivity analyses excluding subjects with RA, defining ACPA positivity by higher titers, and stratifying by age and sex.

Results. The adjusted ORs and 95% CIs did not suggest conclusive associations between ACPA and sunlight exposure or sun-block use, but robust positive relationships were observed between industrial $PM_{2.5}$ emissions and ACPA (OR 1.19 per $\mu g/m^3$ [95% CI 1.03–1.36] in primary analyses).

Conclusion. We did not see clear links between ACPA and sunlight exposure or sun-block use, but we did note positive associations with industrial $PM_{2.5}$. Future studies of sunlight and RA (or ACPA) should take air pollution exposures into account.

INTRODUCTION

There is growing interest in the link between sunlight exposure and the risk of rheumatoid arthritis (RA) (1). Vitamin D synthesis is greatly dependent on ultraviolet B (UVB) rays, and vitamin D insufficiency is a risk factor for RA (2). The higher RA risk in northern versus southern US may be explained by lower UVB exposure (3). Another US study demonstrated an inverse association between UVB light exposure and RA risk, but only included females and assigned a single UVB exposure level to the residents of each state, likely causing significant measurement error (1). Increasing efforts are aimed at reducing sun exposure (e.g., using sun-block) (4); however, these were not considered in previous studies (1,5,6).

Anti-citrullinated protein antibodies (ACPAs) are a characteristic finding in RA and may predate clinical manifestations (7). Despite the interest in sunlight exposure and RA, associations between sunlight exposure and ACPAs have never been examined. Accordingly, we explored for associations between sunlight exposure and ACPA positivity in general population subjects in Quebec, Canada, while controlling for sun-protective behavior.

MATERIALS AND METHODS

Study population and sera samples. The CARTaGENE study (www.cartagene.qc.ca) enrolled 43,000 general population subjects ages 40–69 years as part of the Canadian Partnership for Tomorrow Project, to investigate the health effects of genetics,

Supported by the Canadian Institutes of Health Research (PJT-159682). ACPA assay kits were a gift from Inova Diagnostics, San Diego, California.

[1]Naizhuo Zhao, PhD, Ines Colmegna, MD, Sasha Bernatsky, MD, PhD: McGill University Health Centre, Montreal, Quebec, Canada; [2]Audrey Smargiassi, PhD: Université de Montréal and Institut National de Santé Publique du Québec, Montréal, Quebec, Canada; [3]Marie Hudson, MD: McGill University and Jewish General Hospital, Montréal, Quebec, Canada; [4]Marvin Fritzler, MD, PhD: Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada.

No potential conflicts of interest relevant to this article were reported.

Address correspondence to Sasha Bernatsky, MD, PhD, Centre for Outcomes Research and Evaluation, 5252 Boulevard de Maisonneuve Ouest, 3F.51 Montreal, Quebec, H4A 3S5, Canada. Email: sasha.bernatsky@mcgill.ca.

Submitted for publication April 15, 2020; accepted in revised form September 8, 2020.

SIGNIFICANCE & INNOVATIONS

- Sunlight exposure has been associated with lower rheumatoid arthritis (RA) risk, but no one has ever studied sunlight exposure and anti–citrullinated protein antibodies (ACPAs).
- Controlling for sun-protective behavior and air pollution, we did not detect conclusive associations between sunlight exposure and ACPAs.
- Sun-block use did not correlate with ACPAs, but industrial emissions of fine particulate air pollution were clearly associated with these antibodies.
- Future studies of sunlight exposure and RA/RA-related antibodies should take air pollution exposures into account.

behavior, family history, and environment (8). All Canadian citizens have universal provincial health insurance, and those who resided in the province of Quebec for at least 5 years were randomly selected (from the provincial health insurance registry) and invited to participate in the CARTaGENE cohort. Sociodemographic factors (e.g., age, sex, French Canadian ancestry, family income) and smoking were collected at baseline. These variables are potential confounders or effect modifiers of associations between sunlight and/or other environmental factors (e.g., air pollution) and ACPAs (6,9). Sun-block use (i.e., rarely, sometimes, and often), information also collected in CARTaGENE subjects at baseline, may affect how much UVB reaches the skin and thus was also adjusted for. Baseline CARTaGENE data also include self-report of RA (subjects being asked if they have received an RA diagnoses made by physicians). Consent was obtained from subjects participating in the study prior to conducting the study. The study was reviewed by the McGill University Faculty of Medicine ethics review committee and was given full approval to conduct (#A04-M46-12B). In addition, the CARTaGENE scientific review committee, affiliated with the Centre Hospitalier Universitaire Sainte-Justine (project #582582) provided approval for data and samples to be analyzed.

We chose a random sample (n = 7,600) from CARTaGENE subjects enrolled between 2009 and 2010. Biobanked serum samples were assessed for ACPAs by chemiluminescence immunoassay (CCP3.0; Inova Diagnostics). An ACPA titer of ≥20 units/ml was initially used to define positivity; we conducted sensitivity analyses with titers ≥40 units/ml (10).

Sunlight and industrial particulate matter ($PM_{2.5}$) exposures. The CARTaGENE baseline data set contains 2 categorical variables related to self-reported daily sunlight exposure (i.e., <30 minutes, 30–60 minutes, 1–2 hours, 2–3 hours, 3–4 hours, and >4 hours), 1 for weekdays and 1 for weekends. Given previous work showing associations between industrial $PM_{2.5}$ emissions and ACPAs (9), we calculated total $PM_{2.5}$ emissions from industrial sources within 4 km of each subject's baseline 6-digit postal code, based on National Pollutant Release Inventory data, and a 3-dimensional atmospheric model (11). Fundamental inputs for the atmospheric model come from ground monitoring stations (wind speed, air temperature), and satellite images (terrain) (12).

Statistical analysis. Subjects of the first CARTaGENE recruitment wave were nested in the 4 census metropolitan areas (CMAs) of Quebec, Canada (i.e., Montreal, Quebec City, Sherbrooke, and Saguenay-La-Saint-Jean). In Canada, a CMA is an area consisting of several neighboring municipalities around a major urban core, with a total population of at least 100,000. Though the 4 geographically adjacent CMAs should have similar annual mean UVB radiation and cloud cover (3), to account for spatial differences in average UVB, we used a mixed-effects logistic regression, in which the CMA was set as the random effect and the other variables (i.e., daily sunlight exposure hours for weekdays and weekends, frequency of sun-block use, sex, French Canadian ancestry, smoking, family income level, industrial $PM_{2.5}$ exposure, and age) were treated as fixed effects. Among fixed-effect predictor variables, only $PM_{2.5}$ and age were continuous.

Additional sensitivity analyses were performed sequentially. Since RA patients often have decreased mobility and may remain indoors more than people without RA, we removed all subjects with RA at the time of cohort entry in our first sensitivity analyses and developed another mixed-effects logistic regression model with the same variable settings in the primary analysis. Next, we dichotomized the original daily sunlight exposure variables into 2 categories (i.e., ≤1 hour and >1 hour) and combined the 2 dichotomized categorical variables into 1 with 4 levels (i.e., the 4 possible category combinations of the 2 dichotomized sunlight exposure variables for weekdays and weekends) to indicate an individual's overall weekly sunlight exposure. We used 1 hour as the cut point to ensure the divided 2 groups having the closest numbers of subjects for either weekdays or weekends. We carried out a second sensitivity analysis, in which the 2 original sunlight exposure variables were replaced by the overall sunlight exposure variable, patients with RA were excluded, and the covariates in the primary analysis were maintained. In the third sensitivity analysis, the threshold of ACPA positivity was increased to 40 units/ml, and the other variables were maintained the same as with the second sensitivity analysis. In the fourth sensitivity analysis, we followed the methods of a previous UVB and RA study to stratify the subjects by age (i.e., age <52 and ≥52 years) (1), using the 20 units/ml threshold and excluding subjects with RA. To further compare with previous studies (1,6) that only included females, we conducted another sensitivity analysis that stratified subjects by sex. Finally, we repeated the above sensitivity analyses adjusting for sun exposure and sunblock use but not industrial $PM_{2.5}$ exposure.

Table 1. Adjusted odds ratios for associations between sunlight exposures and ACPA positivity (≥20 units/ml)*

Variables	Overall	Excluding RA
Sunlight exposure for weekdays		
<30 minutes	Ref.	Ref.
30–60 minutes	0.89 (0.70–1.12)	0.85 (0.67–1.07)
1–2 hours	1.08 (0.83–1.41)	1.04 (0.79–1.36)
2–3 hours	0.87 (0.62–1.21)	0.79 (0.56–1.12)
3–4 hours	0.83 (0.53–1.31)	0.83 (0.53–1.31)
>4 hours	0.80 (0.48–1.35)	0.76 (0.45–1.29)
Sunlight exposure for weekends		
<30 minutes	Ref.	Ref.
30–60 minutes	1.03 (0.78–1.36)	1.04 (0.78–1.39)
1–2 hours	0.85 (0.63–1.14)	0.85 (0.63–1.15)
2–3 hours	1.01 (0.74–1.38)	1.04 (0.76–1.44)
3–4 hours	0.99 (0.68–1.43)	1.03 (0.71–1.49)
>4 hours	1.22 (0.82–1.82)	1.29 (0.86–1.93)
Sun-block use		
Rarely	Ref.	Ref.
Sometimes	0.88 (0.70–1.12)	0.89 (0.70–1.12)
Often	1.00 (0.83–1.21)	0.98 (0.81–1.19)
Industrial $PM_{2.5}$ exposure†	1.19 (1.03–1.36)	1.19 (1.03–1.36)
Age	1.01 (1.00–1.02)	1.01 (1.00–1.03)
Sex		
Male	Ref.	Ref.
Female	1.01 (0.85–1.20)	0.98 (0.82–1.16)
Ancestry		
French Canadian	Ref.	Ref.
Other	0.98 (0.83–1.17)	1.05 (0.88–1.25)
Smoking		
Never	Ref.	Ref.
Occasional	0.96 (0.74–1.23)	0.93 (0.72–1.21)
Daily	1.09 (0.85–1.42)	1.08 (0.83–1.41)
Annual income level, Canadian $		
<25,000	Ref.	Ref.
25,000–49,999	0.98 (0.72–1.33)	1.00 (0.72–1.37)
50,000–74,999	0.99 (0.73–1.34)	1.03 (0.75–1.41)
75,000–149,999	0.93 (0.69–1.25)	0.95 (0.70–1.30)
≥150,000	1.02 (0.72–1.44)	1.07 (0.74–1.53)

* Values are the odds ratio (95% confidence interval). ACPA = anti–citrullinated protein antibody; $PM_{2.5}$ = particulate matter; RA = rheumatoid arthritis; Ref. = reference.
† Odds ratios reported per increase in 1 μg/m^3, a value well above median levels.

RESULTS

Of the 7,600 subjects, 76.2% lived in Montreal, 15.3% were residents of Quebec City, 4.3% were from Saguenay-La-Saint-Jean, and 4.1% lived in Sherbrooke. Mean ± SD age was 54.1 ± 7.7 years, 50.1% were female, and French Canadians made up 68.2% of subjects. Many subjects reported that they often (49.1%) or sometimes (21.2%) used sun block, while the remaining (29.7%) rarely did. Two-fifths of the subjects (40.4%) never smoked, 14.3% smoked daily, and the remaining smoked occasionally. Approximately one-tenth of subjects (9.8%) lived below the lowest family income level (for details, see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24448/abstract), 11.5% belonged to the highest level for family incomes, and the remaining subjects fell between these ranges.

Among the 7,600 subjects, 786 had ACPA levels above 20 units/ml, and 292 had a titer above 40 units/ml. A total of 201 individuals in our sample reported physician-diagnosed RA when they entered the cohort, and 37 subjects had both RA and ACPA positivity. Only 35.5% of subjects were exposed to sunlight beyond 1 hour every weekday, and 62.7% had >1 hour of daily sunlight exposure on weekends. Industrial $PM_{2.5}$ concentration estimates across Quebec for 2005–2010 varied from 0.03 to 14.09 μg/m^3 (mean ± SD 0.21 ± 0.40 μg/m^3). More detailed comparisons of baseline information between ACPA-positive and ACPA-negative subjects are shown in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24448/abstract.

The adjusted odds ratios (ORs) and corresponding 95% confidence intervals (95% CIs) did not suggest conclusive relationships between sunlight exposure (treated as either 1 or 2 categorical variables) and ACPAs at either threshold (Tables 1 and 2). As expected, ACPA positivity was more common with older age. However, no clear association between sunlight exposure and

Table 2. Adjusted odds ratios for associations between the overall sunlight exposure and ACPA positivity defined by different thresholds*

	20 units/ml (n = 7,289; +ACPA = 740)	40 units/ml (n = 6,813; +ACPA = 264)
Overall sunlight exposure		
Weekday: ≤1 hour; weekend: ≤1 hour	Ref.	Ref.
Weekday: ≤1 hour; weekend: >1 hour	0.88 (0.71–1.09)	1.06 (0.76–1.49)
Weekday: >1 hour; weekend: ≤1 hour	0.78 (0.36–1.72)	0.99 (0.31–2.22)
Weekday: >1 hour; weekend: >1 hour	0.96 (0.79–1.17)	1.18 (0.87–1.60)
Sun-block use		
Rarely	Ref.	Ref.
Sometimes	0.88 (0.70–1.12)	0.87 (0.60–1.26)
Often	1.00 (0.83–1.21)	0.83 (0.61–1.14)
Industrial $PM_{2.5}$ exposure†	1.18 (1.03–1.36)	1.23 (1.04–1.45)
Age	1.01 (1.00–1.02)	1.01 (0.99–1.03)
Sex		
Male	Ref.	Ref.
Female	1.02 (0.86–1.20)	1.25 (0.95–1.65)
Ancestry		
French Canadian	Ref.	Ref.
Other	1.04 (0.87–1.24)	1.06 (0.80–1.41)
Smoking		
Never	Ref.	Ref.
Occasional	0.93 (0.72–1.21)	0.83 (0.56–1.24)
Daily	1.08 (0.83–1.40)	0.81 (0.54–1.21)
Annual income level, Canadian $		
<25,000	Ref.	Ref.
25,000–49,999	1.00 (0.73–1.37)	1.02 (0.61–1.66)
50,000–74,999	1.03 (0.75–1.41)	0.98 (0.59–1.61)
75,000–149,999	0.96 (0.71–1.30)	0.97 (0.60–1.56)
≥150,000	1.09 (0.77–1.55)	1.07 (0.60–1.89)

* Values are the odds ratio (95% confidence interval). Subjects with rheumatoid arthritis were removed from the regressions. When the 40 units/ml threshold was used to define anti–citrullinated protein antibody (ACPA) positivity, subjects with ACPA ≥20 units/ml but <40 units/ml were excluded from the analysis. $PM_{2.5}$ = particulate matter; Ref. = reference.
† Odds ratios reported per increase in 1 μg/m^3, a value well above median levels.

ACPAs was seen for any age/sex group (Table 3). No clear associations between sun-block use and ACPAs were seen in any of our models. However, significant correlations between industrial $PM_{2.5}$ and ACPA positivity were observed in all of our analyses, although analyses limited to older adults lacked precision. In addition, daily smoking was associated with increased ACPA positivity in women (Table 3 and Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24448/abstract). Without adjusting for industrial $PM_{2.5}$, conclusive relationships between sunlight exposure/sun-block use and ACPAs were not observed either, but most of the adjusted ORs were slightly larger than those with adjusting for the covariate (see Supplementary Tables 2 and 3, at http://onlinelibrary.wiley.com/doi/10.1002/acr.24448/abstract).

DISCUSSION

Our study is the first to assess sunlight exposure and ACPA positivity. We used more sophisticated approaches than prior studies of sunlight and RA. Compared to those studies (1,5,6), we took additional sun-block use and air pollution into account. In a prior study (9), a simple metric of distance to major industrial emitters was used to assess the association between industrial air pollution exposure and ACPA positivity, while in the current study, we used a 3-dimensional atmospheric model (i.e., CALPUFF), which accounts for the effects of spatiotemporally varying meteorologic conditions on transport, transformation, and dissipation of $PM_{2.5}$. Thus, we assigned more accurate estimates of industrial $PM_{2.5}$ exposures in our current study than in our previous work (12). In the study by Arkema et al (1), all subjects living in a given state were assigned the same mean annual exposure, although UVB radiance levels may vary considerably across the region of a state. Moreover, residence in a state with a higher UVB radiance level does not necessarily mean that an individual has a higher UVB exposure than people living in another state, since UVB exposure relates also to the amount of time spent in outdoor activities. By contrast, in our study, we assigned sunlight exposure levels based on self-reported daily exposure hours and reduced the spatial biases of average UVB radiance by setting CMA as a random effect in our models. Additionally, previous population-based studies (1,6) only included females, while we studied both sexes.

In our sample, 2.6% of subjects reported a physician diagnosis of RA before they entered the cohort. These individuals possibly had less sun exposure due to RA, decreasing their ability to

Table 3. Adjusted odds ratio from the mixed-effects logistic regression models for the associations between the overall sunlight exposure and ACPA positivity defined by the 20 units/ml threshold using subsamples stratified by age or sex*

	Stratified by age		Stratified by sex	
	Age < 52 years (n = 3,269; +ACPA = 316)	Age ≥ 52 years (n = 4,020; +ACPA = 424)	Male (n = 3,619; +ACPA = 363)	Female (n = 3,685; +ACPA = 377)
Overall sunlight exposure				
Weekday: ≤1 hour; weekend: ≤1 hour	Ref.	Ref.	Ref.	Ref.
Weekday: ≤1 hour; weekend: >1 hour	1.11 (0.81–1.53)	0.94 (0.68–1.13)	0.75 (0.54–1.53)	1.07 (0.80–1.53)
Weekday: >1 hour; weekend: ≤1 hour	0.43 (0.16–2.21)	1.05 (0.44–2.51)	0.70 (0.21–2.33)	1.01 (0.35–2.91)
Weekday: >1 hour; weekend: >1 hour	1.06 (0.75–1.48)	0.90 (0.71–1.16)	1.01 (0.77–1.34)	0.86 (0.64–1.15)
Sun-block use				
Never	Ref.	Ref.	Ref.	Ref.
Sometimes	0.86 (0.61–1.23)	0.90 (0.65–1.24)	0.76 (0.54–1.06)	1.04 (0.73–1.47)
Often	0.87 (0.64–1.17)	1.08 (0.84–1.40)	1.11 (0.86–1.45)	0.88 (0.66–1.17)
Industrial $PM_{2.5}$ exposure†	1.23 (1.01–1.51)	1.12 (0.91–1.39)	1.23 (1.02–1.48)	1.14 (1.00–1.42)
Age	–	–	1.01 (0.98–1.02)	1.02 (1.00–1.03)
Sex				
Male	Ref.	Ref.	–	–
Female	0.87 (0.67–1.14)	1.08 (0.86–1.36)	–	–
Ancestry				
French Canadian	Ref.	Ref.	Ref.	Ref.
Other	0.90 (0.68–1.19)	1.17 (0.93–1.47)	1.00 (0.78–1.28)	1.10 (0.85–1.41)
Smoking				
Never	Ref.	Ref.	Ref.	Ref.
Occasional	0.97 (0.65–1.46)	0.90 (0.64–1.27)	0.74 (0.52–1.05)	1.28 (0.85–1.93)
Daily	1.25 (0.85–1.86)	0.94 (0.66–1.34)	0.79 (0.55–1.12)	1.57 (1.04–2.37)
Annual income level, Canadian $				
<25,000	Ref.	Ref.	Ref.	Ref.
25,000–49,999	1.17 (0.68–2.01)	0.92 (0.62–1.35)	1.02 (0.62–1.67)	0.99 (0.65–1.49)
50,000–74,999	0.88 (0.51–1.52)	1.11 (0.76–1.63)	1.20 (0.74–1.95)	0.91 (0.60–1.38)
75,000–149,999	1.10 (0.67–1.82)	0.84 (0.57–1.24)	1.08 (0.68–1.73)	0.91 (0.60–1.37)
≥150,000	1.32 (0.76–2.29)	0.89 (0.55–1.43)	1.42 (0.84–2.40)	0.90 (0.54–1.49)

* Values are the odds ratio (95% confidence interval). Subjects with rheumatoid arthritis were removed from the regressions. ACPA = anti–citrullinated protein antibody; $PM_{2.5}$ = particulate matter; Ref. = reference.
† Odds ratio reported per increase in 1 $\mu g/m^3$, a value well above median levels.

participate in occupations (e.g., mail delivery) or activities (e.g., hiking) associated with sun exposure. After excluding subjects with RA from our multivariate analyses, clear associations between ACPAs and sunlight exposure and sun-block use were still not seen. This result does not necessarily mean that our findings are contradictory with previous studies of sunlight exposure and RA (1,5,6), because although ACPA positivity is a specific marker of RA, some RA patients may not have this antibody (particularly when assessed only once), while conversely, in some cases the antibody appears well before disease onset (13). Exposure to sunlight may be tied to exposure to air pollution. The significant relationship between ACPA positivity and industrial $PM_{2.5}$ exposure suggests the need to adjust for air pollution in future studies of sunlight exposure and RA or ACPAs. We were unable to establish a clear relationship between ACPA positivity and industrial $PM_{2.5}$ exposure in the older age group. The diminished sample size after stratifying the subjects by age resulted in low power and relatively wide CIs, which included the null value.

The overall frequency of ACPAs in our sample is relatively high compared with other studies of non-RA subjects but in part this frequency may be due to the fact that age and comorbidity are associated with autoantibody positivity in the absence of RA (14). All of our subjects were age ≥40 years (with well over half age ≥52 years), and many CARTAGENE participants have a comorbidity (8).

The 2 variables regarding sunlight exposure in the CARTaGENE cohort are self-reported, and thus a few participants may not accurately estimate their daily sunlight exposure hours on weekdays and weekends. An additional limitation of our study is that we were unable to completely adjust for race/ethnicity, since we only had information on French (i.e., European) ancestry. Previous studies have demonstrated that UVB absorption is higher in Caucasians (15), which would include those of French ancestry, but also some subjects of non-French ancestry. For people of the same race/ethnicity, their skin types are also likely to be different. Moreover, a very small portion of the participants may have experienced severe sunburns before they entered the cohort. Different skin type and severely sunburned skin may influence UVB absorption and sun behaviors of the participants. Additionally, the CARTaGENE questionnaire only collected the frequency of using sun protection in the summer when participants were in the sun for ≥30 minutes, but the questionnaire lacked more detailed information on sun protection, such as the amount of sun-block used and how much exposed skin was protected by sun block. These incomplete ascertainments in our sample may have prevented us from detecting effect modification related to this variable on ACPAs. Thus, additional studies assessing distinct race/ethnicity groups and skin types would be helpful to compare with and/or reinforce the current findings.

In conclusion, although a few studies have reported that sunlight exposure is associated with a lower RA risk, we did not see any clear associations of sunlight exposure (or sun-block use) with ACPAs. Significant positive relationships between industrial $PM_{2.5}$ emissions and ACPAs were observed. Additional work is needed to follow-up ACPA-positive subjects without RA, to establish whether environmental factors (like air pollution, sunlight exposure, and other variables) could alter later risk of RA development. In addition, future studies of sunlight exposure and RA (or RA-related antibodies) should consider taking air pollution exposures into account.

ACKNOWLEDGMENTS

We thank Dr. Marianne Hatzopoulou at the University of Toronto for directing the development of the CALPUFF estimates. We also acknowledge the laboratory technical efforts of Ms. Haiyan Hou and Ms. Natalia Baeza at Mitogen Advanced Diagnostics (Calgary, Alberta, Canada) and the helpful comments and suggestions from 3 anonymous reviewers.

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. Bernatsky 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. Zhao, Smargiassi, Colmegna, Hudson, Bernatsky.

Acquisition of data. Smargiassi, Fritzler, Bernatsky.

Analysis and interpretation of data. Zhao, Bernatsky.

ROLE OF THE STUDY SPONSOR

Inova Diagnostics 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 Inova Diagnostics.

REFERENCES

1. Arkema EV, Hart JE, Bertrand KA, Laden F, Grodstein F, Rosner BA, et al. Exposure to ultraviolet-B and risk of developing rheumatoid arthritis among women in the Nurses' Health Study. Ann Rheum Dis 2013;72:506–11.
2. Cutolo M, Otsa K, Uprus M, Paolino S, Seriolo B. Vitamin D in rheumatoid arthritis. Autoimmun Rev 2007;7:59–64.
3. Vieira VM, Hart JE, Webster TF, Weinberg J, Puett R, Laden F, et al. Association between residences in U.S. northern latitudes and rheumatoid arthritis: a spatial analysis of the Nurses' Health Study. Environ Health Perspect 2010;118:957–61.
4. Hoel DG, Berwick M, de Gruijl FR, Holick MF. The risks and benefits of sun exposure 2016. Dermatoendocrinol 2016;8:e1248325.
5. Ding C, Cicuttini F, Parameswaran V, Burgess J, Quinn S, Jones G. Serum levels of vitamin D, sunlight exposure, and knee cartilage loss in older adults: the Tasmanian older adult cohort study. Arthritis Rheum 2009;60:1381–9.
6. Merlino LA, Curtis J, Mikuls TR, Cerhan JR, Criswell LA, Saag KG. Vitamin D intake is inversely associated with rheumatoid arthritis: results from the Iowa Women's Health Study. Arthritis Rheum 2004;50:72–7.
7. Puszczewicz M, Iwaszkiewicz C. Role of anti-citrullinated protein antibodies in diagnosis and prognosis of rheumatoid arthritis. Arch Med Sci 2011;7:189–94.
8. Awadalla P, Boileau C, Payette Yves, Idaghdour Y, Goulet JP, Knoppers B, et al. Cohort profile of the CARTaGENE study: Quebec's

population-based biobank for public health and personalized genomics. Int J Epidemiol 2013;42:1285–99.

9. Bernatsky S, Smargiassi A, Joseph L, Awadalla P, Colmegna I, Hudson M, et al. Industrial air emissions, and proximity to major industrial emitters, are associated with anti-citrullinated protein antibodies. Environ Res 2017;157:60–63.
10. Quest Diagnostics. Cyclic citrullinated peptide (CCP) antibody (IgG). Test Center 2019. URL: https://testdirectory.questdiagnostics.com/test/test-detail/11173/cyclic-citrullinated-peptide-ccp-antibody-igg?p=r&q=CCP&cc=MASTER.
11. Exponent. CALPUFF modeling system. URL: http://www.src.com/.
12. Buteau S, Shekarrizfard M, Hatzopoulou M, Gamache P, Liu L, Smargiassi A. Air pollution from industries and asthma onset in childhood: a population-based birth cohort study using dispersion modeling. Environ Res 2020;185:109180.
13. Grootenboer-Mignot S, Nicaise-Roland P, Delaunay C, Meyer O, Chollet-Martin S, Labarre C. Second generation anti-cyclic citrullinated peptide (anti-CCP2) antibodies can replace other anti-filaggrin antibodies and improve rheumatoid arthritis diagnosis. Scand J Rheumatol 2004;33:218–20.
14. Aggarwal R, Liao K, Nair R, Ringold S, Costenbander KH. Anti–citrullinated peptide antibody assays and their role in the diagnosis of rheumatoid arthritis. Arthritis Care Res (Hoboken) 2009, 61: 1472–83.
15. Al-Jamal MS, Griffith JL, Lim HW. Photoprotection in ethnic skin. Dermatologica Sinica 2014;32:217–24.

Arthritis Care & Research
Vol. 74, No. 2, February 2022, pp 243–250
DOI 10.1002/acr.24449

American College of Rheumatology
Empowering Rheumatology Professionals

Inflammation of the Sacroiliac Joints and Spine and Structural Changes on Magnetic Resonance Imaging in Axial Spondyloarthritis: Five-Year Data From the DESIR Cohort

Alexandre Sepriano,[1] Sofia Ramiro,[2] Robert Landewé,[3] Anna Moltó,[4] Pascal Claudepierre,[5] Daniel Wendling,[6] Maxime Dougados,[7] and Désirée van der Heijde[8]

Objective. To test the impact of inflammation on structural changes occurring in the sacroiliac (SI) joints and the spine detected on magnetic resonance imaging (MRI).

Methods. Patients with early axial spondyloarthritis (SpA) from the Devenir des Spondylarthropathies Indifférenciées Récentes (DESIR) cohort were included. MRIs of the SI joints (MRI-SI joints) and spine (MRI-spine), obtained at baseline, 2 years, and 5 years, were scored by 3 central readers. Inflammation and structural damage on MRI-SI joints and MRI-spine were defined by the agreement of ≥2 of 3 readers (binary outcomes) and by the average of 3 readers (continuous outcomes). The effect of inflammation (MRI-SI joints/MRI-spine) on damage (MRI-SI joints/MRI-spine, respectively) was evaluated in 2 models: 1) a baseline prediction model (the effect of baseline inflammation on damage assessed at 5 years); and 2) a longitudinal model (the effect of inflammation on structural damage assessed during a 5-year period).

Results. A total of 202 patients were included. Both the presence of bone marrow edema on MRI-SI joints and on MRI-spine at baseline were predictive of 5-year damage (≥3 fatty lesions) on MRI-SI joints (odds ratio [OR] 4.2 [95% confidence interval (95% CI) 2.4, 7.3]) and MRI-spine (OR 10.7 [95% CI 2.4, 49.0]), respectively, when adjusted for C-reactive protein level. The association was also confirmed in longitudinal models (when adjusted for Ankylosing Spondylitis Disease Activity Score) both in the SI joints (OR 5.1 [95% CI 2.7, 9.6]) and spine (OR 15.6 [95% CI 4.8, 50.3]). Analysis of other structural outcomes (i.e., erosions) on MRI-SI joints yielded similar results. In the spine, a significant association was found for fatty lesions but not for erosions and bone spurs, which occurred infrequently over time.

Conclusion. We found a predictive and longitudinal association between inflammation detected on MRI and several types of structural damage detected on MRI in patients with early axial SpA, which adds to the evidence for a causal relationship.

INTRODUCTION

Axial spondyloarthritis (SpA) is a disease predominantly characterized by involvement of the axial skeleton. Axial involvement often translates into imaging abnormalities, which usually represent either an underlying inflammatory or structural lesion. Magnetic resonance imaging (MRI) of the sacroiliac (SI) joints (MRI-SI joints) and spine (MRI-spine) is a modality to detect, quantify, and evaluate (change of) axial inflammation in axial SpA. Thus far, conventional radiographs have been prescribed

The DESIR cohort is financially supported by an unrestricted grant from Pfizer and is conducted under the control of Assistance publique Hopitaux de Paris via the Clinical Research Unit Paris Centre and under the umbrella of the French Society of Rheumatology and Institut national de la sante et de la recherche medicale (INSERM). Database management is performed within the Department of Epidemiology and Biostatistics. Dr. Sepriano's work was supported by the Fundação para a Ciência e Tecnologia (doctoral grant SFRH/BD/108246/2015).

[1]Alexandre Sepriano, MD: Leiden University Medical Center, Leiden, The Netherlands, and NOVA Medical School, Universidade Nova de Lisboa, Lisbon, Portugal; [2]Sofia Ramiro, MD, PhD: Leiden University Medical Center, Leiden, and Zuyderland Medical Center, Heerlen, The Netherlands; [3]Robert Landewé, MD, PhD: Amsterdam University Medical Center, Amsterdam, and Zuyderland Medical Center, Heerlen, The Netherlands; [4]Anna Moltó, MD, PhD: Hôpital Cochin, Assistance Publique, Hôpitaux de Paris, Université de Paris, INSERM (U1153), CRESS, Paris, France; [5]Pascal Claudepierre, MD: Hôpital Henri-Mondor, Assistance Publique-Hôpitaux de Paris Créteil, Paris, France; [6]Daniel Wendling, MD, PhD: CHRU de Besançon, University Teaching Hospital, and Université de Franche-Comté, EA4266 EPILAB, Besançon, France; [7]Maxime Dougados, MD: Paris Descartes University, Hôpital Cochin, Assistance Publique, Hôpitaux de Paris, INSERM (U1153), and Clinical Epidemiology and Biostatistics, PRES Sorbonne Paris-Cité, Paris, France; [8]Désirée van der Heijde, MD, PhD: Leiden University Medical Center, Leiden, The Netherlands.

No potential conflicts of interest relevant to this article were reported.

Address correspondence to Alexandre Sepriano, MD, Leiden University Medical Center, Department of Rheumatology, Albinusdreef 2, PO Box 9600, 2300 RC Leiden, The Netherlands. Email: alexsepriano@gmail.com.

Submitted for publication March 30, 2020; accepted in revised form September 8, 2020.

SIGNIFICANCE & INNOVATIONS

- There is a predictive and longitudinal association between inflammation detected on magnetic resonance imaging (MRI) and the development of structural damage on MRI in the sacroiliac (SI) joints (fatty lesions and erosions) and spine (fatty lesions) over 5 years in early axial spondyloarthritis (SpA).
- The association between inflammation and damage is detected with more precision in the SI joints where, compared to the spine, structural damage prevails in early disease.
- This is the first time that a relationship is proven between inflammation and damage when both are assessed on MRI, confirming the known relationship between inflammation and structural damage on radiographs in axial SpA.
- These findings suggest that MRI, especially of the SI joints, is a valid alternative to conventional radiographs in detecting the structural consequences of axial inflammation in patients with early axial SpA.

for assessing progression of structural damage in clinical practice and research.

Patients with axial SpA experience varying levels of radiographic progression (e.g., the occurrence of radiographic sacroiliitis and new syndesmophytes) (1–4). Identifying patients with a higher likelihood of damage accrual is key to tailoring treatment strategies early in the disease course. Elevation of C-reactive protein (CRP) level, disease activity as measured with the Ankylosing Spondylitis Disease Activity Score (ASDAS), and bone marrow edema (BME) on MRI-SI joints or MRI-spine have been shown to associate with increased probability of structural progression on conventional radiographs (3,5–12). Evidence is scarce, however, in early disease and mostly limited to studies in which structural damage was measured with conventional radiographs.

The interpretation of data stemming from the above-mentioned studies may be jeopardized by limitations of the instruments used to measure structural progression, especially at the level of the SI joints. It is well established that radiographic sacroiliitis defined by the modified New York criteria (mNY) is poorly reliable (13–15). Investigators have been implementing strategies to improve the signal-to-noise ratio by, for instance, combining judgments from ≥2 trained central readers (3). Still, these strategies cannot fully eliminate the noise.

In recent years, there has been a growing interest in evaluating axial damage with other imaging modalities, such as MRI. Definitions for individual lesions (e.g., fatty lesions, erosions) have been proposed, and composite scores validated (16–19). Although MRI-detected lesions, as any outcome measure, are far from being error free, available literature shows higher reliability for MRI-SI joints compared to pelvic radiographs in detecting structural lesions (20). A better signal-to-noise ratio, in theory, improves the ability to detect change and predictors thereof, especially in early disease where, at the group level, damage is known to be limited and to progress slowly (3,21).

Thus far, no study has assessed the effect of inflammation on structural damage evaluated on MRI. We aimed to test the effect of inflammation on several types of structural lesions both assessed by MRI and at the level of the SI joints and the spine in patients with early axial SpA.

PATIENTS AND METHODS

Patients and study design. Five-year data from patients with early axial SpA from the Devenir des Spondylarthropathies Indifférenciées Récentes (DESIR) cohort have been used (22). Patients had to have ≥2 consecutive MRI images (either of the SI joints or spine) during the 5-year follow-up period to be included. The database used for the current analysis was locked on June 20, 2016. The study was conducted according to Good Clinical Practice guidelines and was approved by the appropriate local ethics committees. Written informed consent had been obtained from participating patients before inclusion.

Imaging scoring procedures. MRI-SI joints and MRI-spine were performed at baseline for all patients. By protocol, at 2 and 5 years of follow-up, MRIs were only performed in participating centers in Paris (n = 9 of the 25 participating centers). Each image was independently scored by 3 trained central readers blinded to chronology and clinical data. MRI-SI joints and MRI-spine were performed on a 1.0–1.5T scanner providing T1-weighted turbo spin-echo and short tau inversion recovery sequences. Scanning was performed in a coronal oblique plane for the SI joints and in a sagittal plane for the spine, with a slice thickness of 4 mm. A detailed description of the MRI protocol in the DESIR cohort has been reported previously (23,24).

Structural damage on MRI. The Spondyloarthritis Research Consortium of Canada (SPARCC) MRI-SI joints structural score by Weber et al was used to define individual structural lesions on MRI-SI joints (18). In the absence of a formal definition for structural damage on MRI-SI joints, we considered 3 definitions previously shown most discriminatory between axial SpA and no axial SpA: ≥5 fatty lesions and/or erosions; ≥3 erosions; and ≥3 fatty lesions (25). Continuous structural lesions on MRI-SI joints were defined as number of erosions, number of fatty lesions (range of both 0–40), number of fatty lesions and/or erosions (range 0–80), and as the total number of lesions including fatty lesions, erosions, and partial ankylosis/total ankylosis with the addition of sclerosis (not in the original score) (range 0–144).

Structural lesions on MRI-spine were scored according to the Canada–Denmark (CANDEN) method, modified to include only corner lesions (16,17). Similar to MRI-SI joints, in the absence of a formal definition, we defined structural damage on MRI-spine

as ≥5 fatty lesions, which has been previously shown highly specific for axial SpA (25,26). In addition, we also considered ≥5 fatty lesions and/or erosions; ≥3 erosions; ≥3 fatty lesions; and ≥3 bone spurs. The total number of fatty lesions, erosions, bone spurs (range 0–92 for each), fatty lesions and/or erosions (range 0–184), and the total number of structural lesions (fatty lesions, erosions, bone spurs, including also ankylosis; range 0–322) was assessed as continuous structural outcomes.

Inflammation on MRI. Inflammation on MRI-SI joints was assessed using the Assessment of SpondyloArthritis international Society (ASAS) definition (positive/negative) and the SPARCC score (range 0–72) (27–29). BME on MRI-spine was defined according to the ASAS definition (≥3 vertebral corner lesions; positive/negative) (30). In addition, a cutoff of at least 5 lesions was assessed, as it has been shown to be highly specific of axial SpA (25). The total spine SPARCC score was used as a continuous inflammatory outcome (range 0–414) (31). The interreader reliability of the MRI scores used in this study has been reported elsewhere (32).

Table 1. Baseline patient and disease characteristics comparing patients with magnetic resonance imaging (MRI) results available for ≥2 consecutive (included) visits to those without (excluded)*

Characteristic	MRI on ≥2 consecutive visits (n = 202)	MRI on <2 consecutive visits (n = 60)
Age at baseline, mean ± SD years	34 ± 9	33 ± 8
Male sex	96 (48)	27 (45)
Symptom duration, mean ± SD years	2 ± 1	1 ± 1
HLA–B27	125 (62)	32 (53)
ASAS axial SpA criteria	133 (66)	35 (60)
Sacroiliitis on MRI-SI joints (ASAS)†	58 (29)	15 (28)
BME on MRI-spine (ASAS)†	14 (7)	3 (6)
≥5 BME lesions on MRI-spine	10 (5)	2 (4)
Radiographic sacroiliitis (mNY)†	25 (13)	8 (14)
≥3 fatty lesions on MRI-SI joints	23 (12)	7 (14)
≥3 erosions on MRI-SI joints	29 (15)	9 (17)
≥3 fatty lesions on MRI-spine	3 (2)	0 (0)
≥3 erosions on MRI-spine	0 (0)	0 (0)
≥3 bone spurs on MRI-spine	0 (0)	0 (0)
BASDAI score, mean ± SD (range 0–10)	4 ± 2	47 ± 21
ASDAS-CRP score, mean ± SD	3 ± 1	3 ± 1
Elevated CRP (≥6 mg/liter)	52 (27)	12 (21)
BASFI score, mean ± SD (range 0–10)	3 ± 2	33 ± 28
Treatment with NSAIDs	192 (95)	57 (95)
Treatment with TNFi	0 (0)	0 (0)

* Values are the number (%) unless indicated otherwise. The following variables had <5% missing data: radiographic sacroiliitis (mNY), bone marrow edema (BME) on MRI-spine (ASAS), ≥5 BME lesions on MRI-spine, ≥3 fatty lesions on MRI-spine, ≥3 erosions on MRI-spine, ≥3 bone spurs on MRI-spine, and ASDAS-CRP score. The following categories had <1% missing data: sacroiliitis on MRI-SI joints (ASAS), ≥3 fatty lesions on MRI-SI joints, ≥3 erosions on MRI-SI joints, BASDAI score, and BASFI score. ASAS = Assessment of SpondyloArthritis international Society; ASDAS = Ankylosing Spondylitis Disease Activity Score; BASDAI = Bath Ankylosing Spondylitis Disease Activity Index; BASFI = Bath Ankylosing Spondylitis Functional Index; CRP = C-reactive protein; mNY = modified New York criteria for radiographic sacroiliitis; MRI-SI joints = MRI of the sacroiliac joints; MRI-spine = MRI of the spine; NSAIDs = nonsteroidal antiinflammatory drugs; SpA = spondyloarthritis; TNFi = tumor necrosis factor inhibitors.
† Agreement between 2 of 3 readers.

Statistical analysis. Structural progression of binary scores was assessed in clinically relevant subgroups according to the CRP level and BME status at baseline and defined by the agreement of ≥2 of 3 readers as the percentage of net progression: the number of progressors (change from negative to positive) minus the number of regressors (change from positive to negative) divided by the total number of patients, a method previously described in detail (33).

The effect of inflammation, both on MRI-SI joints and MRI-spine, on structural outcomes, again both on MRI-SI joints and MRI-spine, respectively, was evaluated by 2 types of generalized estimating equation (GEE) models: 1) a baseline model (the effect of baseline inflammation on 5 years of structural damage incorporating measurements from all readers [1-level GEE model adjusted for the reader]); and 2) a longitudinal model (the effect of BME at t on structural outcomes at $t + 1$ over 5 years [longitudinal time-lagged, 2-level GEE models with autoregression]). Binary variables of inflammation (i.e., BME) were modeled using binary damage outcomes (binomial GEE), while continuous variables of inflammation (i.e., SPARCC score) were modeled using continuous outcomes of damage (linear GEE).

The final multivariable models included variables that were found to confound the association of interest (i.e., that importantly changed the effect of inflammation on structural outcomes). The following variables were tested as possible confounders: age (in years), sex (male versus female), HLA–B27 (positive versus negative), smoking status (smoker versus nonsmoker), CRP level (mg/liter), Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) score, the ASDAS (BASDAI score plus CRP level and ASDAS tested in separate models to avoid collinearity), treatment with nonsteroidal antiinflammatory drugs (yes/no), and treatment with tumor necrosis factor inhibitors (TNFi) (yes/no). Variables with a potential to change over time were modeled as such (i.e., all the above except sex and HLA–B27) in the longitudinal models.

RESULTS

Baseline characteristics. Of the total 708 patients from the DESIR cohort, 262 could have imaging at follow-up according to the protocol, and 202 had at least 2 consecutive visits with data available either on MRI-SI joints or MRI-spine (196 had both modalities, 3 had MRI-SI joints only, and 3 had MRI-spine only) and were therefore included. No significant baseline differences were found between patients included and not included in this study (Table 1). The presence of BME at baseline was more

frequent in the SI joints (29%) than in the spine (7% [ASAS definition]; 5% for ≥5 BME lesions). Likewise, structural damage was higher in the SI joints (e.g., ≥3 fatty lesions on MRI-SI joints: 12%) than in the spine (e.g., ≥3 fatty lesions on MRI-spine: 2%).

Structural progression according to the presence of objective inflammation at baseline. In total, 155 patients had complete MRI data at baseline and 5 years (141 both modalities, 10 MRI-SI joints only, and 4 MRI-spine only). Net progression, defined by ≥5 fatty lesions and/or erosions, ≥3 fatty lesions, and ≥3 erosions on MRI-SI joints, according to baseline objective inflammatory markers, is shown in Figure 1. Patients with BME on MRI-SI joints present at baseline had higher net progression rates compared to those who were BME-negative for all outcomes, irrespective of the CRP status (range if BME positive: 7–24%; range if BME negative: 0–4%). On MRI-spine overall, net progression was –0.7% both for ≥5 fatty lesions and/or erosions and for ≥5 fatty lesions; 0.7% for ≥3 fatty lesions, and 0% for ≥3 erosions and for ≥3 bone spurs. These low numbers precluded further analysis according to the presence of inflammatory markers at baseline.

Effect of inflammation on structural progression (multivariable models). *Sacroiliac joints.* The presence of BME on MRI-SI joints at baseline was predictive of the development of fatty lesions and erosions on MRI-SI joints 5 years later for all binary definitions (range odds ratio [OR] 4.1–5.6) after adjustment for CRP at baseline (Table 2). Similar results were found in the longitudinal models (after adjustment for ASDAS). On average, patients with BME on MRI-SI joints had a 5 times higher likelihood of having at least 3 fatty lesions in the subsequent visit as compared to those without BME (OR 5.1 [95% confidence interval (95% CI) 2.7, 9.6]) (Figure 2). The association between the continuous SPARCC score on MRI-SI joints and the various continuous structural outcomes was also always statistically significant and present in both models.

Spine. Testing the association of interest on MRI-spine was hampered by a low number of lesions, leading to imprecise estimates and, for some outcomes (i.e., ≥3 erosions and ≥5 fatty lesions/erosions), precluded the estimation of the effect (Table 3). Only the association between BME and ≥3 fatty lesions was statistically significant. The presence of baseline BME (ASAS definition) on MRI-spine was positively associated with ≥3 fatty lesions at 5 years on MRI-spine (OR 10.7 [95% CI 2.4, 49]). This effect was also positive in the longitudinal model (OR 15.6 [95% CI 4.8, 50.3]) (Figure 2). As in MRI-SI joints, CRP level (baseline models), and ASDAS (longitudinal models) have been found to confound the association of interest. Testing the effect of ≥5 BME lesions yielded similar results, but with wider a 95% CI (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24449/abstract). For continuous variables, a positive association could be found for fatty lesions alone or in combination with erosions, but not for erosions alone and bone spurs, both in baseline and longitudinal models.

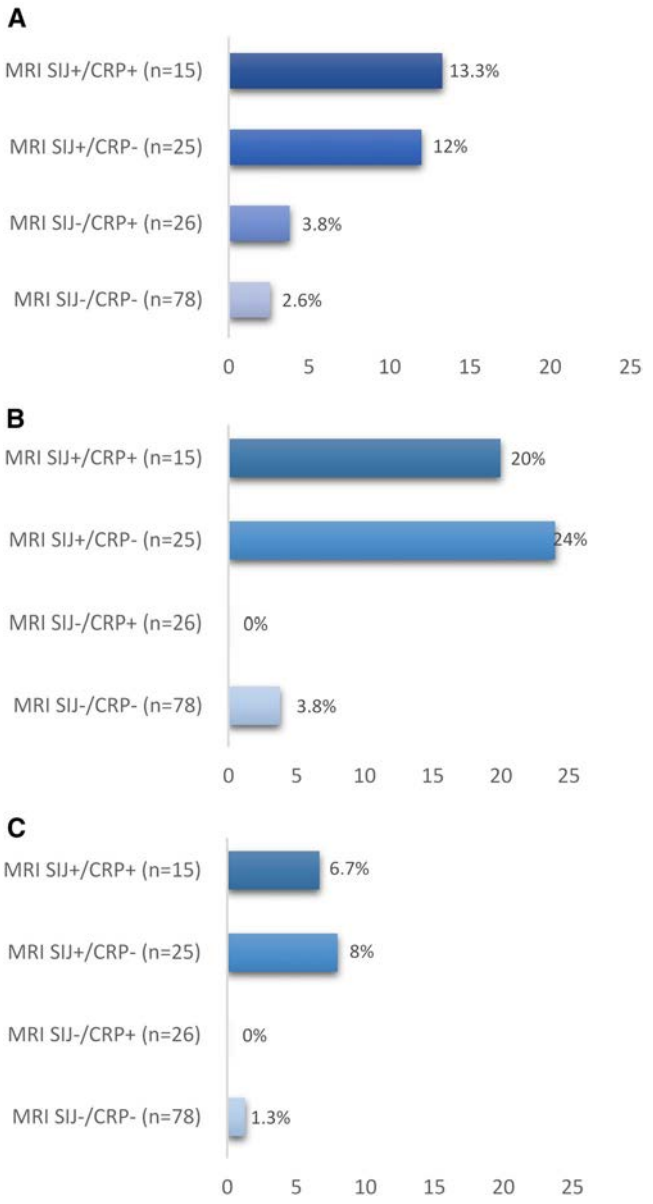

Figure 1. Net progression from magnetic resonance imaging (MRI) of the sacroiliac (SI) joints (MRI-SI joints) without structural lesions (MRI SIJ–) to MRI-SI joints with structural lesions (MRI SIJ+) defined by ≥5 fatty lesions and/or erosions (**A**), ≥3 fatty lesions (**B**), and ≥3 erosions (**C**) according to baseline objective inflammatory markers (MRI-SI joints inflammation and C-reactive protein [CRP] level). MRI-SIJ+ is defined as the presence of bone marrow edema on MRI-SI joints according to the Assessment of SpondyloArthritis international Society definition. CRP+ is defined as a CRP level ≥6 mg/liter at baseline. Net progression from MRI SIJ– to MRI SIJ+ at year 5 is defined as the number of progressors minus the number of regressors divided by the total number of patients in each category (n = 144; MRI-SI joints available both at baseline and year 5, and CRP level available at baseline).

DISCUSSION

In this prospective observational cohort study, we have shown that axial inflammation detected on MRI predicts subsequent development of structural lesions (especially fatty lesions) also on MRI over 5 years in patients with early axial SpA. This effect is independent of systemic inflammation and is seen at the level of both the SI joints and the spine but is measured more precisely in the SI joints where damage prevails in early disease. Our results add to the existing evidence by showing that the association between axial inflammation and some lesions reflecting structural damage can be measured with MRI in patients with early axial SpA.

In the current study, we have demonstrated an association between local inflammation and structural damage both measured on MRI in patients with early axial SpA. Involvement of the axial skeleton in axial SpA usually starts at the level of the SI joints (21,34,35). In line with the literature, we found that 6 times more patients showed structural damage (e.g., ≥3 fatty lesions) on MRI-SI joints (12%) than on MRI-spine (2%) at baseline. Consequently, the longitudinal association between BME and structural damage (e.g., ≥3 fatty lesions) on MRI-SI joints (OR 5.1 [95% CI 2.7, 9.6]) was found with a substantially higher precision (narrower confidence intervals) compared to the same effect in the spine (OR 15.6 [95% CI 4.8, 50.3]). Although it may seem that the effect of inflammation on damage is stronger on the spine than on the SI joints (OR 16 versus 5), this is not necessarily the case. It is well known that imprecise estimates tend to overestimate effect sizes (36).

Evidence that inflammation on MRI drives structural damage in early axial SpA is relevant to the practicing rheumatologist because it argues in favor of its use for prognostic stratification. In addition, if inflammation drives damage, it is logical to expect that interventions targeting the former will prevent, or at least inhibit, the latter. However, thus far, trial data do not support this claim (37). The complex, and yet not fully understood, pathophysiology of new bone formation in axial SpA may, at least in part, explain this disappointing result. For instance, it has been shown that systemic inflammation, measured by the ASDAS, predicts spinal radiographic progression in radiographic axial SpA (6,8). However, progression was still found in patients with inactive disease. Similarly, in another study, inflammation at the level of vertebral unit increased the likelihood of the formation of a new syndesmophyte in the same location 2 years later, but most new syndesmophytes appeared in vertebral units without signs of inflammation (12). These data highlight the relevance of inflammation in driving structural progression but also suggest that other mechanisms may play a role.

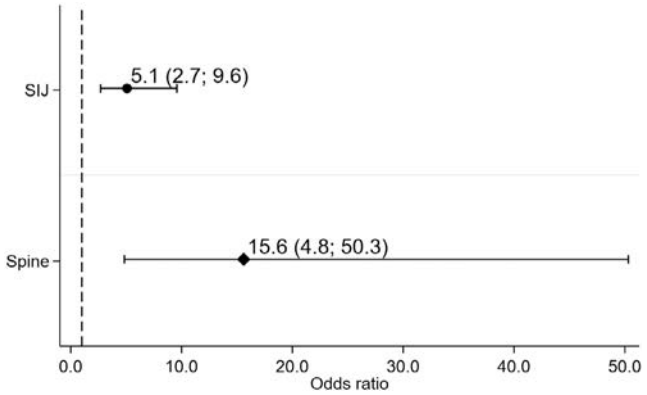

Figure 2. The effect of bone marrow edema (according to the Assessment of SpondyloArthritis international Society definition) on structural damage (defined as ≥3 fatty lesions) over 5 years both in the sacroiliac joints (SIJ) and spine (longitudinal time-lagged models with autoregression). Circles represent ≥3 fatty lesions on MRI of the SI joints. Diamonds represent ≥3 fatty lesions on MRI of the spine. Bars show the 95% confidence interval. MRI = magnetic resonance imaging.

Table 2. The effect of inflammation detected by magnetic resonance imaging (MRI) on structural damage detected by MRI in the sacroiliac joints (multivariable models)*

	≥5 fatty lesions/ erosions, OR (95% CI)	≥3 fatty lesions, OR (95% CI)	≥3 erosions, OR (95% CI)	Fatty lesions/ erosions, β (95% CI)	Fatty lesions, β (95% CI)	Erosions β (95% CI)
Binary scores						
BME at baseline (range 144–151)†	5.6 (3.1, 10.0)‡	4.2 (2.4, 7.3)‡	4.1 (2.1, 7.8)	–	–	–
BME over 5 years (range 197–199)§	7.7 (4.5, 13.4)¶	5.1 (2.7, 9.6)¶	3.2 (1.9, 5.3)	–	–	–
Continuous scores						
SPARCC at baseline (range 144–151)†	–	–	–	0.23 (0.15, 0.31)‡	0.12 (0.05, 0.19)‡	0.12 (0.06, 0.18)
SPARCC over 5 years (range 197–199)§	–	–	–	0.13 (0.07, 0.19)¶	0.10 (0.04, 0.16)¶	0.04 (0.01, 0.06)

* 95% CI = 95% confidence interval; ASDAS = Ankylosing Spondylitis Disease Activity Score; BME = bone marrow edema (according to the Assessment of SpondyloArthritis international Society definition [positive/negative]); OR = odds ratio; SPARCC = Spondyloarthritis Research Consortium of Canada.
† Multilevel generalized estimating equation (GEE) models (i.e., effect of inflammation at baseline on the outcome at 5 years, taking the scores from the individual readers into account).
‡ Adjusted for C-reactive protein (CRP) level at baseline.
§ Longitudinal multilevel time-lagged GEE models with autoregression (i.e., effect of inflammation at t on the outcome at $t + 1$, adjusted for the outcome at t, taking the scores from the individual readers into account).
¶ Adjusted for time-lagged ASDAS-CRP score.

Table 3. The effect of inflammation detected by magnetic resonance imaging (MRI) on structural damage detected by MRI in the spine (multivariable models)*

	≥5 fatty lesions/ erosions	≥5 fatty lesions	≥3 fatty lesions	≥3 erosions	≥3 bone spurs
Binary scores					
BME at baseline (n = 139)†	‡	‡	10.7 (2.4, 49.0)§	‡	3.2 (0.4, 27.8)§
BME over 5 years (n = 197)¶	‡	0.9 (0.8, 1.2)#	15.6 (4.8, 50.3)#	‡	2.8 (0.8, 9.6)#
Continuous scores					
SPARCC at baseline (range 139–145)†	0.10 (0.01, 0.18)§	0.08 (0.02, 0.14)	0.08 (0.02, 0.14)	0.02 (0.00, 0.03)†	0.01 (–0.01, 0.03)†
SPARCC over 5 years (n = 197)¶	0.06 (0.02, 0.11)#	0.07 (0.02, 0.11)#	0.07 (0.02, 0.11)#	0.00 (–0.01, 0.01)#	0.01 (0.00, 0.02)

* Values are the odds ratio (95% confidence interval). ASDAS = Ankylosing Spondylitis Disease Activity Score; BME = bone marrow edema (according to the Assessment of SpondyloArthritis international Society definition [≥3 lesions; positive/negative]); SPARCC = Spondyloarthritis Research Consortium of Canada.
† Multilevel generalized estimating equation (GEE) model (i.e., effect of inflammation at baseline on the outcome at 5 years, taking the scores from the individual readers into account).
‡ Model fails to find a mathematical solution due to low number of events.
§ Adjusted for C-reactive protein (CRP) level at baseline.
¶ Longitudinal multilevel time-lagged GEE models with autoregression (i.e., effect of inflammation at t on the outcome at $t + 1$, adjusted for the outcome at t, taking the scores from the individual readers into account).
# Adjusted for time-varying lagged ASDAS-CRP score.

Biology, however, cannot fully explain the failure of antiinflammatory drugs in modifying the effect of inflammation on structural damage. The lack of sensitivity to change of the outcome measures has also been proposed previously as a likely explanation (38). If an intervention truly prevents further damage by reducing inflammation (or by any other means), low sensitivity to change of the outcome measure may prevent that such effect becomes evident (e.g., no significant difference between active drug and placebo). Thus far, progression of structural damage has been measured mostly using conventional radiographs, with the modified Stoke Ankylosing Spondylitis Spine Score (mSASSS) and the mNY grading system as the outcome measures used most often in the spine and SI joints, respectively. However, both the mSASSS and the mNY have low sensitivity to change, and assessing radiographic progression with the latter is further challenged by its poor reliability (3,14,15,39). It remains to be proven that structural lesions detected on MRI are more sensitive to change than those on radiographs. However, our study suggests that different lesions may yield different results. For instance, compared with erosions or bony spurs, fatty lesions were more prevalent in our population of patients with early axial SpA, especially in the SI joints, leading to more precise estimates. Thus, our data may inform future research aiming at clarifying whether MRI is a valid alternative to conventional radiography in detecting structural treatment effects in patients with axial SpA.

Our study is not without limitations. First, inflammatory and structural lesions, per patient, were read together by the same reader, which obviously may result in overestimating the association between both. This contrasts with other studies in which inflammation and damage were blindly measured using different imaging modalities. However, it should be stressed that readers were still blinded to time order. That is, they did not know if a certain lesion (e.g., BME) pertained to a baseline or to a follow-up image. Thus, causality by reading, although not impossible, is unlikely to fully explain the impressive associations found in our study. Second, the lack of an association between vertebral corner inflammation on MRI-spine and erosions and bone spurs should be interpreted with caution. Even though a true lack of association cannot be ruled out, as mentioned above, this also may be due to low statistical power driven by a low number of these lesions in the spine. The role of inflammation on sites other than vertebral corners for the progression of spinal damage should be addressed in future studies.

In conclusion, we have shown that local inflammation is associated with development of structural damage (e.g., fatty lesions), both measured with MRI, over 5 years in the SI joints and spine in early axial SpA. This association is detected with more precision on the SI joints, where structural damage prevails, compared to the spine in early disease. These findings support the concept that MRI is a valid alternative to conventional radiographs in detecting the structural consequences of axial inflammation in patients with early axial SpA.

ACKNOWLEDGMENTS

The authors thank the following at the different regional participating centers: Maxime Dougados (Paris-Cochin B), Andre Kahan (Paris-Cochin A), Philippe Dieudé (Paris-Bichat), Bruno Fautrel (Paris-La Pitie-Salpetriere), Francis Berenbaum (Paris-Saint-Antoine), Pascal Claudepierre (Creteil), Maxime Breban (Boulogne-Billancourt), Bernadette Saint-Marcoux (Aulnay-sous-Bois), Philippe Goupille (Tours), Jean Francis Maillefert (Dijon), Emmanuelle Dernis (Le Mans), Daniel Wendling (Besancon), Bernard Combe (Montpellier), Liana Euller-Ziegler (Nice), Pascal Richette (ParisLariboisiere), Pierre Lafforgue (Marseille), Patrick Boumier (Amiens), Martin Soubrier (ClermontFerrand), Nadia Mehsen (Bordeaux), Damien Loeuille (Nancy), Rene-Marc Flipo (Lille), Alain Saraux (Brest), Xavier Mariette (LeKremlin-Bicetre), Alain Cantagrel (Toulouse), and Olivier Vittecoq (Rouen). They also thank the research nurses, the staff members of the Clinical Research Unit of Paris Centre, the staff members of the Biological Resource Center of Bichat Hospital, the staff members of the Department of Statistics of Nîmes, and all the

investigators, and in particular Jerome Allain, Emmanuelle Dernis, Salah Ferkal, Clement Prati, Marie-Agnes Timsit, and Eric Toussirot, for active patient recruitment and monitoring. The authors are grateful to Miranda van Lunteren for her work in preparing the database used in this study, and to all DESIR readers, without whom this study would not be possible.

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. Sepriano 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. Sepriano, Ramiro, Landewé, van der Heijde.

Acquisition of data. Moltó, Claudepierre, Wendling, Dougados.

Analysis and interpretation of data. Sepriano, Ramiro, Landewé, Moltó, Claudepierre, Wendling, Dougados, van der Heijde.

REFERENCES

1. Ramiro S, Stolwijk C, van Tubergen A, van der Heijde D, Dougados M, van den Bosch F, et al. Evolution of radiographic damage in ankylosing spondylitis: a 12 year prospective follow-up of the OASIS study. Ann Rheum Dis 2015;74:52–9.
2. Poddubnyy D, Brandt H, Vahldiek J, Spiller I, Song IH, Rudwaleit M, et al. The frequency of non-radiographic axial spondyloarthritis in relation to symptom duration in patients referred because of chronic back pain: results from the Berlin early spondyloarthritis clinic. Ann Rheum Dis 2012;71:1998–2001.
3. Dougados M, Sepriano A, Molto A, van Lunteren M, Ramiro S, de Hooge M, 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.
4. Van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis: a proposal for modification of the New York criteria. Arthritis Rheum 1984;27:361–8.
5. Baraliakos X, Listing J, Rudwaleit M, Sieper J, Braun J. The relationship between inflammation and new bone formation in patients with ankylosing spondylitis. Arthritis Res Ther 2008;10:R104.
6. Poddubnyy D, Protopopov M, Haibel H, Braun J, Rudwaleit M, Sieper J. High disease activity according to the Ankylosing Spondylitis Disease Activity Score is associated with accelerated radiographic spinal progression in patients with early axial spondyloarthritis: results from the German Spondyloarthritis Inception Cohort. Ann Rheum Dis 2016;75:2114–8.
7. Poddubnyy D, Rudwaleit M, Haibel H, Listing J, Märker-Hermann E, Zeidler H, et al. Rates and predictors of radiographic sacroiliitis progression over 2 years in patients with axial spondyloarthritis. Ann Rheum Dis 2011;70:1369–74.
8. Ramiro S, van der Heijde D, van Tubergen A, Stolwijk C, Dougados M, van den Bosch F, et al. Higher disease activity leads to more structural damage in the spine in ankylosing spondylitis: 12-year longitudinal data from the OASIS cohort. Ann Rheum Dis 2014;73:1455–61.
9. Machado PM, Baraliakos X, van der Heijde D, Braun J, Landewé R. MRI vertebral corner inflammation followed by fat deposition is the strongest contributor to the development of new bone at the same vertebral corner: a multilevel longitudinal analysis in patients with ankylosing spondylitis. Ann Rheum Dis 2016;75:1486–93.
10. Maksymowych WP, Chiowchanwisawakit P, Clare T, Pedersen SJ, Østergaard M, Lambert RG. Inflammatory lesions of the spine on magnetic resonance imaging predict the development of new syndesmophytes in ankylosing spondylitis: evidence of a relationship between inflammation and new bone formation. Arthritis Rheum 2009;60:93–102.
11. Baraliakos X, Heldmann F, Callhoff J, Listing J, Appelboom T, Brandt J, et al. Which spinal lesions are associated with new bone formation in patients with ankylosing spondylitis treated with anti-TNF agents? A long-term observational study using MRI and conventional radiography. Ann Rheum Dis 2014;73:1819–25.
12. Van der Heijde D, Machado P, Braun J, Hermann KG, Baraliakos X, Hsu B, et al. MRI inflammation at the vertebral unit only marginally predicts new syndesmophyte formation: a multilevel analysis in patients with ankylosing spondylitis. Ann Rheum Dis 2012;71:369–73.
13. Sepriano A, Rudwaleit M, Sieper J, van den Berg R, Landewé R, van der Heijde D. Five-year follow-up of radiographic sacroiliitis: progression as well as improvement? Ann Rheum Dis 2016;75:1262–3.
14. Van den Berg R, Lenczner G, Feydy A, van der Heijde D, Reijnierse M, Saraux A, et al. Agreement between clinical practice and trained central reading in reading of sacroiliac joints on plain pelvic radiographs: results from the DESIR cohort. Arthritis Rheumatol 2014;66:2403–11.
15. Van Tubergen A, Heuft-Dorenbosch L, Schulpen G, Landewé R, Wijers R, van der Heijde D, et al. Radiographic assessment of sacroiliitis by radiologists and rheumatologists: does training improve quality? Ann Rheum Dis 2003;62:519–25.
16. Ostergaard M, Maksymowych WP, Pedersen SJ, Chiowchanwisawakit P, Lambert RG. Structural lesions detected by magnetic resonance imaging in the spine of patients with spondyloarthritis: definitions, assessment system, and reference image set. J Rheumatol 2009;36 Suppl 84:18–34.
17. Krabbe S, Sørensen IJ, Jensen B, Møller JM, Balding L, Madsen OR, et al. Inflammatory and structural changes in vertebral bodies and posterior elements of the spine in axial spondyloarthritis: construct validity, responsiveness and discriminatory ability of the anatomy-based CANDEN scoring system in a randomised placebo-controlled trial. RMD Open 2018;4:e000624.
18. Weber U, Lambert RG, Østergaard M, Hodler J, Pedersen SJ, Maksymowych WP. The diagnostic utility of magnetic resonance imaging in spondylarthritis: an international multicenter evaluation of one hundred eighty-seven subjects. Arthritis Rheum 2010;62:3048–58.
19. Maksymowych WP, Lambert RG, Østergaard M, Pedersen SJ, Machado PM, Weber U, et al. MRI lesions in the sacroiliac joints of patients with spondyloarthritis: an update of definitions and validation by the ASAS MRI working group. Ann Rheum Dis 2019;78:1550–8.
20. Lukas C, Cyteval C, Dougados M, Weber U. MRI for diagnosis of axial spondyloarthritis: major advance with critical limitations 'Not everything that glisters is gold (standard)'. RMD Open 2018;4:e000586.
21. Ramiro S, Claudepierre P, Sepriano A, van Lunteren M, Molto A, Feydy A, et al. Which scoring method depicts spinal radiographic damage in early axial spondyloarthritis best? Five-year results from the DESIR cohort. Rheumatology (Oxford) 2018;57:1991–2000.
22. Dougados M, Etcheto A, Molto A, Alonso A, Bouvet S, Daurès JP, et al. Clinical presentation of patients suffering from recent onset chronic inflammatory back pain suggestive of spondyloarthritis: the DESIR cohort. Joint Bone Spine 2015;82:345–51.
23. De Hooge M, Pialat JB, Reijnierse M, van der Heijde D, Claudepierre P, Saraux A, et al. Assessment of typical SpA lesions on MRI of the spine: do local readers and central readers agree in the DESIR-cohort at baseline? Clin Rheumatol 2017;36:1551–9.
24. Jacquemin C, Rubio Vargas R, van den Berg R, Thévenin F, Lenczner G, Reijnierse M, et al. What is the reliability of non-trained investigators in recognising structural MRI lesions of sacroiliac joints in patients with recent inflammatory back pain? Results of the DESIR cohort. RMD Open 2016;2:e000303.
25. De Hooge M, van den Berg R, Navarro-Compan V, Reijnierse M, van Gaalen F, Fagerliet K, al. Patients with chronic back pain of short duration from the SPACE cohort: which MRI structural lesions in the

sacroiliac joints and inflammatory and structural lesions in the spine are most specific for axial spondyloarthritis? Ann Rheum Dis 2016; 75:1308–14.

26. Bennett AN, Rehman A, Hensor EM, Marzo-Ortega H, Emery P, McGonagle D. The fatty Romanus lesion: a non-inflammatory spinal MRI lesion specific for axial spondyloarthropathy. Ann Rheum Dis 2010;69:891–4.
27. Lambert RG, Bakker PA, van der Heijde D, Weber U, Rudwaleit M, Hermann KG, et al. Defining active sacroiliitis on MRI for classification of axial spondyloarthritis: update by the ASAS MRI working group. Ann Rheum Dis 2016;75:1958–63.
28. Maksymowych WP, Inman RD, Salonen D, Dhillon SS, Williams M, Stone M, et al. Spondyloarthritis Research Consortium of Canada magnetic resonance imaging index for assessment of sacroiliac joint inflammation in ankylosing spondylitis. Arthritis Rheum 2005;53:703–9.
29. Rudwaleit M, Jurik AG, Hermann KG, Landewé R, van der Heijde D, Baraliakos X, et al. Defining active sacroiliitis on magnetic resonance imaging (MRI) for classification of axial spondyloarthritis: a consensual approach by the ASAS/OMERACT MRI group. Ann Rheum Dis 2009; 68:1520–7.
30. Hermann KG, Baraliakos X, van der Heijde DM, Jurik AG, Landewé R, Marzo-Ortega H, et al. Descriptions of spinal MRI lesions and definition of a positive MRI of the spine in axial spondyloarthritis: a consensual approach by the ASAS/OMERACT MRI study group. Ann Rheum Dis 2012;71:1278–88.
31. Maksymowych WP, Inman RD, Salonen D, Dhillon SS, Krishnananthan R, Stone M, et al. Spondyloarthritis Research Consortium of Canada magnetic resonance imaging index for assessment of spinal inflammation in ankylosing spondylitis. Arthritis Rheum 2005; 53:502–9.
32. Madari Q, Sepriano A, Ramiro S, Molto A, Claudepierre P, Wendling D, et al. 5-year follow-up of spinal and sacroiliac MRI abnormalities in early axial spondyloarthritis: data from the DESIR cohort. RMD Open 2020;6
33. Sepriano A, Ramiro S, Landewé R, Dougados M, van der Heijde D. Percentage of progressors in imaging: can we ignore regressors? RMD Open 2019;5:e000848.
34. Rudwaleit M, Haibel H, Baraliakos X, Listing J, Märker-Hermann E, Zeidler H, et al. The early disease stage in axial spondylarthritis: results from the German Spondyloarthritis Inception Cohort. Arthritis Rheum 2009;60:717–27.
35. Wanders AJ, Landewé RB, Spoorenberg A, Dougados M, van der Linden S, Mielants H, et al. What is the most appropriate radiologic scoring method for ankylosing spondylitis? A comparison of the available methods based on the Outcome Measures in Rheumatology Clinical Trials filter. Arthritis Rheum 2004;50: 2622–32.
36. Van Calster B, Steyerberg EW, Collins GS, Smits T. Consequences of relying on statistical significance: some illustrations. Eur J Clin Invest 2018;48:e12912.
37. Van der Heijde D, Landewé R. Inhibition of spinal bone formation in AS: 10 years after comparing adalimumab to OASIS. Arthritis Res Ther 2019;21:225.
38. Maksymowych WP. The role of imaging in the diagnosis and management of axial spondyloarthritis. Nat Rev Rheumatol 2019;15: 657–72.
39. Ramiro S, van der Heijde D, Sepriano A, van Lunteren M, Moltó A, Feydy A, et al. Spinal radiographic progression in early axial spondyloarthritis: five-year results from the DESIR cohort. Arthritis Care Res (Hoboken) 2019;71:1678–84.

Arthritis Care & Research
Vol. 74, No. 2, February 2022, pp 251–258
DOI 10.1002/acr.24459

American College of Rheumatology
Empowering Rheumatology Professionals

Imaging Outcomes for Axial Spondyloarthritis and Sensitivity to Change: A Five-Year Analysis of the DESIR Cohort

Alexandre Sepriano,[1] Sofia Ramiro,[2] Désirée van der Heijde,[3] Maxime Dougados,[4] Pascal Claudepierre,[5] Antoine Feydy,[6] Monique Reijnierse,[3] Damien Loeuille,[7] and Robert Landewé[8]

Objective. To compare the sensitivity to change of different imaging scoring methods in patients with early axial spondyloarthritis (SpA).

Methods. Patients from the Devenir des Spondylarthropathies Indifférenciées Récentes (DESIR) cohort fulfilling the Assessment of SpondyloArthritis international Society criteria for axial SpA were included. Radiographs and magnetic resonance imaging (MRI) of the sacroiliac (SI) joints and spine were obtained at baseline, 1, 2, and 5 years. Each image was scored by 2 or 3 readers in 3 separate reading waves. The rate of change of outcomes measuring inflammation of the spine and SI joints (e.g., Spondyloarthritis Research Consortium of Canada [SPARCC] score) and structural damage on MRI (e.g., ≥3 fatty lesions) and radiographs (e.g., modified New York grading) was assessed using multilevel generalized estimating equation models (taking all readers and waves into account). To allow comparisons across outcomes, rates were standardized (difference between the individual's value and the population mean divided by the SD).

Results. In total, 345 patients were included. Inflammation detected on MRI of the SI joints (MRI-SI joints) (standardized rate range –0.278, –0.441) was more sensitive to change compared to spinal inflammation (range –0.030, –0.055). Structural damage in the SI joints showed a higher standardized rate of change on MRI-SI joints (range 0.015, 0.274) compared to radiography of the SI joints (range 0.043, 0.126). MRI-SI joints damage defined by ≥3 fatty lesions showed the highest sensitivity to change (0.274). Spinal structural damage slowly progressed over time with no meaningful difference between radiographic (range 0.037, 0.043) and MRI structural outcomes (range 0.008, 0.027).

Conclusion. Structural damage assessed in pelvic radiographs has low sensitivity to change, while fatty lesions detected on MRI-SI joints are a promising alternative. In contrast, MRI of the spine is not better than radiography of the spine in detecting structural changes in patients with early axial SpA.

INTRODUCTION

Several imaging outcomes have been developed to assess inflammation and structural damage over time in patients with axial spondyloarthritis (SpA). A recent systematic literature review informing the European Alliance of Associations for Rheumatology recommendations for the use of imaging in the diagnosis and management of SpA in clinical practice identified several studies testing

The DESIR study is supported by an unrestricted grant from Pfizer and is conducted as a program of the Hospitalier de Recherche Clinique with Assistance Publique Hopitaux de Paris as the sponsor and under the umbrella of the French Society of Rheumatology, which supports the cohort. The DESIR cohort is conducted under the Assistance Publique Hopitaux de Paris via the Clinical Research Unit Paris Centre and under the umbrella of the French Society of Rheumatology and Institut national de la sante et de la recherche medicale (INSERM). Database management is performed within the Department of Epidemiology and Biostatistics. Dr. Sepriano's work was supported by the Fundação para a Ciência e Tecnologia (doctoral grant SFRH/BD/108246/2015).

[1]Alexandre Sepriano, MD: Leiden University Medical Center, Leiden, The Netherlands, and NOVA Medical School, Universidade Nova de Lisboa, Lisbon, Portugal; [2]Sofia Ramiro, MD, PhD: Leiden University Medical Center, Leiden, and Zuyderland Medical Center, Heerlen, The Netherlands; [3]Désirée van der Heijde, MD, PhD, Monique Reijnierse, MD, PhD: Leiden University Medical Center, Leiden, The Netherlands; [4]Maxime Dougados, MD: Department of Rheumatology, Hôpital Cochin, Assistance Publique, Hôpitaux de Paris, INSERM (U1153), and Clinical Epidemiology and Biostatistics, PRES Sorbonne Paris-Cité, Paris, France; [5]Pascal Claudepierre, MD: Department of Rheumatology, Hôpital Henri-Mondor, Assistance Publique-Hôpitaux de Paris, and Université Paris Est Créteil, Créteil, France; [6]Antoine Feydy, MD, PhD: Radiology B Department, Cochin Hôpital Paris and Descartes University, Paris, France; [7]Damien Loeuille, MD, PhD: Department of Rheumatology, Hôpital Brabois, Nancy, France; [8]Robert Landewé, MD, PhD: Amsterdam University Medical Center, Amsterdam, and Zuyderland Medical Center, Heerlen, The Netherlands.

No potential conflicts of interest relevant to this article were reported.

Address correspondence to Alexandre Sepriano, MD, Leiden University Medical Center, Department of Rheumatology, Albinusdreef 2, PO Box 9600, 2300 RC Leiden, The Netherlands. Email: alexsepriano@gmail.com.

Submitted for publication March 9, 2020; accepted in revised form September 15, 2020.

SIGNIFICANCE & INNOVATIONS

- Several imaging outcomes are available to measure inflammation and damage over time in patients with axial spondyloarthritis (SpA); however, direct comparisons of their sensitivity to change are scarce, especially in early disease.
- In early axial SpA, outcomes of inflammation measured on magnetic resonance imaging (MRI) are more sensitive to change on the sacroiliac (SI) joints than on the spine.
- MRI of the SI joints is more sensitive in capturing change in structural damage, especially fatty lesions, than pelvic radiographs, while MRI of the spine is not better than spinal radiographs in detecting structural changes in patients with early axial SpA.
- Results from this study may help in prioritizing imaging scoring methods in subsequent observational or interventional studies in early axial SpA.

the utility of magnetic resonance imaging (MRI) and radiographs of the sacroiliac (SI) joints and spine on monitoring disease activity and structural damage over time (1). However, these studies mostly assessed only 1 score each and focused on comparing imaging to clinical measures of disease activity, disability, and mobility, which means that they mostly addressed their validity.

In addition to validity, in order to prioritize imaging outcomes measuring similar aspects of the disease (i.e., inflammation or structural damage), the other aspects of the Outcome Measures in Rheumatology (OMERACT) filter, namely discrimination (sensitivity to change and reliability) and feasibility, should also be taken into account (2). However, direct comparisons of the discriminative ability and feasibility of imaging outcomes in axial SpA have been seldom performed, and almost only in later phases of the disease (radiographic axial SpA) (3–5). An exception to this is the comparison of the different spinal radiographic scoring methods performed in the Devenir des Spondylarthropathies Indifférenciées Récentes (DESIR) cohort and previously reported by our team (6).

A better understanding of which imaging findings (reflecting inflammation or structural damage), imaging modality (MRI or radiographs), and anatomic location (SI joints or spine) are most informative to monitor axial changes in the entire spectrum of axial SpA (also including nonradiographic axial SpA) over time is still a major unmet need. We aimed to compare the sensitivity to change of different MRI and radiographic scoring methods in patients with early axial SpA.

PATIENTS AND METHODS

Patients and study design. Five-year data from patients with early axial SpA from the DESIR cohort have been used (ClinicalTrials.gov identifier: NCT01648907) (7). Patients had to fulfill the Assessment of SpondyloArthritis international Society (ASAS) criteria for axial SpA and to have ≥1 radiograph and/or MRI reading available during the 5-year follow-up period to be included in the current study. The database used for the current analysis was locked on June 20, 2016. The study was approved by the appropriate local medical ethical committees. All patients provided signed informed consent upon participation.

Imaging scoring procedures. Radiographs of the SI joints and spine and MRIs of the SI joints (MRI-SI joints) and spine (MRI-spine) were obtained at baseline, 1, 2, and 5 years. Each image was independently scored in 3 reading waves by trained central readers blinded to chronology, clinical data, and to the results of other imaging modalities. In wave 1, baseline images were scored by 2 readers and 1 adjudicator (in case of disagreement). In wave 2, images from baseline, 1, and 2 years were also scored by 2 readers and 1 adjudicator. In wave 3, images from baseline, 2, and 5 years were scored by 3 central readers. Readers and adjudicators varied across modalities and waves (8) (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24459/abstract). By protocol, radiographs were performed in all 25 participating centers at each time point, but MRIs were only performed in all centers at baseline, while MRIs at 1, 2, and 5 years were only obtained in 9 centers from Paris.

Inflammation outcomes. Inflammation on MRI-SI joints was assessed using the ASAS definition (positive/negative) and the Spondyloarthritis Research Consortium of Canada (SPARCC) score (range 0–72) (9–11). Bone marrow edema (BME) on MRI-spine was defined according to the ASAS definition (≥3 vertebral corner lesions; yes/no) (12). In addition, a cutoff of 5 vertebral corner BME lesions (typical of axial SpA and present in ≥2 consecutive slices) was also assessed according to the Canada–Denmark method, as it has been shown to be highly specific of axial SpA (13). The total spine SPARCC score (range 0–414) and Berlin score (range 0–69) were used as continuous inflammatory outcomes (3,14).

Structural outcomes. Structural damage on radiography of the SI joints was assessed according to the modified New York (mNY) system as continuous (range 0–8) and as a binary (positive/negative) score (15). Two additional binary definitions were assessed: worsening of ≥1 grade in ≥1 SI joints (yes/no); and worsening of ≥1 grade in ≥1 SI joints, with a 5-year grade ≥2 in the worsened joint (yes/no) (16).

An adaptation of the MRI-SI joints structural score by Weber et al, previously described by our team (17), was used to define individual structural lesions on MRI-SI joints (18). In summary, fatty lesions, erosions, and ankylosis/partial ankylosis are scored as originally described. Sclerosis was added. Fatty lesions, erosions, and sclerosis were marked as present if seen on ≥2

consecutive slices (maximum 5 lesions in 6 slices per each of the 8 quadrants in both SI joints). Ankylosis or partial ankylosis was considered present if seen on a single slice. Partial ankylosis and ankylosis cannot occur simultaneously in a quadrant, and ankylosis always involves 2 quadrants; therefore, the corresponding scoring range is 0–24. In the absence of a formal definition of presence of structural damage on MRI-SI joints, we considered 3 definitions previously shown most discriminatory in early axial SpA: ≥5 fatty lesions and/or erosions; ≥3 erosions; and ≥3 fatty lesions (13). Continuous structural lesions on MRI-SI joints were defined as number of fatty lesions and/or erosions (range 0–80), number of erosions (range 0–40), number of fatty lesions (range 0–40), and total number of lesions with (range 0–144) and without (range 0–104) sclerosis.

Structural lesions on radiography of the spine were assessed as the presence of ≥1 syndesmophyte (yes/no) and by using the modified Stoke Ankylosing Spondylitis Spine Score (mSASSS; range 0–72) (19). Structural lesions on MRI-spine were scored according to the Canada–Denmark method (20,21). In the absence of a formal definition, we defined structural damage as ≥5 fatty lesions, also previously shown to be highly specific for axial SpA (13). The total number of structural lesions (fatty lesions, erosions, bone spurs, ankylosis) (range 0–322) was assessed, as well as the total number of fatty lesions, erosions, and bone spurs (range 0–92 for all).

A detailed description of all scores is provided in Supplementary Tables 2–10, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24459/abstract. The interreader reliability of the radiographic and MRI outcomes used in this study has been reported in detail elsewhere (6,17) and is summarized in Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24459/abstract.

Statistical analysis. The baseline value for each outcome was defined by a combination algorithm of the scores from the 3 readers from wave 3 (agreement between ≥2 of 3 for binary, and a mean of 3 readers for continuous outcomes). The rate of change of each outcome was analyzed using generalized estimating equations (GEEs), with time in years as the explanatory variable of interest. Each outcome was analyzed per patient, per time point and per individual reader, and the yearly rate of change estimated using so-called integrated analysis, including all patients with ≥1 score from ≥1 reader from ≥1 reading wave. Different to traditional measures of sensitivity to change (e.g., Cohen's effect size), this method, which we have previously explained in detail (8), appropriately handles the multilevel data structure of our data. All patients had to have ≥1 score from all outcomes, thus ensuring that the same patients are used across all analyses. All variables were standardized. A standardized variable (metric free) was defined at the patient level as the difference between the individual's value and the population mean divided by the population SD. Each standardized variable has a mean of 0 and a variance of 1 and reads as the number of SD above (positive) or below (negative) the mean.

In addition, the relative standardized rate of change (i.e., the standardized yearly rate of change of an outcome divided by the corresponding rate of a reference imaging outcome) was calculated. For this calculation, a value >1 means larger sensitivity, and a value <1 lower sensitivity compared to the reference (the further away from 1, the larger the difference). Three types of references were defined: 1) inflammation common reference (comparing all inflammation outcomes to sacroiliitis on MRI-SI joints [ASAS definition]); 2) structural common reference (comparing all structural outcomes to sacroiliitis on radiography of the SI joints [mNY]); and 3) modality reference (comparing outcomes to a reference within each modality and anatomic site).

Goodness-of-fit statistics (quasi-likelihood under the independence model criterion [QIC]) were used to get an impression on how much of the outcome variability was explained by each model. Different transformations of time were tested to assess which one yielded the lowest QIC (better fit). A nonlinear model was chosen if it best fit the data and if the nonlinear factor (e.g., quadratic term) added to the model was significant ($P < 0.05$). Stata, version 15.1, was used for the analyses.

RESULTS

Baseline characteristics. In total, 345 patients were included (mean ± SD symptom duration 1.6 ± 0.9 years; 53% were male patients, and 89% HLA–B27 positive [Table 1]). Baseline inflammation on MRI was more frequently present at the SI joints (active sacroiliitis: 39%) than at the spine level (BME ≥5 lesions: 6%) (Table 2). Structural damage at baseline was limited in the SI joints (21% mNY positive) and even more in the spine (≥1 syndesmophyte: 6%) (Table 3).

Sensitivity to change of the different imaging outcomes. Inflammation on MRI-SI joints showed a higher sensitivity to change than on MRI-spine, the latter remaining essentially unchanged over time. This was true for the dichotomous ASAS MRI-SI joints score (standardized yearly rate of change –0.278) and especially for the continuous SPARCC score (standardized yearly rate of change –0.441), while the standardized yearly rates of change for MRI-spine ranged only between –0.030 and –0.055 (Table 2). The differences between SI joints and spine inflammation outcomes become especially evident with the relative standardized rate of change. Compared to the ASAS definition of a positive MRI-SI joints (inflammation common reference, i.e., a value of 1), all inflammation outcomes in the spine were much less sensitive to change (range of relative standardized rates 0.094–0.531; i.e., all values far below 1).

Structural damage in the SI joints increased over time but with a larger yearly rate on MRI-SI joints (standardized rate range 0.015–0.274) compared to radiography of the SI joints

Table 1. Patient and disease characteristics at baseline and during follow-up*

Characteristic	Baseline (n = 345)	1 year (n = 345)	2 years (n = 342)	5 years (n = 320)
Age at baseline, mean ± SD years	31.0 ± 7.0	–	–	–
Male sex	183 (53)	–	–	–
Symptoms duration, mean ± SD years	1.6 ± 0.9	–	–	–
Current smokers†	135 (39)	127 (39)	118 (37)	92 (34)
HLA–B27	307 (89)	–	–	–
Radiographic sacroiliitis (mNY)‡	73 (21)	NA	68 (23)	68 (27)
BASDAI score, mean ± SD (range 0–10)†	4.1 ± 2.0	3.2 ± 2.2	3.1 ± 2.2	2.9 ± 2.0
ASDAS-CRP score, mean ± SD‡	2.6 ± 1.0	2.1 ± 0.9	2.0 ± 0.9	2.0 ± 0.9
Elevated CRP (≥6 mg/liter)‡	109 (33)	64 (20)	69 (22)	57 (22)
BASFI score, mean ± SD (0–10)†	2.7 ± 2.2	2.1 ± 2.1	2.1 ± 2.2	2.0 ± 2.0
TNFi treatment‡	0 (0)	76 (24)	94 (29)	111 (42)
NSAID treatment†	329 (95)	250 (77)	216 (68)	180 (66)

* Values are the number (%) unless indicated otherwise. ASDAS = Ankylosing Spondylitis Disease Activity Score; BASDAI = Bath Ankylosing Spondylitis Disease Activity Index; BASFI = Bath Ankylosing Spondylitis Functional Index; CRP = C-reactive protein; mNY = modified New York criteria (scored in wave 3); NA = not applicable (imaging in wave 3 is only scored at baseline, 2 years, and 5 years); NSAID = nonsteroidal anti-inflammatory drugs; TNFi = tumor necrosis factor inhibitors.
† Missing data <15% in each visit.
‡ Missing data <20% in each visit.

(standardized rate range 0.043–0.126) (Table 3). Three or more fatty lesions on MRI-SI joints was the SI joints structural outcome with highest sensitivity to change (standardized rate 0.274; relative rate of 6.227 comparing to mNY). On the contrary, ≥3 erosions on MRI-SI joints was the least sensitive (standardized rate 0.015) of all SI joints structural outcomes (including both MRI-SI joints and radiography of the SI joints). Importantly, ≥3 fatty lesions alone was slightly more sensitive to change than combining fatty lesions with erosions, i.e., ≥5 fatty lesion and/or erosions (relative rate of 1.151 for the former compared to the latter).

Among the radiography of the SI joints structural outcomes, worsening of ≥1 grade in ≥1 SI joints and worsening of ≥1 grade in ≥1 SI joints, with a 5-year grade ≥2 in the worsened joint, were far more sensitive to change compared to the mNY binary definition as the modality reference (relative rate 2.864 and 2.705, respectively). Of note, the mNY continuous grading and the mNY binary score had comparable sensitivity to change (relative rate of the continuous versus the reference binary score = 0.977).

Overall, the standardized yearly rate of change of the spinal radiographic outcomes (range 0.037–0.043) was higher as compared to MRI-spine structural outcomes (range 0.012–0.027) (Table 3), although all are relatively low. Among MRI-spine outcomes, the total number of bone spurs was the outcome that most captured change (standardized rate 0.027; and relative rate of 2.077 compared to ≥5 fatty lesions, i.e., the modality reference). Yet, the best MRI-spine outcome is still less sensitive to change as compared to radiography of the spine outcomes, with

Table 2. Baseline score and standardized yearly rate of change (ROC) of inflammatory imaging outcomes over 5 years of follow-up in patients with early axial spondyloarthritis (SpA) fulfilling the Assessment of SpondyloArthritis international Society (ASAS) classification criteria for axial SpA*

Imaging outcomes	Baseline score (range 334–344)†	Standardized ROC per year‡	Relative standardized ROC§	Relative standardized ROC per modality and anatomic site
Inflammatory lesions (MRI of the SI joints)¶				
Sacroiliitis (ASAS criteria), no. (%)	134 (39.2)	–0.278#	1	1
SPARCC SI joint score (range 0–72)	4.7 ± 7.9	–0.441#	1.586	1.586
Inflammatory lesions (MRI of the spine)**				
BME ≥3 lesions, no. (%)	32 (9.4)	–0.032	0.319	1
BME ≥5 lesions, no. (%)	19 (5.6)	–0.030	0.094	0.938
23-DVU SPARCC spine score (range 0–414)	2.6 ± 7.7	–0.050	0.531	1.563
Berlin spine score (range 0–69)	0.9 ± 2.7	–0.055	0.104	1.719

* Values are the mean ± SD unless indicated otherwise. BME = bone marrow edema; DVU = discovertebral unit; MRI = magnetic resonance imaging; SI = sacroiliac; SPARCC = Spondyloarthritis Research Consortium of Canada.
† Agreement of ≥2 of 3 readers for binary variables and of 3 readers for continuous variables from wave 3.
‡ Estimated from a model in which all independent variables (time, reader, and wave) and the outcome were standardized.
§ Common reference: ASAS MRI of the SI joints.
¶ Refs. 9–11.
# Quadratic transformation led to a better model goodness of fit (quasi-likelihood under the independence model criterion).
** Refs. 3 and 12–14.

Table 3. Baseline score and standardized yearly rate of change (ROC) of structural imaging outcomes over 5 years of follow-up in patients with early axial spondyloarthritis (SpA) fulfilling the Assessment of SpondyloArthritis international Society (ASAS) classification criteria for axial SpA*

Imaging outcomes	Baseline score (range 313–344)†	Standardized ROC per year‡	Relative standardized ROC§	Relative standardized ROC per modality and anatomic site
Structural lesions (radiograph of the SI joints)¶				
mNY dichotomous, no. (%)	73 (21.2)	0.044	1	1
mNY 1-grade change#	NA	0.126	2.864	2.864
mNY 1-grade change and value ≥2**	NA	0.119	2.705	2.705
mNY continuous grade (range 0–8)	1.7 ± 1.8	0.043	0.977	0.977
Structural lesions (MRI of the SI joints)††				
≥5 fatty lesions and/or erosions, no. (%)	66 (19.5)	0.238‡‡	5.409	1
≥3 erosions, no. (%)	60 (17.7)	0.015	0.341	0.063
≥3 fatty lesions, no. (%)	56 (16.5)	0.274‡‡	6.227	1.151
No. of fatty lesions and/or erosions (range 0–80)	2.9 ± 4.9	0.111	2.523	0.466
No. of erosions (range 0–40)	1.3 ± 2.2	0.030	0.682	0.126
No. of fatty lesions (range 0–40)	1.5 ± 3.5	0.140	3.182	0.588
Total structural lesions (range 0–144)§§	3.4 ± 5.9	0.115	2.614	0.483
Total structural lesions without sclerosis (range 0–104)	3.2 ± 5.8	0.124	2.818	0.521
Structural lesions (radiograph of the spine)¶¶				
≥1 syndesmophyte, no. (%)	19 (5.5)	0.037	0.841	1
mSASSS score (range 0–72)	0.3 ± 1.3	0.043	0.977	1.162
Structural lesions (MRI of the spine)##				
≥5 fatty lesions, no. (%)	5 (1.6)	–0.013	0.295	1
Total structural lesions (range 0–322)***	0.4 ± 1.0	0.016	0.364	1.231
No. of fatty lesions (range 0–92)	0.3 ± 0.8	0.008	0.182	0.615
No. of corner erosions (range 0–92)	0.1 ± 0.2	0.012	0.273	0.923
No. of corner bone spurs (range 0–92)	0.1 ± 0.3	0.027	0.614	2.077

* Values are the mean ± SD unless indicated otherwise. mNY = modified New York criteria; MRI = magnetic resonance imaging; mSASSS = modified Stoke Ankylosing Spondylitis Spine Score; NA = not applicable; SI = sacroiliac.
† Agreement of ≥2 of 3 readers for binary variables and of 3 readers for continuous variables from wave 3.
‡ Estimated from a model in which all independent variables (time, reader, and wave) and the outcome were standardized.
§ Common reference: mNY.
¶ Refs. 15 and 16.
# Change of at least 1 grade in at least 1 SI joint.
** Change of at least 1 grade in at least 1 SI joint, but with a 5-year grade ≥2 in the worsened joint.
†† Ref. 18.
‡‡ Quadratic transformation led to a better model goodness of fit (quasi-likelihood under the independence model criterion).
§§ Fatty lesions, erosions, sclerosis, and partial ankylosis/total ankylosis.
¶¶ Ref. 19.
Refs. 20 and 21.
*** Erosions, fat infiltration, bone spurs, and ankylosis.

a standardized rate of 0.037 for ≥1 syndesmophyte and of 0.043 for the continuous mSASSS.

DISCUSSION

In this prospective observational study, we have shown that in patients with early axial SpA, MRI outcomes of inflammation are more sensitive to change in the SI joints than in the spine. In addition, pelvic radiographs yield low sensitivity to change in detecting structural damage, while fatty lesions detected on MRI-SI joints emerges as a promising alternative. In contrast, MRI-spine is not better than radiography of the spine in detecting structural changes in patients with early axial SpA.

In the current study, we directly compared, for the first time, inflammation outcomes on MRI-SI joints and MRI-spine and have shown that the former are more sensitive to change. Inflammation on MRI-spine remained low and essentially unchanged over a period of 5 years. Different from previous studies evaluating the sensitivity to change of imaging outcomes over shorter periods, we have applied an analytical technique (integrated analysis) that we have previously shown to be robust for the evaluation of change over long periods of follow-up, especially with outcomes that are expected to occur infrequently over time (8). Of note, combination algorithms (e.g., agreement between 2 of 3 readers) are not needed when using this method. Instead, each individual reader score is analyzed as it is in an assumption-free manner that, to some extent, handles across-reader variability.

The ASAS/OMERACT MRI working group has previously compared different (continuous) scores to quantify inflammation on MRI-SI joints (22). In a multireader exercise, the SPARCC method has been shown to be the most reliable and sensitive to change among patients with radiographic axial SpA. The current

study adds to these data by showing that both the continuous SPARCC score and the binary ASAS definition of a positive MRI-SI joint yield good sensitivity to change in the entire spectrum of axial SpA (including nonradiographic axial SpA) during the early phases of the disease.

The same group performed a similar exercise for MRI-spine (also in radiographic axial SpA) (3). This experiment has shown discrepant reliability results for the comparison between the 6-discovertebral unit (DVU) SPARCC score, the Ankylosing Spondylitis Spine MRI Activity score, and the Berlin method (SPARCC performed better when using the intraclass correlation coefficient but worse when using the smallest detectable change). All methods yielded excellent sensitivity to change according to Guyatt's effect size. Here, we compared the 23-DVU SPARCC score to the Berlin method and 2 binary outcomes and found that all yield very poor sensitivity to change. Of note, these studies differ in several aspects, including the reading methods and population. In fact, our early axial SpA population had lower baseline levels of inflammation compared to that in patients from the ASAS/OMERACT exercise (mean $\pm$ SD Berlin score 0.9 ± 2.7 versus 6 ± 9.0, respectively), which may hinder the detection of change, which we have shown before to be small in early axial SpA (17). Of note, in patients with nonradiographic axial SpA and high disease activity selected for randomized controlled trials, inflammation on MRI-spine performed well both in terms of sensitivity to change and in discriminating response between treatment arms (23,24). This confirms that the ability of the scoring methods to detect change is not only dependent on their intrinsic characteristics, but also on the population in which they are applied.

A recent study, also from the DESIR cohort, has shown that net progression from mNY negative to mNY positive (i.e., considering measurement error) is very limited (16). In the current study, we have additionally shown that the change in the mNY (continuous) grading is as poorly sensitive to change as the mNY binary score (relative rate of ~1). However, the change of at least 1 grade in at least 1 SI joint, with or without considering the change between grade 0 and grade 1, performs better in detecting change (16,25).

Information on the sensitivity to change of MRI-SI joints structural outcomes is very scarce (26). To the best of our knowledge, no previous formal comparison with radiography of the SI joints scores has been performed thus far. We have found that $\geq$3 fatty lesions on MRI-SI joints largely outperform all radiography of the SI joints outcomes. Erosions, however, performed poorly in this early population. Thus, our study yields encouraging data supporting MRI (in particular fatty lesions) as an alternative to radiographs in detecting change of structural damage at the SI joints. In contrast, in the spine, we found no evidence that MRI is better than radiographs in detecting change of structural damage. Despite the disappointing results with MRI, our results are in line with previous studies, showing that spinal radiographic progression can be detected even in early phases of the disease (4,27). A recent study has shown that low-dose computerized tomography of the spine is more sensitive at detecting new syndesmophytes than conventional radiographs, which promises to further expand our ability to detect change in axial damage (28).

Our study has some limitations. First, not all available scoring systems were assessed. However, to the best of our knowledge, this is, so far, the largest direct comparison across scores, which includes those currently more often used in research and clinical practice. Second, we did not assess all domains of the OMERACT filter, namely validity, reliability, and feasibility (2). Thus, we cannot, and do not claim to, evoke superiority of one score over others based on our data alone. Instead, our results should be interpreted in light of the literature already informing on these aspects but falling short on direct comparisons of sensitivity to change. Third, the observed levels of inflammation, structural damage, and changes over time are limited in this cohort, especially in the spine, which reduces the possibility of detecting differences across methods. Finally, our data are limited to patients with early axial SpA; thus, our findings cannot be generalized to all patients with axial SpA from clinical practice, especially those with more advanced disease (i.e., with radiographic axial SpA).

In conclusion, we have shown that MRI inflammation scores are more sensitive to change in the SI joints than in the spine. Also, radiography of the SI joints structural outcomes are less sensitive to change compared to fatty lesions on MRI-SI joints. In contrast, MRI-spine is no better than radiography of the spine in detecting structural changes in this early axial SpA cohort. These data may help in prioritizing imaging scoring methods in subsequent observational or interventional studies in early axial SpA.

ACKNOWLEDGMENTS

The authors thank the different regional participating centers: Maxime Dougados (Paris-Cochin B), Andre Kahan (Paris-Cochin A), Philippe Dieudé (Paris-Bichat), Bruno Fautrel (Paris-La Pitie-Salpetriere), Francis Berenbaum (Paris-Saint-Antoine), Pascal Claudepierre (Creteil), Maxime Breban (Boulogne-Billancourt), Bernadette Saint-Marcoux (Aulnay-sous-Bois), Philippe Goupille (Tours), Jean Francis Maillefert (Dijon), Emmanuelle Dernis (Le Mans), Daniel Wendling (Besancon), Bernard Combe (Montpellier), Liana Euller-Ziegler (Nice), Pascal Richette (Paris Lariboisiere), Pierre Lafforgue (Marseille), Patrick Boumier (Amiens), Martin Soubrier (Clermont Ferrand), Nadia Mehsen (Bordeaux), Damien Loeuille (Nancy), Rene-Marc Flipo (Lille), Alain Saraux (Brest), Xavier Mariette (LeKremlin-Bicetre), Alain Cantagrel (Toulouse), and Olivier Vittecoq (Rouen). They also thank the research nurses, the staff members of the Clinical Research Unit of Paris Centre, the staff members of the Biological Resource Center of Bichat Hospital, the staff members of the Department of Statistics of Nîmes, and all the investigators, and in particular Jerome Allain, Emmanuelle Dernis, Salah Ferkal, Clement Prati, Marie-Agnes Timsit, and Eric Toussirot, for active patient recruitment and monitoring. The authors are grateful to Anna Moltó and Miranda van Lunteren for their work in preparing the database used in this study; and to all DESIR readers, without whom this study would not be possible.

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. Sepriano 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. Sepriano, Ramiro, van der Heijde, Dougados, Landewé.

Acquisition of data. Claudepierre, Feydy, Reijnierse, Loeuille.

Analysis and interpretation of data. Sepriano, Ramiro, van der Heijde, Dougados, Landewé.

REFERENCES

1. Mandl P, Navarro-Compan V, Terslev L, Aegerter P, van der Heijde D, D'Agostino MA, et al. EULAR recommendations for the use of imaging in the diagnosis and management of spondyloarthritis in clinical practice. Ann Rheum Dis 2015;74:1327–39.
2. Boers M, Kirwan JR, Wells G, Beaton D, Gossec L, d'Agostino MA, et al. Developing core outcome measurement sets for clinical trials: OMERACT filter 2.0. J Clin Epidemiol 2014;67:745–53.
3. Lukas C, Braun J, van der Heijde D, Hermann KG, Rudwaleit M, Ostergaard M, et al. Scoring inflammatory activity of the spine by magnetic resonance imaging in ankylosing spondylitis: a multireader experiment. J Rheumatol 2007;34:862–70.
4. Ramiro S, van Tubergen A, Stolwijk C, Landewé R, van de Bosch F, Dougados M, et al. Scoring radiographic progression in ankylosing spondylitis: should we use the modified Stoke Ankylosing Spondylitis Spine Score (mSASSS) or the Radiographic Ankylosing Spondylitis Spinal Score (RASSS)? Arthritis Res Ther 2013;15:R14.
5. Wanders AJ, Landewé RB, Spoorenberg A, Dougados M, van der Linden S, Mielants H, et al. What is the most appropriate radiologic scoring method for ankylosing spondylitis? A comparison of the available methods based on the Outcome Measures in Rheumatology Clinical Trials filter. Arthritis Rheum 2004;50:2622–32.
6. Ramiro S, Claudepierre P, Sepriano A, van Lunteren M, Molto A, Feydy A, et al. Which scoring method depicts spinal radiographic damage in early axial spondyloarthritis best? Five-year results from the DESIR cohort. Rheumatology (Oxford) 2018;57:1991–2000.
7. Dougados M, d'Agostino MA, Benessiano J, Berenbaum F, Breban M, Claudepierre P, 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.
8. Sepriano A, Ramiro S, van der Heijde D, Dougados M, Claudepierre P, Feydy A, et al. Integrated longitudinal analysis does not compromise precision and reduces bias in the study of imaging outcomes: a comparative 5-year analysis in the DESIR cohort. Semin Arthritis Rheum 2020;50:1394–9.
9. Maksymowych WP, Inman RD, Salonen D, Dhillon SS, Williams M, Stone M, et al. Spondyloarthritis Research Consortium of Canada magnetic resonance imaging index for assessment of sacroiliac joint inflammation in ankylosing spondylitis. Arthritis Rheum 2005;53:703–9.
10. Rudwaleit M, Jurik AG, Hermann KG, Landewé R, van der Heijde D, Baraliakos X, et al. Defining active sacroiliitis on magnetic resonance imaging (MRI) for classification of axial spondyloarthritis: a consensual approach by the ASAS/OMERACT MRI group. Ann Rheum Dis 2009;68:1520–7.
11. Lambert RG, Bakker PA, van der Heijde D, Weber U, Rudwaleit M, Hermann KG, et al. Defining active sacroiliitis on MRI for classification of axial spondyloarthritis: update by the ASAS MRI working group. Ann Rheum Dis 2016;75:1958–63.
12. Hermann KG, Baraliakos X, van der Heijde DM, Jurik AG, Landewé R, Marzo-Ortega H, et al. Descriptions of spinal MRI lesions and definition of a positive MRI of the spine in axial spondyloarthritis: a consensual approach by the ASAS/OMERACT MRI study group. Ann Rheum Dis 2012;71:1278–88.
13. De Hooge M, van den Berg R, Navarro-Compan V, Reijnierse M, van Gaalen F, Fagerli K, et al. Patients with chronic back pain of short duration from the SPACE cohort: which MRI structural lesions in the sacroiliac joints and inflammatory and structural lesions in the spine are most specific for axial spondyloarthritis? Ann Rheum Dis 2016;75:1308–14.
14. Maksymowych WP, Inman RD, Salonen D, Dhillon SS, Krishnananthan R, Stone M, et al. Spondyloarthritis Research Consortium of Canada magnetic resonance imaging index for assessment of spinal inflammation in ankylosing spondylitis. Arthritis Rheum 2005;53:502–9.
15. Van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis: a proposal for modification of the New York criteria. Arthritis Rheum 1984;27:361–8.
16. Dougados M, Sepriano A, Molto A, van Lunteren M, Ramiro S, de Hooge M, 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.
17. Madari Q, Sepriano A, Ramiro S, Molto A, Claudepierre P, Wendling D, et al. 5-year follow-up of spinal and sacroiliac MRI abnormalities in early axial spondyloarthritis: data from the DESIR cohort. RMD Open 2020;6:e001093.
18. Weber U, Lambert RG, Østergaard M, Hodler J, Pedersen SJ, Maksymowych WP. The diagnostic utility of magnetic resonance imaging in spondylarthritis: an international multicenter evaluation of one hundred eighty-seven subjects. Arthritis Rheum 2010;62:3048–58.
19. Creemers MC, Franssen MJ, van't Hof MA, Gribnau FW, van de Putte LB, van Riel PL. Assessment of outcome in ankylosing spondylitis: an extended radiographic scoring system. Ann Rheum Dis 2005;64:127–9.
20. Ostergaard M, Maksymowych WP, Pedersen SJ, Chiowchanwisawakit P, Lambert RG. Structural lesions detected by magnetic resonance imaging in the spine of patients with spondyloarthritis: definitions, assessment system, and reference image set. J Rheumatol 2009;36 Suppl 84:18–34.
21. Krabbe S, Sorensen IJ, Jensen B, Moller JM, Balding L, Madsen OR, et al. Inflammatory and structural changes in vertebral bodies and posterior elements of the spine in axial spondyloarthritis: construct validity, responsiveness and discriminatory ability of the anatomy-based CANDEN scoring system in a randomised placebo-controlled trial. RMD Open 2018;4:e000624.
22. Van der Heijde DM, Landewé RB, Hermann KG, Jurik AG, Maksymowych WP, Rudwaleit M, et al. Application of the OMERACT filter to scoring methods for magnetic resonance imaging of the sacroiliac joints and the spine: recommendations for a research agenda at OMERACT 7. J Rheumatol 2005;32:2042–7.
23. Maksymowych WP, Dougados M, van der Heijde D, Sieper J, Braun J, Citera G, et al. Clinical and MRI responses to etanercept in early non-radiographic axial spondyloarthritis: 48-week results from the EMBARK study. Ann Rheum Dis 2016;75:1328–35.
24. Sieper J, van der Heijde D, Dougados M, Mease PJ, Maksymowych WP, Brown MA, et al. Efficacy and safety of adalimumab in patients with non-radiographic axial spondyloarthritis: results of a randomised placebo-controlled trial (ABILITY-1). Ann Rheum Dis 2013;72:815–22.
25. Poddubnyy D, Rudwaleit M, Haibel H, Listing J, Marker-Hermann E, Zeidler H, et al. Rates and predictors of radiographic sacroiliitis progression over 2 years in patients with axial spondyloarthritis. Ann Rheum Dis 2011;70:1369–74.
26. Baraliakos X, van der Heijde D, Braun J, Landewé RB. OMERACT magnetic resonance imaging initiative on structural and inflammatory lesions in ankylosing spondylitis: report of a special interest group at OMERACT 10 on sacroiliac joint and spine lesions. J Rheumatol 2011;38:2051–4.

27. Cho SK, Sakai R, Nanki T, Koike R, Watanabe K, Yamazaki H, et al. A comparison of incidence and risk factors for serious adverse events in rheumatoid arthritis patients with etanercept or adalimumab in Korea and Japan. Mod Rheumatol 2014;24: 572–9.

28. De Koning A, de Bruin F, van den Berg R, Ramiro S, Baraliakos X, Braun J, et al. Low-dose CT detects more progression of bone formation in comparison to conventional radiography in patients with ankylosing spondylitis: results from the SIAS cohort. Ann Rheum Dis 2018;77:293–9.

Arthritis Care & Research
Vol. 74, No. 2, February 2022, pp 259–267
DOI 10.1002/acr.24442

American College of Rheumatology
Empowering Rheumatology Professionals

Performance and Predictors of Minimal Disease Activity Response in Patients With Peripheral Spondyloarthritis Treated With Adalimumab

Laura C. Coates,[1] Sonya Abraham,[2] William Tillett,[3] Philip J. Mease,[4] Sofia Ramiro,[5] Tianshuang Wu,[6] Xin Wang,[6] Aileen L. Pangan,[6] and In-Ho Song[6]

Objective. To examine the concurrent validity and discrimination of criteria for modified minimal disease activity (MDA) in peripheral spondyloarthritis (SpA) following filter principles of Outcome Measures in Rheumatology (OMERACT) and to determine predictors of modified MDA response.

Methods. Four modified MDA versions were derived in the ABILITY-2 study using the Spondyloarthritis Research Consortium of Canada (SPARCC) enthesitis index or the Leeds Enthesitis Index (LEI) while excluding psoriasis. To assess concurrent validity, modified MDA versions were correlated with Peripheral Spondyloarthritis Response Criteria (PSpARC) remission, Ankylosing Spondylitis Disease Activity Score showing inactive disease (ASDAS ID), and physician global assessment of disease activity. Treatment discrimination was assessed between adalimumab and placebo at week 12. Multiple logistic regression was used to determine baseline predictors of long-term modified MDA responses and sustained modified MDA.

Results. The 4 modified MDA versions showed a stronger positive correlation with PSpARC remission ($r_{tet} > 0.95$) versus ASDAS ID ($r_{tet} > 0.75$) at week 12 and years 1–3 and were able to show discrimination ($P < 0.001$). Responsiveness was shown at week 12; significantly more patients receiving adalimumab versus placebo achieved all 4 versions of modified MDA. Approximately 40–60% of patients treated with adalimumab achieved modified MDA using the LEI or SPARCC enthesitis index at years 1–3. Achieving modified MDA response after 12 weeks of adalimumab treatment was a robust positive predictor of attaining long-term modified MDA through 3 years (odds ratio [OR] 11.38–27.13 for modified MDA using the LEI; OR 17.98–37.85 for modified MDA using the SPARCC enthesitis index).

Conclusion. All 4 versions of modified MDA showed concurrent validity and discriminated well between adalimumab and placebo treatment groups. Early modified MDA response is a more consistent predictor of long-term modified MDA achievement than baseline characteristics. The 5 of 6 versions of modified MDA could be an appropriate treatment target in patients with peripheral SpA.

INTRODUCTION

Peripheral spondyloarthritis (SpA) encompasses patients with predominantly peripheral symptoms such as peripheral arthritis, dactylitis, and/or enthesitis (1,2). To date, most studies in the field of peripheral SpA have focused on psoriatic arthritis (PsA). Relatively few outcome measures have been developed specifically for nonpsoriatic peripheral SpA. Due to a lack of

ClinicalTrials.gov identifier: NCT01064856.

The views expressed herein are those of the authors and not necessarily those of the National Health Service, the National Institute for Health Research, or the Department of Health.

Supported by AbbVie and the National Institute for Health Research and Oxford Biomedical Research Centre. Dr. Coates' work was supported by the National Institute for Health Research (Clinician Scientist award).

[1]Laura C. Coates, MBChB, PhD: University of Oxford, Oxford, UK; [2]Sonya Abraham, MD: National Institute for Health Research/Wellcome CRF, Imperial College Healthcare NHS Trust, London, UK; [3]William Tillett, MBChB, BSc, PhD: Royal National Hospital for Rheumatic Diseases, Pharmacy and Pharmacology, University of Bath, Bath, UK; [4]Philip J. Mease, MD: Swedish Medical Center/Providence St. Joseph Health and University of Washington, Seattle; [5]Sofia Ramiro, MD: Leiden University Medical Center, Leiden, and Zuyderland Medical Center, Heerlen, The Netherlands; [6]Tianshuang Wu, PhD, Xin Wang, PhD, Aileen L. Pangan, MD, In-Ho Song, MD: AbbVie, North Chicago, Illinois.

Dr. Coates has received consulting fees and/or speaking fees from Amgen, Bristol Myers Squibb, Celgene, Galapagos, Gilead, Eli Lilly and Company, Janssen, MSD, Pfizer, Prothena, Sun Pharma, UCB (less than $10,000 each), AbbVie, and Novartis (more than $10,000 each) and research grants from AbbVie, Celgene, Novartis, and Pfizer. Dr. Abraham has received consulting fees and/or speaking fees from AbbVie, Celgene, Novartis, Pfizer, and UCB (less than $10,000 each) and research grants from those companies. Dr. Tillett has received consulting fees and/or speaking fees from AbbVie, Celgene, Eli Lilly and Company, Janssen, Novartis, Pfizer, and UCB (less than $10,000 each) and research grants from those companies. Dr. Mease has received consulting fees and/or speaking fees from AbbVie, Bristol Myers Squibb, Boehringer Ingelheim, Celgene, Galapagos, Gilead, GlaxoSmithKline,

SIGNIFICANCE & INNOVATIONS

- In patients with peripheral spondyloarthritis (SpA), using 4 modified minimal disease activity (MDA; excluding psoriasis) versions following aspects of the Outcome Measures in Rheumatology (OMERACT) filter criteria, the modified MDA version that used either of the psoriatic arthritis (PsA)–validated entheseal indices (Leeds Enthesitis Index and Spondyloarthritis Research Consortium of Canada enthesitis index) discriminated well between adalimumab and placebo treatment groups.
- Similar to MDA definitions used in PsA, the data presented here support the concurrent validity and discrimination of modified MDA using either of the entheseal indices, depending on physician preference.
- Early modified MDA response is a more consistent predictor of long-term modified MDA achievement than baseline characteristics, and identification of factors that predict long-term modified MDA response in peripheral SpA patients would help to facilitate treatment decisions.

validated outcome measures in nonpsoriatic peripheral SpA, recent studies have used varying outcome measures such as improvement in the patient global assessment of disease activity (PtGA) used in the TIPES trial (3), ≥40% improvement in Peripheral Spondyloarthritis Response Criteria (PSpARC40) developed as a novel primary end point in the ABILITY-2 trial (4), or clinical remission, defined as absence of peripheral arthritis, enthesitis, and dactylitis used in the golimumab CRESPA trial (5).

The discriminatory capacity of different outcome measures was evaluated in patients with peripheral SpA from the TIPES and ABILITY-2 studies (6). Although most of the outcome measures used in studies of peripheral SpA distinguished between active treatment and placebo, not all of the relevant disease manifestations of peripheral SpA were fully captured. Therefore, it may be worthwhile to develop and validate alternative peripheral SpA–specific composite indices that better capture relevant disease aspects of peripheral SpA.

The minimal disease activity (MDA) measure was developed and validated in patients with PsA to define a specific disease activity state and has been validated in interventional clinical trials and observational studies and recommended as a treatment target in patients with PsA (7–12). However, the potential applicability of MDA to other forms of peripheral SpA has not yet been established. If the measure shows validity in defining a disease state in peripheral SpA, identification of factors that predict long-term modification of the MDA response in patients with peripheral SpA would help to facilitate decisions regarding treatment initiation and maintenance.

The purpose of this analysis was to examine the concurrent validity and discrimination of modified MDA criteria (excluding psoriasis) following aspects of the Outcome Measures in Rheumatology (OMERACT) filter (including truth and discrimination) and to identify predictors of long-term modified MDA response following treatment with adalimumab in patients with peripheral SpA included in the ABILITY-2 study (13).

PATIENTS AND METHODS

Patient population. Results from ABILITY-2 (ClinicalTrials.gov identifier: NCT01064856), a phase 3, randomized, double-blind, placebo-controlled study, were reported previously (4,14). Briefly, ABILITY-2 included adult patients (≥18 years) with peripheral SpA who fulfilled the Assessment of SpondyloArthritis international Society criteria for peripheral SpA, with symptom onset at least 3 months prior to study entry (1). To avoid overlap in patient populations, patients with a history of psoriasis or PsA or ankylosing spondylitis (AS) were excluded from the study. Patients were randomized 1:1 to receive adalimumab 40 mg or placebo every other week during the 12-week placebo-controlled period, followed by an open-label extension during which they received open-label adalimumab for up to 144 weeks. Patients underwent a total of 16 visits during the open-label period of the study (14). In this analysis, data from the ABILITY-2 study were used to examine the concurrent validity and discrimination of modified versions of MDA and to identify predictors of long-term modified MDA following adalimumab treatment.

Outcome measures. The original MDA measure for PsA is defined as achieving ≥5 of the following 7 criteria: tender joint count of 78 joints (TJC78) of ≤1; swollen joint count of 76 joints

Genentech, Merck, Sun Pharma, and UCB (less than $10,000 each), Amgen, Janssen, Eli Lilly and Company, Novartis, and Pfizer (more than $10,000 each) and research grants from those companies. Dr. Ramiro has received consulting fees and/or speaking fees from AbbVie, Eli Lilly and Company, MSD, Novartis, Sanofi, and UCB (less than $10,000 each) and research grants from MSD. Drs. Wu, Wang, Pangan, and Song own stock or stock options in AbbVie. No other disclosures relevant to this article were reported.

AbbVie is committed to responsible data sharing regarding the clinical trials we sponsor. This includes access to anonymized, individual, and trial-level data (analysis data sets), as well as other information (e.g., protocols and clinical study reports), as long as the trials are not part of an ongoing or planned regulatory submission. This includes requests for clinical trial data for unlicensed products and indications. These clinical trial data can be requested by any qualified researchers who engage in rigorous, independent scientific research and will be provided following review and approval of a research proposal and Statistical Analysis Plan (SAP) and execution of a Data Sharing Agreement (DSA). Data requests can be submitted at any time, and the data will be accessible for 12 months, with possible extensions considered. For more information on the process, or to submit a request, visit the following link: https://www.abbvie.com/our-science/clinical-trials/clinical-trials-data-and-information-sharing/data-and-information-sharing-with-qualified-researchers.html.

Address correspondence to Laura C. Coates, MBChB, PhD, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, B4495, Headington, Oxford OX3 7LD, UK. Email: laura.coates@ndorms.ox.ac.uk.

Submitted for publication May 6, 2020; accepted in revised form September 1, 2020.

(SJC76) of ≤1; Psoriasis Area and Severity Index (PASI) score of ≤1; patient assessment of pain score of ≤15 on a visual analog scale (VAS) (0–100 mm); PtGA score of ≤20 on a VAS (0–100 mm); Health Assessment Questionnaire disability index (HAQ DI) score of ≤0.5; and ≤1 tender entheseal points (assessed bilaterally at 2 sites) (7).

Skin outcome measures (using the PASI or body surface area), however, were not included in ABILITY-2, as patients with psoriasis or PsA were excluded. Hence, the MDA criteria were modified for the nonpsoriatic, peripheral SpA population by removing the psoriasis skin component (i.e., the PASI score), and 2 modifications were tested defined as achieving at least 4 or 5 of the 6 modified MDA components mentioned above, but excluding the PASI score. A previous study showed that the Leeds Enthesitis Index (LEI) and the Spondyloarthritis Research Consortium of Canada (SPARCC) enthesitis index were able to better discriminate between adalimumab and placebo treatment responses compared with the Maastricht Ankylosing Spondylitis Enthesitis Score in patients with peripheral SpA (15), but it has not been investigated whether the SPARCC enthesitis index or the LEI instrument is better in peripheral SpA. Thus, enthesitis was assessed by either the LEI or SPARCC enthesitis index in this analysis. The different enthesitis measures are described in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24442/abstract. To summarize, at each time point, the following 4 versions of modified MDA were evaluated: modified MDA 4 of 6 (LEI) (achieving 4 of 6 modified MDA components, and use of the LEI); modified MDA 5 of 6 (LEI) (achieving 5 of 6 modified MDA components); modified MDA 4 of 6 (SPARCC) (use of the SPARCC enthesitis index); and modified MDA 5 of 6 (SPARCC).

Remission according to the Peripheral SpA Response Criteria (PSpARC) and remission according to the Ankylosing Spondylitis Disease Activity Score showing inactive disease (ASDAS ID; ASDAS score <1.3) were used as outcome measures (4,16). Disease remission based on the PSpARC was defined as achieving an SJC of ≤1 and ≥4 of following 5 criteria: PtGA score of ≤20; patient assessment of pain score of ≤20; TJC78 of ≤1; an enthesitis count (based on 29 enthesitis sites) of ≤1; and a dactylitis count of ≤1 (4).

Statistical methods. Given that there is no true gold standard for disease control in peripheral SpA, criterion validity could not be assessed. Instead, concurrent validity was assessed by correlation of the 4 versions of the modified MDA with related disease outcome measures. Correlation analyses were therefore performed using PSpARC remission and ASDAS ID by tetrachoric correlation (r_{tet}) applicable for binary outcomes (17); in addition, the correlation between the 4 versions of modified MDA and physician global assessment of disease activity (PhGA) was evaluated by point-biserial correlation (r_{pb}) applicable for 1 continuous and 1 binary outcome (18).

The discriminatory ability of all 4 versions of modified MDA, PSpARC remission, and ASDAS ID comparing adalimumab treatment versus placebo at week 12 was assessed using Pearson's χ^2 test (higher score meaning better discrimination). To differentiate between the 4 candidate modified MDA measures and to examine their threshold of meaning, residual disease in different domains despite achieving these candidate targets was analyzed.

Patients who received at least 12 weeks of adalimumab treatment during the placebo-controlled period or open-label extension with data available after 12 weeks of adalimumab exposure and at years 1, 2, and 3 were included in this analysis (data are presented as observed). The number and proportion of patients achieving a modified MDA response over time was calculated.

Multiple logistic regression with stepwise variable selection was used to determine predictors of long-term, 5 of 6 modified MDA responses at year 1, year 2, and year 3, respectively, and sustained modified MDA response at any time point (defined as achieving modified MDA for at least 24 consecutive weeks (14). Two sets of candidates predicting variables were considered in the model selection: baseline patient demographic and disease characteristics alone, and also the same potential predictors along with modified MDA response after 12 weeks of adalimumab exposure (modified MDA12). Candidate baseline variables included in the analysis were age, sex, duration of peripheral SpA (in years), HLA–B27 status (positive versus negative), high-sensitivity C-reactive protein (hs-CRP) status (elevated versus normal), prior treatment with disease-modifying antirheumatic drugs (yes versus no), TJC78, SJC76, enthesitis count (0–29), dactylitis count (0–20), Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) score, ASDAS, PtGA (0–100 mm VAS), PhGA (0–100 mm VAS), and treatment groups (adalimumab versus placebo).

RESULTS

Patients enrolled in ABILITY-2 had generally comparable demographic and baseline disease characteristics except for mean age, which was higher in the adalimumab group, and the percentage of patients with a dactylitis count of >1, which was lower in the adalimumab group versus placebo (4). Patients with peripheral SpA receiving adalimumab achieved significantly greater clinical responses compared with placebo at week 12, and the efficacy was maintained over 3 years (4,14).

Correlation between modified MDA and other outcome measures for peripheral SpA. To explore the relationship between modified MDA and outcome measures for peripheral SpA used in prior clinical trials, response rates were compared at week 12 and years 1–3. All 4 versions of modified MDA response showed a stronger positive correlation with PSpARC remission ($r_{tet} > 0.95$) compared with ASDAS ID

(r_{tet} > 0.75) at week 12 and years 1–3 (Table 1). There was a moderate negative correlation between the 4 versions of modified MDA response and PhGA (r_{pb} = 0.43–0.61) at week 12 and years 1–3. However, the 4 versions of modified MDA did not correlate with CRP levels (data not shown). Correlation with PtGA was not performed as PtGA is part of the MDA measure.

Achievement of modified MDA over 3 years. Among 163 patients (82 receiving adalimumab, 81 receiving placebo) who completed week 12 of the double-blind period of ABILITY-2, a significantly greater proportion of patients receiving adalimumab achieved modified MDA (regardless of the definition) compared with placebo (P < 0.001 for all comparisons) (Figure 1). At week 12, 40.2%, 28.0%, 35.4%, and 26.8% of patients treated with adalimumab achieved modified MDA based on LEI 4 of 6, LEI 5 of 6, SPARCC 4 of 6, and SPARCC 5 of 6, respectively, compared with 13.6%, 4.9%, 12.3%, and 4.9% of patients treated with placebo.

Among patients who achieved modified MDA at week 12, 48.5% (LEI 4 of 6; n = 16), 69.6% (LEI 5 of 6; n = 16), 48.3% (SPARCC 4 of 6; n = 14), and 63.6% (SPARCC 5 of 6; n = 14) of patients treated with adalimumab attained all 6 components of modified MDA compared with only 18.2% (n = 2), 50.0% (n = 2), 20% (n = 2), and 50.0% (n = 2) of patients treated with placebo.

During the open-label extension, the proportion of patients who achieved modified MDA (as observed) were maintained over 3 years across all 4 versions (see Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24442/abstract). However, the rate of modified MDA achievement, including achievement of sustained modified MDA, was numerically higher at every time point among patients initially randomized to receive adalimumab compared with patients switching from placebo to adalimumab.

Table 1. Correlation between modified minimal disease activity (MDA) and peripheral spondylitis outcome measures*

Modified MDA	Week 12 (DB)	Year 1 (OLE)	Year 2 (OLE)	Year 3 (OLE)
4 of 6 LEI components				
PSpARC remission†	0.99‡	0.99	0.96	0.99
ASDAS ID†	0.82	0.79	0.89	0.86
PhGA§	–0.50	–0.51	–0.51	–0.61
5 of 6 LEI components				
PSpARC remission†	0.98	0.98	0.97	0.99‡
ASDAS ID†	0.82	0.80	0.88	0.84
PhGA§	–0.44	–0.47	–0.54	–0.55
4 of 6 SPARCC components				
PSpARC remission†	0.99	0.99	0.96	0.99
ASDAS ID†	0.76	0.76	0.89	0.85
PhGA§	–0.49	–0.55	–0.54	–0.58
5 of 6 SPARCC components				
PSpARC remission†	0.97	0.98	0.97	0.99
ASDAS ID†	0.80	0.79	0.86	0.83
PhGA§	–0.43	–0.46	–0.50	–0.54

* ASDAS ID = Ankylosing Spondylitis Disease Activity Score showing inactive disease; DB = double-blind period; LEI = Leeds Enthesitis Index; OLE = open-label extension; PhGA = physician global assessment of disease activity; PSpARC = Peripheral Spondyloarthritis Response Criteria; SPARCC = Spondyloarthritis Research Consortium of Canada enthesitis index.
† Tetrachoric correlation.
‡ Value >0.99.
§ Point-biserial correlation.

Ability of modified MDA for detection of treatment effect. All 4 versions of modified MDA had numerically higher Pearson's χ^2 values compared to PSpARC remission and the ASDAS ID. All 4 versions showed the ability to detect significant efficacy difference between adalimumab and placebo treatment groups (Table 2).

Individual modified MDA components and criteria not met. To establish the threshold of meaning of the measure, the individual modified MDA components that were not achieved in the patients treated with adalimumab achieving each version of modified MDA were assessed to better understand their contribution to the overall response. Among patients receiving adalimumab who fulfilled 4 of 6 modified MDA criteria (LEI or SPARCC) at week 12, the joint count criteria were most frequently not met (TJC: LEI 4 of 6 [27.3%]; SPARCC 4 of 6 [20.7%]; SJC: LEI 4 of 6 [30.3%]; SPARCC 4 of 6 [24.1%]) (Table 3). However, the 5 of 6 criteria (LEI or SPARCC) were more stringent, with approximately only 4% and 13% not meeting the TJC or SJC criterion, respectively. Approximately 5–10% of patients treated with adalimumab achieving modified MDA did not attain the HAQ DI criterion. For enthesitis, there was no big difference between the 4 of 6 or the 5 of 6 modified MDA versions; but using the LEI versions of modified MDA, patients treated with adalimumab all attained the enthesitis criterion compared with 9% and 10% of patients not meeting this criterion using the modified MDA versions and the SPARCC enthesitis index, respectively.

In patients treated with adalimumab achieving 4 of 6 modified MDA criteria (LEI or SPARCC), the mean TJC and SJC were at or above the modified MDA cutoff (see Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24442/abstract), while all other components were lower than the modified MDA cutoff. The 5 of 6 modified MDA criteria were more stringent, with the mean of each MDA component lower than the modified MDA cutoff. Among patients with peripheral SpA who did not achieve modified MDA at week 12, 83–88% of patients treated with adalimumab achieved 1 of 6 modified MDA criteria (LEI or SPARCC), while the rate was 66–79% in patients treated with placebo (see Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24442/abstract).

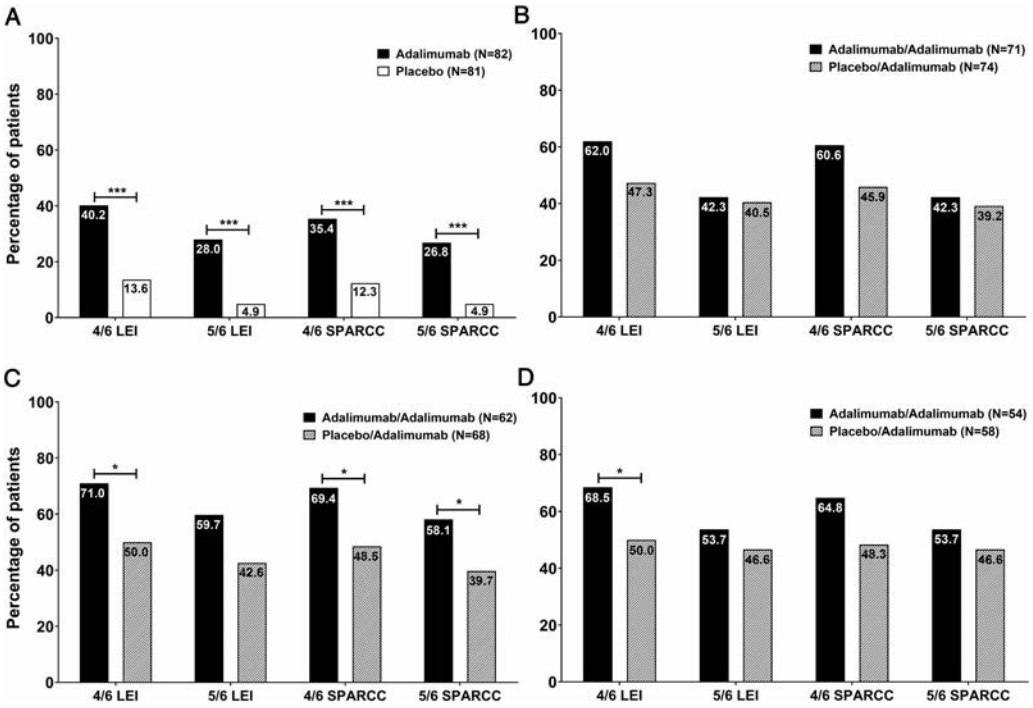

Figure 1. Achievement of modified minimal disease activity (MDA) (as observed) at week 12 (**A**), and years 1 (**B**), 2 (**C**), and 3 (**D**) in patients receiving adalimumab, placebo, or switching from placebo to adalimumab at week 12 using 4 definitions of minimal MDA. *P* values for difference between adalimumab and placebo or placebo/adalimumab treatment groups: * = $P < 0.05$; *** = $P < 0.001$. LEI = Leeds Enthesitis Index; SPARCC = Spondyloarthritis Research Consortium of Canada enthesitis index.

Predictors of long-term modified MDA response. Multiple logistic regression analyses were performed for the more stringent 5 of 6 modified MDA definitions based on better face validity and identified modified MDA response after 12 weeks of adalimumab exposure (modified MDA12) as a strong and robust positive predictor of attaining both long-term modified MDA at years 1–3 and sustained modified MDA ($P < 0.001$ for the 5 of 6 modified MDA definitions, using either the LEI [odds ratio (OR) range 11.38–27.13] or the SPARCC enthesitis index [OR range 17.98–37.85] at years 1–3 and sustained over time) (Figure 2). In contrast, baseline BASDAI score was a consistent negative predictor of modified MDA achievement at years 1–3 and sustained over time, irrespective of the inclusion of the LEI (OR range 0.36–0.66) or the SPARCC enthesitis index (OR range 0.51–0.68) in the modified MDA definition (Figure 2). Although prior DMARD use and baseline PhGA score were selected as positive predictors, and baseline enthesitis was selected as a negative predictor, these variables did not consistently predict achievement of modified MDA at every time point or sustained over time and were only weakly associated. The ASDAS showed positive association with achievement of modified MDA at year 3.

An analysis excluding the ASDAS was performed, as the ASDAS and BASDAI score are highly correlated outcomes. The analysis showed that both modified MDA12 and BASDAI score were still significantly associated with modified MDA achievement at year 3 (see Supplementary Figure 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24442/abstract). In addition, hs-CRP level showed a positive association with achievement of modified MDA at year

Table 2. Discrimination between adalimumab and placebo treatment at week 12*

Outcome measure	Adalimumab	Placebo	Pearson's χ^2	*P*
Modified MDA, 4 of 6 LEI components	33/82 (40.2)	11/81 (13.6)	14.70	<0.001
Modified MDA, 5 of 6 LEI components	23/82 (28.0)	4/81 (4.9)	15.75	<0.001
Modified MDA, 4 of 6 SPARCC components	29/82 (35.4)	10/81 (12.3)	11.86	<0.001
Modified MDA, 5 of 6 SPARCC components	22/82 (26.8)	4/81 (4.9)	14.57	<0.001
PSpARC remission	33/81 (40.7)	16/80 (20.0)	8.18	0.004
ASDAS ID	27/80 (33.8)	12/78 (15.4)	7.17	0.007

* Values are the no./total no. (%) unless indicated otherwise. ASDAS ID = Ankylosing Spondylitis Disease Activity Score showing inactive disease; LEI = Leeds Enthesitis Index; MDA = minimal disease activity; PSpARC = Peripheral Spondyloarthritis Response Criteria; SPARCC = Spondyloarthritis Research Consortium of Canada enthesitis index.

Table 3. Criteria not met in achievers of modified minimal disease activity (MDA) receiving adalimumab at week 12*

Modified MDA	TJC78 ≤ 1	SJC76 ≤ 1	Patient pain VAS score ≤ 15†	PtGA VAS score ≤ 20	HAQ DI score ≤ 0.5	Enthesitis index score ≤ 1
4 of 6 LEI components (n = 33)	9 (27.3)	10 (30.3)	2 (6.1)	3 (9.1)	3 (9.1)	0 (0.0)
5 of 6 LEI components (n = 23)	1 (4.3)	3 (13.0)	0 (0.0)	1 (4.3)	2 (8.7)	0 (0.0)
4 of 6 SPARCC components (n = 29)	6 (20.7)	7 (24.1)	1 (3.4)	2 (6.9)	3 (10.3)	3 (10.3)
5 of 6 SPARCC components (n = 22)	1 (4.5)	3 (13.6)	0 (0.0)	1 (4.5)	1 (4.5)	2 (9.1)

* Values are the number (%). HAQ DI = Health Assessment Questionnaire disability index (based on 20 questions); LEI = Leeds Enthesitis Index; PtGA = patient global assessment of disease activity; SJC76 = swollen joint count of 76 joints; SPARCC = Spondyloarthritis Research Consortium of Canada enthesitis index; TJC78 = tender joint count of 78 joints; VAS = visual analog scale.
† Patient pain = patient global assessment of pain.

3; elevated baseline hs-CRP was associated with increased likelihood of achieving modified MDA.

In the model examining the baseline variables alone (model without modified MDA12), age, enthesitis, and BASDAI scores were most commonly selected as negative predictors for achieving long-term response over 3 years and sustained modified MDA (Figure 3). Baseline PhGA, hs-CRP level, and male sex were selected as positive predictors, and dactylitis was selected as negative predictor; however, these predictors were not consistently selected for all time points or sustained modified MDA.

The modified MDA response rates (probability of achieving modified MDA response) at years 1–3 and sustained over time in patients who achieved modified MDA after 12 weeks of adalimumab exposure compared with patients who did not achieve modified MDA12 are shown in Supplementary Table 4, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24442/abstract. The response rate was consistent with the model selection results, indicating that patients who achieved modified MDA after 12 weeks of adalimumab exposure were 80–90% more likely to achieve modified MDA response at years 1–3 and sustained over time.

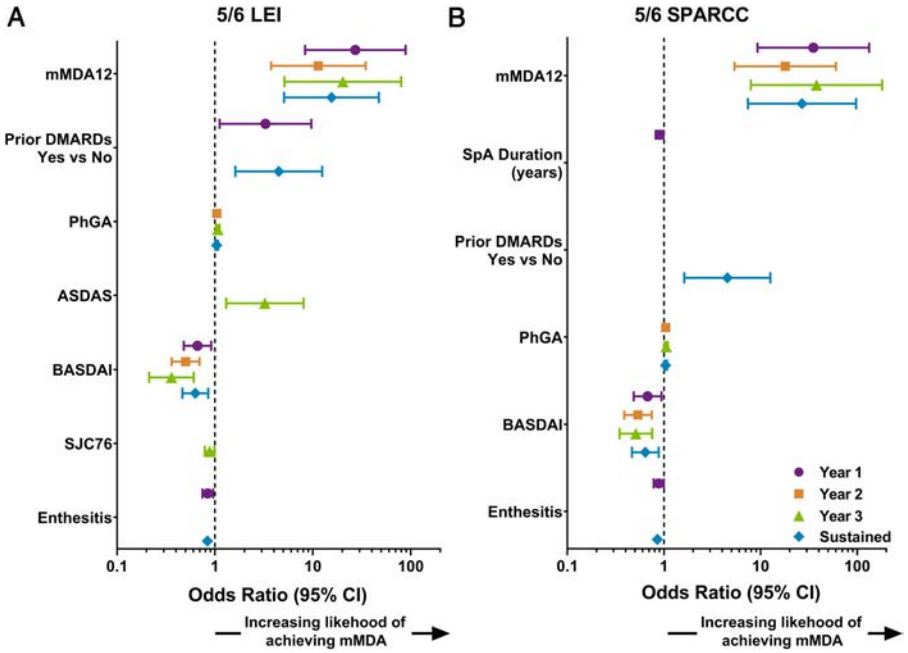

Figure 2. Factors associated with long-term and sustained modified minimal disease activity (mMDA) response predictors of long-term (years 1–3) and sustained mMDA responses at baseline and week 12 using a multiple logistic regression model, including mMDA response after 12 weeks of adalimumab treatment (mMDA12), for the 5 of 6 Leeds Enthesitis Index (LEI) components (**A**) and the 5 of 6 Spondyloarthritis Research Consortium of Canada (SPARCC) enthesitis index components (**B**). Only variables selected by a stepwise selection model are shown (variables selected by the model are significant at $P < 0.05$). 95% CI = 95% confidence interval; ASDAS = Ankylosing Spondylitis Disease Activity Score; BASDAI = Bath Ankylosing Spondylitis Disease Activity Index; DMARDs = disease-modifying antirheumatic drugs; PhGA = physician global assessment of disease activity; SJC76 = swollen joint count of 76 joints; SpA = spondyloarthritis.

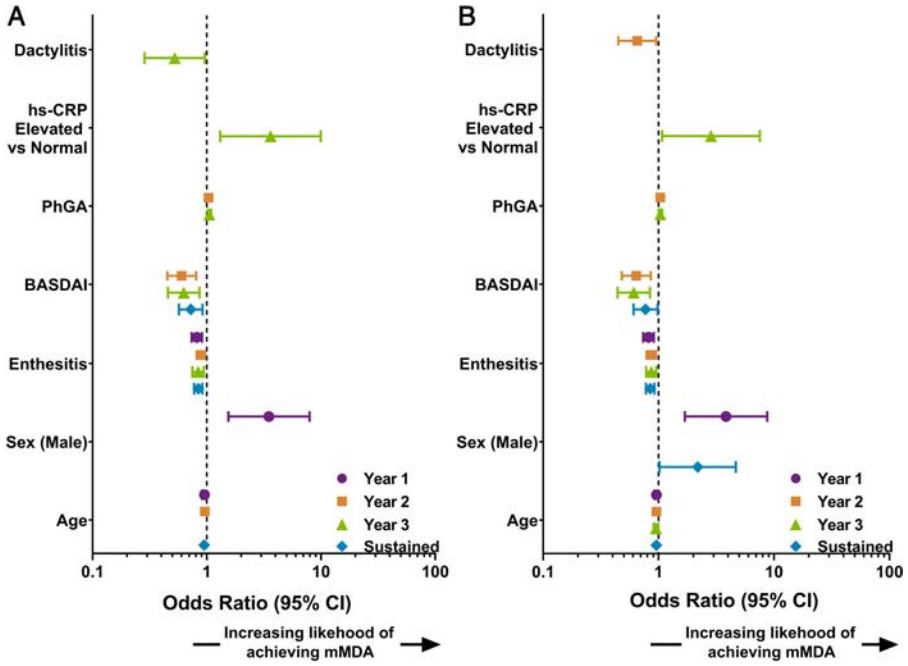

Figure 3. Factors associated with long-term and sustained modified minimal disease activity (mMDA) response predictors of long-term (years 1–3) and sustained mMDA responses at baseline using a multiple logistic regression model, without mMDA response after 12 weeks of adalimumab treatment, for the 5 of 6 Leeds Enthesitis Index (LEI) components (**A**) and the 5 of 6 Spondyloarthritis Research Consortium of Canada (SPARCC) enthesitis index components (**B**). Only variables selected by a stepwise selection model are shown (variables selected by the model are significant at $P < 0.05$). 95% CI = 95% confidence interval; BASDAI = Bath Ankylosing Spondylitis Disease Activity Index; hs-CRP = high sensitivity C-reactive protein; PhGA = physician global assessment of disease activity.

DISCUSSION

The MDA measure has been established as a valid composite outcome measure and appropriate treatment target in patients with PsA (19,20). The present analyses evaluated the potential applicability and performance of a modification of the MDA following principles of the OMERACT filter in patients with nonpsoriatic peripheral SpA from the ABILITY-2 study and identified predictors of long-term modified MDA responses following treatment with adalimumab.

Given that there is no established gold standard in peripheral SpA, criterion validity could not be assessed. Therefore, we evaluated concurrent validity, which assesses how the modified MDA compares with other measures, to measure similar constructs (the PSpARC and the ASDAS). All 4 versions of modified MDA criteria showed strong positive correlations with PSpARC remission and ASDAS ID, disease-specific outcome measures in peripheral SpA, underlining the validity of applying modified MDA criteria to peripheral SpA. Among the response criteria evaluated, all 4 modified MDA versions, the PSpARC remission, and ASDAS ID overall discriminated well between adalimumab and placebo treatment groups, and the modified MDA version that used either of the PsA-validated entheseal indices (the LEI and the SPARCC enthesitis index) performed comparably.

Similar to definitions of MDA used in PsA, the data presented here support the concurrent validity and discrimination of modified MDA using either of the entheseal indices depending on physician preference. At week 12, nearly 35–40% of patients treated with adalimumab achieved 4 of 6 versions of modified MDA, while 27–28% achieved the more stringent 5 of 6 versions of modified MDA. The proportions of patients achieving modified MDA was maintained over 3 years across the different definitions of modified MDA, reaching up to 58–60% in patients initially randomized to adalimumab and continued therapy. Throughout the duration of the study (3 years), sustained modified MDA was achieved by 55% (4 of 6 versions) and 44% (5 of 6 versions) of patients initially treated with adalimumab. The rates of sustained modified MDA using the more stringent 5 of 6 versions of the LEI or the SPARCC enthesitis index are similar to the rates observed in previous studies (21,22). Also, at week 12, 17–20% of patients receiving adalimumab achieved 6 of 6 versions of modified MDA, which is in accordance with the rates of MDA using 7 of 7 criteria reported in patients with PsA (23,24).

In PsA, MDA requires achievement of 5 of 7 criteria. However, the MDA criteria were modified for the nonpsoriatic peripheral SpA population by removing the psoriasis skin component (i.e., the PASI score); with this modification, it was unclear if 4 of 6 or 5 of 6 would perform best in peripheral SpA. Concurrent validity and treatment discrimination were similar for both the 4 of 6 and the 5 of 6 versions, so residual disease was examined to compare these further. Among patients treated with adalimumab who achieved 4 of 6 versions of modified MDA at week 12, joint responses (TJC and SJC)

were most often the limiting factors in modified MDA attainment. In patients achieving 5 of 6 versions of modified MDA, SJC, followed by HAQ DI (LEI) or enthesitis (SPARCC), appeared to limit modified MDA achievement, but residual levels of disease were much lower. Interestingly, all patients treated with adalimumab achieving modified MDA based on the LEI met the enthesitis criterion, whereas 9–10% failed to meet the enthesitis criterion based on the SPARCC enthesitis index. This difference is likely attributed to more sites being assessed in the SPARCC enthesitis index compared with the LEI. More studies are needed to assess whether the LEI or SPARCC enthesitis index might be more appropriate to use in peripheral SpA.

Although most modified MDA components were below the modified MDA cutoff in patients treated with adalimumab achieving 4 of 6 modified MDA criteria, the mean TJC and SJC were at or above the modified MDA cutoff. In contrast, the mean for all modified MDA components was below the modified MDA cutoff in patients treated with adalimumab achieving 5 of 6 versions. A measure of MDA or a treatment target should be associated with low levels of residual disease for face validity. The 5 of 6 modified MDA versions were more stringent than the 4 of 6 versions and more closely represent the concept of MDA with low values for most of the modified MDA components. Thus, to summarize: the 5 of 6 modified MDA criteria using either of the entheseal measures, rather than the 4 of 6 modified MDA criteria, could be an appropriate response measure in patients with peripheral SpA.

Multiple logistic regression analysis identified achievement of either of the 5 of 6 versions of modified MDA (modified MDA 5 of 6 [LEI] or modified MDA 5 of 6 [SPARCC]) after 12 weeks of adalimumab exposure as the strongest and most consistent predictors of long-term modified MDA, whether at 1, 2, or 3 years or sustained over time. That early clinical response is predictive of long-term response is in line with other studies in PsA, AS, or peripheral SpA (22,25,26). In PsA patients treated with certolizumab pegol, early clinical response at week 12 was identified as a positive predictor of MDA response at week 48 (27). Previously, achievement of early response at week 12 in patients with AS receiving adalimumab was found to be most predictive of long-term treatment response (28). Recently, in patients with peripheral SpA from ABILITY-2, ASDAS ID or PSpARC remission at week 12 were shown to predict subsequent long-term and sustained treatment response (14). All other variables including baseline BASDAI score were either only marginal predictors and/or did not reliably or consistently predict modified MDA response at every time point or sustained over time. In several studies evaluating predictors of MDA response in patients with PsA, baseline HAQ DI was most often reported as a negative predictor of long-term MDA (22,25,26).

Limitations of the current analysis include the limited sample size in each treatment arm and the lack of PASI score in the definition of MDA. However, ABILITY-2 had both a placebo-controlled period and long-term open-label extension, allowing the validation of modified MDA in a population of patients with peripheral SpA and identification of predictors of long-term treatment response. As there is no gold standard in peripheral SpA, we used concurrent validity to assess the performance of modified MDA versions compared to other available outcome measures that have been used in the past in peripheral SpA. Correlation analyses were only performed with a limited number of other end points such as the PSpARC or ASDAS ID, as few outcome measures are established in this disease. This analysis did not aim to address all aspects of the OMERACT filter systematically, so further work is required using different methodology to address truth (content and face validity), feasibility, and reliability. However, it should be noted that the original MDA has been proven to be feasible, so the modifications by definition should also be feasible given that 1 domain was removed. The performance of the modified MDA could not be analyzed in subgroups of peripheral SpA patients with dactylitis, inflammatory bowel disease (IBD), or uveitis due to the limited sample size of these subgroups. While the modified MDA criteria could reasonably be applied to peripheral SpA patients with dactylitis, as they evaluate tender and swollen joints, further studies are needed for patients with IBD or uveitis.

In conclusion, modified MDA using 5 of 6 criteria in peripheral SpA appears to be a valid, discriminative measure to assess treatment differences in patients with peripheral SpA. Achievement of early modified MDA response following adalimumab treatment was the most robust and consistent predictor of long-term modified MDA response.

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. Coates 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. Mease, Pangan.

Acquisition of data. Mease, Pangan, Song.

Analysis and interpretation of data. Coates, Abraham, Tillett, Mease, Ramiro, Wu, Wang, Pangan, Song.

ROLE OF THE STUDY SPONSOR

AbbVie funded this study, contributed to its design, and participated in the data collection, analysis, and interpretation of the data and in the writing and review of the manuscript. Medical writing support was provided by Deepa Venkitaramani, PhD, of AbbVie. Publication of this article was contingent upon approval by AbbVie.

REFERENCES

1. Rudwaleit M, van der Heijde D, Landewe R, Listing J, Akkoc N, Brandt J, 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.

2. Sepriano A, Landewe R, van der Heijde D, Sieper J, Akkoc N, Brandt J, et al. Predictive validity of the ASAS classification criteria for axial and peripheral spondyloarthritis after follow-up in the ASAS cohort: a final analysis. Ann Rheum Dis 2016;75:1034–42.
3. Paramarta JE, De Rycke L, Heijda TF, Ambarus CA, Vos K, Dinant HJ, et al. Efficacy and safety of adalimumab for the treatment of peripheral arthritis in spondyloarthritis patients without ankylosing spondylitis or psoriatic arthritis. Ann Rheum Dis 2013;72:1793–9.
4. Mease P, Sieper J, van den Bosch F, Rahman P, Karunaratne PM, Pangan AL. Randomized controlled trial of adalimumab in patients with nonpsoriatic peripheral spondyloarthritis. Arthritis Rheumatol 2015;67:914–23.
5. Carron P, Varkas G, Cypers H, van Praet L, Elewaut D, van den Bosch F, et al. Anti-TNF-induced remission in very early peripheral spondyloarthritis: the CRESPA study. Ann Rheum Dis 2017;76:1389–95.
6. Turina MC, Ramiro S, Baeten DL, Mease P, Paramarta JE, Song IH, et al. A psychometric analysis of outcome measures in peripheral spondyloarthritis. Ann Rheum Dis 2016;75:1302–7.
7. Coates LC, Fransen J, Helliwell PS. Defining minimal disease activity in psoriatic arthritis: a proposed objective target for treatment. Ann Rheum Dis 2010;69:48–53.
8. Coates LC, Helliwell PS. Validation of minimal disease activity criteria for psoriatic arthritis using interventional trial data. Arthritis Care Res (Hoboken) 2010;62:965–9.
9. Coates LC, Kavanaugh A, Mease PJ, Soriano ER, Laura Acosta-Felquer M, Armstrong AW, et al. Group for research and assessment of psoriasis and psoriatic arthritis 2015 treatment recommendations for psoriatic arthritis. Arthritis Rheumatol 2016;68:1060–71.
10. Gossec L, McGonagle D, Korotaeva T, Lubrano E, de Miguel E, Ostergaard M, et al. Minimal disease activity as a treatment target in psoriatic arthritis: a review of the literature. J Rheumatol 2018;45:6–13.
11. Gossec L, Smolen JS, Ramiro S, de Wit M, Cutolo M, Dougados M, 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.
12. Smolen JS, Schols M, Braun J, Dougados M, FitzGerald O, Gladman DD, 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.
13. Beaton DE, Maxwell LJ, Shea BJ, Wells GA, Boers M, Grosskleg S, et al. Instrument selection using the OMERACT filter 2.1: the OMERACT methodology. J Rheumatol 2019;46:1028–35.
14. Van den Bosch F, Mease PJ, Sieper J, Baeten DL, Xia Y, Chen S, et al. Long-term efficacy and predictors of remission following adalimumab treatment in peripheral spondyloarthritis: 3-year results from ABILITY-2. RMD Open 2018;4:e000566.
15. Mease PJ, van den Bosch F, Sieper J, Xia Y, Pangan AL, Song IH. Performance of 3 enthesitis indices in patients with peripheral spondyloarthritis during treatment with adalimumab. J Rheumatol 2017;44:599–608.
16. Machado P, Landewe R, Lie E, Kvien TK, Braun J, Baker D, et al. Ankylosing Spondylitis disease activity score (ASDAS): defining cut-off values for disease activity states and improvement scores. Ann Rheum Dis 2011;70:47–53.
17. Agresti A. Categorical data analysis. 3rd ed. Hoboken (NJ): John Wiley & Sons; 2013. p. 714.
18. Lincare JM. The expected value of a point-biserial (or similar) correlation. Rasch Measurement Trans 2008;22:1154.
19. Coates LC. Treating to target in psoriatic arthritis. Curr Opin Rheumatol 2015;27:107–10.
20. Coates LC, Moverley AR, McParland L, Brown S, Navarro-Coy N, O'Dwyer JL, et al. Effect of tight control of inflammation in early psoriatic arthritis (TICOPA): a UK multicentre, open-label, randomised controlled trial. Lancet 2015;386:2489–98.
21. Coates LC, Cook R, Lee KA, Chandran V, Gladman DD. Frequency, predictors, and prognosis of sustained minimal disease activity in an observational psoriatic arthritis cohort. Arthritis Care Res (Hoboken) 2010;62:970–6.
22. Kavanaugh A, van der Heijde D, Beutler A, Gladman D, Mease P, Krueger GG, et al. Radiographic progression of patients with psoriatic arthritis who achieve minimal disease activity in response to golimumab therapy: results through 5 years of a randomized, placebo-controlled study. Arthritis Care Res (Hoboken) 2016;68:267–74.
23. Lubrano E, Perrotta FM. Defining low disease activity states in psoriatic arthritis using novel composite disease instruments. J Rheumatol 2016;43:1765–6.
24. Mease PJ, Kavanaugh A, Coates LC, McInnes IB, Hojnik M, Zhang Y, et al. Prediction and benefits of minimal disease activity in patients with psoriatic arthritis and active skin disease in the ADEPT trial. RMD Open 2017;3:e000415.
25. Lubrano E, Parsons WJ, Perrotta FM. Assessment of response to treatment, remission, and minimal disease activity in axial psoriatic arthritis treated with tumor necrosis factor inhibitors. J Rheumatol 2016;43:918–23.
26. Theander E, Husmark T, Alenius GM, Larsson PT, Teleman A, Geijer M, et al. Early psoriatic arthritis: short symptom duration, male gender and preserved physical functioning at presentation predict favourable outcome at 5-year follow-up. Results from the Swedish early psoriatic arthritis register (SwePsA). Ann Rheum Dis 2014;73:407–13.
27. Van der Heijde D, Deodhar A, Fleischmann R, Mease PJ, Rudwaleit M, Nurminen T, et al. Early disease activity or clinical response as predictors of long-term outcomes with certolizumab pegol in axial spondyloarthritis or psoriatic arthritis. Arthritis Care Res (Hoboken) 2017;69:1030–9.
28. Sieper J, van der Heijde D, Dougados M, Brown LS, Lavie F, Pangan AL. Early response to adalimumab predicts long-term remission through 5 years of treatment in patients with ankylosing spondylitis. Ann Rheum Dis 2012;71:700–6.

Arthritis Care & Research
Vol. 74, No. 2, February 2022, pp 268–273
DOI 10.1002/acr.24465

American College of Rheumatology
Empowering Rheumatology Professionals

BRIEF REPORT

Determinants of the Physician Global Assessment of Disease Activity and Influence of Contextual Factors in Early Axial Spondyloarthritis

Fumio Hirano,[1] Robert B. M. Landewé,[2] Floris A. van Gaalen,[1] Désirée van der Heijde,[1] Cécile Gaujoux-Viala,[3] and Sofia Ramiro[4]

Objective. To investigate determinants of the physician global assessment (PhGA) of disease activity and the influence of the contextual factors on this relationship in patients with early axial spondyloarthritis (SpA).

Methods. Five-year data of DESIR, a cohort of early axial SpA, were analyzed. Univariable generalized estimating equations (GEEs) were used to investigate contributory explanatory effects of various potential determinants of PhGA. Effect modification by contextual factors (age, sex, and educational level) was tested, and if significant, models were stratified. Autoregressive GEE models (i.e., models adjusted for PhGA at the previous time point) were used to confirm a longitudinal relationship.

Results. A total of 708 patients were included. Higher Bath Ankylosing Spondylitis Disease Activity Index individual questions, swollen joint count in 28 joints (SJC28), tender joint count in 53 joints, Maastricht Ankylosing Spondylitis Enthesitis Score, C-reactive protein (CRP) level, and Bath Ankylosing Spondylitis Metrology Index score were associated with a higher PhGA. Sex and age were effect modifiers of SJC28; the contributory effect of SJC28 was largest in the younger male stratum ($\beta = 1.07$ [95% confidence interval (95% CI) 0.71, 1.43]), and the smallest in the older female stratum ($\beta = 0.13$ [95% CI 0.04, 0.22]). Autoregressive GEE models revealed the same determinants as having a longitudinal association with PhGA and the same pattern of effect modification.

Conclusion. Patients' subjective symptoms, peripheral arthritis and enthesitis, higher CRP level, and impaired spinal mobility contribute to explaining PhGA in patients with early axial SpA, irrespective of sex and age. Intriguingly, physicians consider the presence of swollen joints as more important in males than in females.

INTRODUCTION

Axial spondyloarthritis (SpA) is a rheumatic musculoskeletal disease primarily affecting the axial skeleton. Although axial SpA is mainly a disease of the spine, it may affect the peripheral joints and manifest itself in extramusculoskeletal manifestations. The goal of treatment in axial SpA is to improve the quality of life by abrogating inflammation that causes symptoms, structural damage, and physical disability. To achieve this goal, treatment needs to be adjusted through a shared decision based on a proper assessment of disease activity.

While the patient global assessment is considered an important item and is incorporated in the disease activity measurement in axial SpA, namely in the Ankylosing Spondylitis Disease Activity Score (1), far less attention is paid to the physician global assessment (PhGA). However, PhGA is still a major factor in the

The DESIR study is conducted as a Programme Hospitalier de Recherche Clinique with Assistance Publique Hopitaux de Paris as the sponsor and is supported by the French Society of Rheumatology. An unrestricted grant from Pfizer has been allocated for the first 10 years. The DESIR cohort is conducted under the control of Assistance Publique Hopitaux de Paris via the Clinical Research Unit Paris Center and under the umbrella of the French Society of Rheumatology and Institut National de la Santé et de la Recherche Medicale (Inserm). Database management is performed within the Department of Epidemiology and Biostatistics by Jean-Pierre Daures, DIM, Nîmes, France. Dr. Hirano's work was supported by a scholarship from the Japan College of Rheumatology (Japan College of Rheumatology–European Alliance of Associations for Rheumatology Young Rheumatologist Training Program).

[1]Fumio Hirano, MD, PhD, Floris A. van Gaalen, MD, PhD, Désirée van der Heijde, MD, PhD: Leiden University Medical Center, Leiden, The Netherlands; [2]Robert B. M. Landewé, MD, PhD: Amsterdam University Medical Center, Amsterdam, and Zuyderland Medical Center, Heerlen, The Netherlands; [3]Cécile Gaujoux-Viala, MD, PhD: University of Montpellier, Montpellier, and Nîmes University Hospital, Nîmes, France; [4]Sofia Ramiro, MD, PhD: Leiden University Medical Center, Leiden, and Zuyderland Medical Center, Heerlen, The Netherlands.

No potential conflicts of interest relevant to this article were reported.

Address correspondence to Sofia Ramiro, MD, PhD, Department of Rheumatology, Leiden University Medical Center, Albinusdreef 2, 2333 GA Leiden, The Netherlands. Email: sofiaramiro@gmail.com.

Submitted for publication March 31, 2020; accepted in revised form September 22, 2020.

SIGNIFICANCE & INNOVATIONS

- Physician global assessment (PhGA) of disease activity is a major factor in the therapeutic decision-making process, but little is known about which disease manifestations actually contribute to PhGA in axial spondyloarthritis.
- The patient's subjective symptoms, the peripheral joints and enthesitis findings, C-reactive protein level, and spinal mobility determine PhGA.
- The presence of swollen joints is most associated with higher PhGA in young male patients and the least in older females.

therapeutic decision-making process (2,3). Holding responsibility for the outcomes and knowledge of the therapeutic options, the treating physician supposedly summarizes the patient's complaints and objective findings into his/her own assessment and takes the initiative in therapeutic decisions. This situation is especially true with the advent of new therapies, which are effective, yet expensive and with distinct safety profiles.

Only a few studies have investigated PhGA as an outcome in patients with axial SpA (4–6). The main interest of these studies was the discordance with patient global assessment, and only 1 study has investigated the factors explaining PhGA (4). Cervical rotation, swollen joints, C-reactive protein (CRP) level, intermalleolar distance, and finger-to-floor were reported to explain PhGA. This analysis was based on cross-sectional data only and in patients with established ankylosing spondylitis, currently known as radiographic axial SpA. Nevertheless, little is known about which disease manifestations actually contribute to the PhGA over time and in particular in the early phase of the disease. Moreover, the patients' characteristics (such as age, sex, and educational level) are also hypothesized to have influence on the physician's interpretation of the disease manifestations. These patient characteristics are referred to as contextual factors and are important because they may have influence on outcomes as effect modifiers or confounders (7). The objectives of this study were to investigate the determinants of PhGA and the influence of contextual factors (age, sex, and educational level) on the contributory effects of the determinants of PhGA over time in patients with early axial SpA.

PATIENTS AND METHODS

Study design, study population, and outcome. DESIR is a cohort of patients with early inflammatory back pain highly suggestive of axial SpA. The protocol has been described previously (8). Briefly, the inclusion criteria were patients ages 18–50 years, with inflammatory back pain of >3 months and <3 years duration and symptoms highly suggestive of axial SpA according to the rheumatologist. A total of 708 patients were included consecutively in 25 French centers between December 2007 and April 2010. Clinical data were collected every 6 months up to 2 years and annually up to 5 years. Magnetic resonance imaging (MRI) of the spine and sacroiliac (SI) joints was performed in all patients at baseline and in patients from 9 centers in Paris at 2 years and 5 years. Patients with PhGA scores collected at least once during the 5-year follow-up were subjects of the present analyses. The PhGA was collected on a 0–10 numerical rating scale by asking the physician's "overall assessment of the activity of the rheumatic disease during the last 48 hours," with inactive disease and active disease as anchors (a higher score means higher disease activity). The database used for this analysis was locked in June 2016. DESIR has been approved by the appropriate ethics committees and patients signed the informed consent upon participation.

Potential explanatory variables of PhGA. Potential explanatory variables were as follows: individual component questions of the Bath Ankylosing Spondylitis Disease Activity Index (BASDAI; question 1 [Q1]: fatigue, Q2: back pain, Q3: peripheral joint pain, Q4: enthesitis, Q5: severity, and Q6: duration of morning stiffness; range 0–10 each), swollen joint counts in 28 joints (SJC28; range 0–28), tender joint counts in 53 joints (TJC53; range 0–159, with each joint graded as no tenderness = 0; tenderness = 1; tenderness and grimace = 2; tenderness, grimace, and withdrawal = 3), enthesitis measured with the Maastricht Ankylosing Spondylitis Enthesitis Score (MASES; range 0–39) (9), CRP level in mg/liter, the presence of any extramusculoskeletal manifestations (EMM, i.e., cumulative presence of any of uveitis, psoriasis, or inflammatory bowel disease), the Bath Ankylosing Spondylitis Metrology Index linear definition (BASMI linear; range 0–10), Spondyloarthritis Research Consortium of Canada MRI indices for the spine (SPARCC-spine; range 0–414) (10) and for the SI joints (SPARCC-SI joints; range 0–72) (11). SPARCC-spine and SPARCC-SI joints were mean scores of 3 central readers from wave 3 (blinded for chronologic order and clinical information) (12).

Contextual factors. The patient's age, sex, and educational level were the contextual factors tested as potential effect modifiers or confounders of the relationship between the determinants of PhGA and PhGA measurements. If stratification for age, sex, or educational level was needed, the population was dichotomized by age at baseline (younger or older than the median age at baseline, 33.3 years), sex (male or female), or educational level at baseline (university level or not).

Statistical analysis. To investigate the relationship between the independent variables and PhGA, we used generalized estimating equations (GEEs) (13). This method enabled us to make use of all available data and estimate a population-averaged parameter, correcting for within-patient correlation of outcomes at multiple time points. A linear GEE model was used

Table 1. Factors associated with the physician global assessment over time in sex- and age-stratified groups in univariable analysis*

	Female (younger) (n = 181)	Female (older) (n = 200)	Male (younger) (n = 173)	Male (older) (n = 154)
BASDAI Q1 (fatigue, 0–10)	0.39 (0.34, 0.44)	0.39 (0.34, 0.44)	0.46 (0.41, 0.51)	0.41 (0.35, 0.46)
BASDAI Q2 (back pain, 0–10)	0.53 (0.49, 0.57)	0.49 (0.45, 0.54)	0.58 (0.54, 0.63)	0.48 (0.43, 0.53)
BASDAI Q3 (peripheral joint pain, 0–10)	0.36 (0.31, 0.41)	0.31 (0.27, 0.36)	0.43 (0.37, 0.48)	0.32 (0.27, 0.37)
BASDAI Q4 (enthesitis, 0–10)	0.42 (0.37, 0.46)	0.37 (0.33, 0.41)	0.52 (0.47, 0.56)	0.36 (0.31, 0.41)
BASDAI Q5 (severity of morning stiffness, 0–10)	0.45 (0.40, 0.49)	0.42 (0.37, 0.46)	0.58 (0.54, 0.63)	0.44 (0.40, 0.49)
BASDAI Q6 (duration of morning stiffness, 0–10)	0.35 (0.30, 0.39)	0.30 (0.25, 0.35)	0.50 (0.45, 0.56)	0.36 (0.31, 0.41)
BASMI linear (0–10)	0.67 (0.48, 0.86)	0.61 (0.45, 0.78)	0.95 (0.75, 1.15)	0.49 (0.30, 0.68)
SJC28 (0–28)	0.52 (0.31, 0.73)	0.13 (0.04, 0.22)	1.07 (0.71, 1.43)	0.58 (0.40, 0.76)
TJC53 (0–159)†	0.13 (0.11, 0.16)	0.05 (0.04, 0.06)	0.15 (0.13, 0.18)	0.13 (0.11, 0.16)
MASES (0–39)	0.15 (0.12, 0.17)	0.10 (0.08, 0.12)	0.30 (0.25, 0.35)	0.18 (0.14, 0.23)
CRP, mg/liter	0.03 (0.01, 0.05)	0.02 (0.01, 0.04)	0.04 (0.03, 0.05)	0.06 (0.04, 0.07)
Any EMM (presence vs. absence)	–0.20 (–0.58, 0.19)	–0.13 (–0.49, 0.23)	–0.28 (–0.69, 0.14)	–0.26 (–0.68, 0.17)
SPARCC-spine (0–414)‡	0.05 (–0.11, 0.20) (n = 56)	0.06 (–0.11, 0.22) (n = 59)	0.05 (–0.04, 0.14) (n = 57)	0.02 (–0.03, 0.06) (n = 46)
SPARCC-SI joints (0–72)‡	0.01 (–0.08, 0.10) (n = 56)	–0.02 (–0.13, 0.09) (n = 60)	0.01 (–0.04, 0.06) (n = 57)	0.05 (–0.01, 0.11) (n = 46)

* Values are the coefficient (95% confidence interval). Univariable generalized estimating equation models with stratification for sex and age were used to investigate contributory explanatory effects of each factor on physician global assessment (PhGA). Age and sex were shown to be effect modifiers of the relationship between swollen joint count in 28 joints (SJC28) and PhGA and therefore analyses were conducted in Strata. BASDAI questions (Q) 1–6 = individual component questions of the Bath Ankylosing Spondylitis Disease Activity Index; BASMI linear = linear definition of Bath Ankylosing Spondylitis Mobility Index; CRP = C-reactive protein; EMM = extramusculoskeletal manifestation; MASES = Maastricht Ankylosing Spondylitis Enthesitis Score; SPARCC-spine/SPARCC-SI joints = Spondyloarthritis Research Consortium of Canada Magnetic Resonance Imaging indices for the spine/sacroiliac joints; TJC53 = tender joint count in 53 joints.
† Total score of the 53 joints with each joint graded 0–3 (0 = no tenderness, 1 = tenderness, 2 = tenderness + grimace, 3 = tenderness + grimace + withdrawal).
‡ Coefficients of SPARCC-spine/SPARCC-SI joints were estimated in a subgroup of patients with magnetic resonance imaging performed at least once at either 2 years or 5 years.

because the outcome was continuous. An exchangeable correlation matrix was selected because it showed the best fit.

First, we tested interactions between the contextual factors and the potential explanatory variables of PhGA. If the interaction term was significant with a predefined criterion of a *P* value less than 0.15 (14), analyses were further stratified. Effect modification was judged based on a clinically relevant difference in the regression coefficients across the strata. Univariable analysis was chosen to better assess the contributory explanatory effect of each of the determinants in each of the strata. This method was preferred to multivariable analysis, in which only the independent effect of a specific determinant, i.e., independent of confounders, would be considered. We were rather interested in the overall effect that a given determinant had on PhGA across different strata, also to allow proper comparisons across strata.

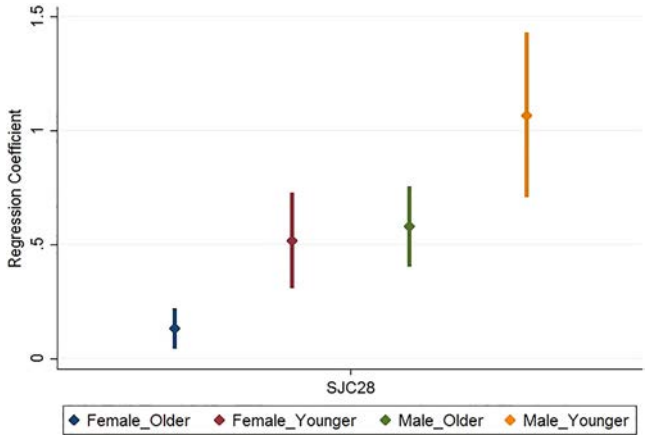

Figure 1. Impact of swollen joint count on the physician global assessment of disease activity across sex and age groups. Regression coefficients of the relationship of interest are plotted in an ascending order, reflecting an increasing impact of the swollen joint count on 28 joints (SJC28) on the physician global assessment of disease activity from older females to younger males (stratification according to median age at baseline, 33.3 years).

As relationships found in GEE models can be attributable both to cross-sectional and longitudinal effects, an autoregressive GEE model (i.e., a model adjusted for the outcome [PhGA] at the previous time point) was used to investigate whether the determinants had a true longitudinal association with PhGA. We used only data from yearly assessments (baseline, 12, 24, 36, 48, and 60 months), so that the intervals between the time points were similar. Contextual factors that proved not to be effect-modifiers were tested as confounders. If the addition of the contextual factor to a univariable GEE model made a relevant difference in the regression coefficient, it was deemed a confounder.

As MRI was repeated only in the patients from the centers in Paris, univariable GEE analyses with the MRI scores as potential explanatory variables could only be conducted in a subgroup of patients. A sensitivity analysis, similar to the main analysis, was conducted in patients fulfilling the Assessment

Table 2. Factors longitudinally associated with the change of physician global assessment from previous time points in sex- and age-stratified groups in univariable analysis*

	Female (younger) (n = 181)	Female (older) (n = 200)	Male (younger) (n = 173)	Male (older) (n = 154)
BASDAI Q1 (fatigue, 0–10)	0.39 (0.33, 0.45)	0.37 (0.31, 0.43)	0.37 (0.31, 0.44)	0.34 (0.27, 0.41)
BASDAI Q2 (back pain, 0–10)	0.53 (0.47, 0.58)	0.48 (0.42, 0.54)	0.55 (0.49, 0.60)	0.41 (0.35, 0.47)
BASDAI Q3 (peripheral joint pain, 0–10)	0.41 (0.35, 0.48)	0.33 (0.27, 0.38)	0.41 (0.34, 0.48)	0.27 (0.20, 0.33)
BASDAI Q4 (enthesitis, 0–10)	0.42 (0.37, 0.48)	0.37 (0.31, 0.42)	0.47 (0.41, 0.54)	0.34 (0.27, 0.40)
BASDAI Q5 (severity of morning stiffness, 0–10)	0.45 (0.39, 0.51)	0.38 (0.32, 0.44)	0.55 (0.49, 0.61)	0.38 (0.32, 0.44)
BASDAI Q6 (duration of morning stiffness, 0–10)	0.34 (0.28, 0.40)	0.29 (0.23, 0.35)	0.45 (0.37, 0.52)	0.29 (0.22, 0.35)
BASMI linear (0–10)	0.65 (0.42, 0.88)	0.54 (0.35, 0.72)	0.62 (0.40, 0.85)	0.34 (0.11, 0.56)
SJC28 (0–28)	0.73 (0.36, 1.10)	0.10 (–0.00, 0.21)	1.33 (0.73, 1.93)	0.61 (0.30, 0.92)
TJC53 (0–159)†	0.15 (0.11, 0.18)	0.05 (0.03, 0.06)	0.14 (0.11, 0.17)	0.13 (0.09, 0.17)
MASES (0–39)	0.13 (0.10, 0.16)	0.10 (0.07, 0.12)	0.25 (0.20, 0.31)	0.18 (0.13, 0.23)
CRP, mg/liter	0.02 (0.00, 0.04)	0.03 (0.01, 0.05)	0.02 (0.00, 0.03)	0.04 (0.02, 0.07)
Any EMM (presence vs. absence)	0.04 (–0.35, 0.44)	0.32 (–0.03, 0.68)	–0.26 (–0.63, 0.11)	–0.23 (–0.71, 0.24)
SPARCC-spine (0–414)‡	–0.11 (–0.28, 0.06) (n = 50)	0.03 (–0.13, 0.19) (n = 58)	0.08 (–0.02, 0.17) (n = 52)	–0.07 (–0.14, 0.00) (n = 46)
SPARCC-SI joints (0–72)‡	–0.01 (–0.13, 0.10) (n = 50)	0.12 (–0.01, 0.26) (n = 58)	0.02 (–0.05, 0.08) (n = 52)	0.01 (–0.08, 0.10) (n = 46)

* Values are the coefficient (95% confidence interval). Univariable autoregressive generalized estimating equation models (i.e., models adjusted for physician global assessment [PhGA] at the previous time point using data at 0, 12, 24, 36, 48, and 60 months) with stratification for sex and age were used to investigate longitudinal contributory effects of each factor on PhGA. Age and sex were shown to be effect modifiers of the relationship between swollen joint count in 28 joints (SJC28) and PhGA and therefore analyses were conducted in Strata. BASDAI questions (Q) 1–6 = individual component questions of the Bath Ankylosing Spondylitis Disease Activity Index; BASMI linear = linear definition of Bath Ankylosing Spondylitis Mobility Index; CRP = C-reactive protein; EMM = extramusculoskeletal manifestation; MASES = Maastricht Ankylosing Spondylitis Enthesitis Score; SPARCC-spine/SPARCC-SI joints = Spondyloarthritis Research Consortium of Canada Magnetic Resonance Imaging indices for the spine/sacroiliac joints; TJC53 = tender joint count in 53 joints.
† Total score of the 53 joints with each joint graded 0–3 (0 = no tenderness, 1 = tenderness, 2 = tenderness + grimace, 3 = tenderness + grimace + withdrawal).
‡ Coefficients of SPARCC-spine/SPARCC-SI joints were estimated in a subgroup of patients with magnetic resonance imaging performed at least once at either 2 years or 5 years.

of SpondyloArthritis international Society (ASAS) classification criteria for axial SpA. *P* values less than 0.05 were considered significant unless specified otherwise. All statistical analyses were conducted using Stata software, version 14.

RESULTS

A total of 708 patients were the subjects of the analyses. The subgroup of patients with repeated MRI consisted of 220 patients. Baseline characteristics of the total study population and the subgroup are shown in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24465/abstract. No important differences were found between the groups.

Higher scores of the BASDAI individual questions on fatigue (Q1), back pain (Q2), peripheral joint pain (Q3), enthesitis (Q4), severity and duration of morning stiffness (Q5 and Q6), SJC28, TJC53, MASES, CRP level, and BASMI linear were all associated with a higher PhGA (Table 1). Neither the presence of EMM nor inflammation captured on MRI had a contributory effect on PhGA. Sex and age were found to be effect modifiers of the relationship between SJC28 and PhGA; the contributory effect of SJC28 was largest in the younger male stratum, moderate in the younger female and the older male strata, and the smallest in the older female stratum (Table 1 and Figure 1). Sensitivity analyses in patients fulfilling the ASAS classification criteria retrieved similar results except for the lack of effect modification by age in female patients. The baseline characteristics of the age- and sex-stratified groups are shown in Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24465/abstract. Patients in the younger male stratum showed numerically more positivity for HLA–B27, higher modified New York grading, higher SPARCC-spine, and higher CRP level than the other strata. SPARCC-spine score was higher in male patients than in female patients.

Autoregressive GEE models yielded the same determinants as being longitudinally associated. A similar pattern of effect modification on SJC28 by sex and age (Table 2) was found. In other words, the determinants of PhGA (absolute value) are also associated with a change in PhGA over time.

DISCUSSION

In this study we have shown that patients' subjective symptoms, peripheral arthritis and enthesitis, higher CRP level, and impaired spinal mobility contribute to explaining PhGA in patients with early axial SpA, irrespective of sex and age. Interestingly, age and sex modify the impact of swollen joints on PhGA;

rheumatologists consider the presence of swollen joints as more important for the overall assessment of disease activity in male patients than in female patients, and more important in younger patients than in older patients.

Sex differences have been reported in outcomes of patients with axial SpA: male sex is associated with more progression of spinal structural damage in radiographic axial SpA (15). At the same time, a higher prevalence of fibromyalgia has been reported in female patients than in male patients with axial SpA (16), and a higher level of fatigue was observed in female patients in DESIR (17). In our particular cohort, male patients had higher CRP levels and more structural damage in the SI joints at baseline than the female patients. As a result, physicians may have related these characteristics to a higher risk of further progression. However, this explanation does not account for the fact that the different impact on PhGA across sex and age groups was only present for swollen joints. As we chose to use univariable analyses to properly compare the effect of each determinant across the different strata, rather than to investigate the independent effects, a higher SJC was possibly associated with other determinants in male patients compared to female patients and thus have a higher effect on PhGA; however, adjusted analyses, e.g., for CRP level, did not show this (results not shown). Or perhaps swollen joints are a finding that is likely to be attributed to other causes than disease activity in older and female patients. This different impact of the SJC across sex and age has not been previously reported. Physicians should become aware of this different impact, as it may represent a source of inequity for patients, especially if it influences therapeutic choices.

With regard to the determinants of PhGA, the individual BASDAI questions could be expected to contribute to PhGA, because these questions are, by definition, relevant items chosen for the assessment of disease activity. Likewise, the physicians reasonably took the findings of peripheral joints and entheses as a representation of the peripheral involvement and CRP level as a marker of inflammation. Spinal mobility was also a determinant of PhGA over time. Physicians probably relied on spinal mobility due to the lack of clinical objective findings reflecting spinal inflammation well. Also, this may be based on the knowledge that spinal mobility impairment is caused by inflammation as well as irreversible structural damage (18). However, by taking spinal mobility into account, physicians may be falsely reflecting structural damage on disease activity. Nevertheless, given the fact that DESIR is a cohort of early axial SpA with low levels of structural damage, this impact will be low in these patients (19).

The previous study that reported determinants of PhGA in radiographic axial SpA identified, through factor analysis, 4 latent factors of PhGA, which were labeled as "patient assessment," "mobility and function," "physician assessment," and "lab." Based on the individual items, these factors can be interpreted as "the patient's subjective symptoms," "spinal mobility and physical function," "peripheral arthritis," and "acute phase reactants" (4). Our results are similar except that we chose not to include physical function in the analyses as a potential explanatory variable, because it is usually considered a remote and indirect consequence of disease activity. Another previous study investigating PhGA as an outcome was from DESIR (5). This analysis focused on the discordance between patient global assessment and PhGA over time and the factors of the discordance and did not investigate the determinants of PhGA.

Our study has some limitations. First, whether the physician assessed PhGA while aware of the MRI findings for each patient is not clear. At least the physicians could not be aware of the MRI scores by the central readers at the time of assessment, as the scoring sessions only took place later. Therefore, a possible interpretation for the lack of impact of the MRI scores on PhGA is that the physicians did not have access to the MRI findings when making their judgment on the disease activity, because MRIs were usually made after the visit to the rheumatologist. Nevertheless, physicians possibly rate the findings of MRI as less important for their impression of the patient's disease than we may have thought. Likewise, in some centers, CRP levels were not available at the time of assessment. This absence may have led to a possible underestimation of the impact of CRP level. As a second limitation, we used a summary variable of the cumulative presence for EMM. This variable did not necessarily represent the presence of EMM at the time of assessment and using this summary variable may be the reason why the presence of EMM was not contributing to PhGA.

In conclusion, patients' subjective symptoms, peripheral arthritis and enthesitis, higher CRP level, and impaired spinal mobility contribute to explain PhGA in patients with early axial SpA irrespective of sex and age. However, physicians consider the presence of swollen joints as more important in males than in females.

ACKNOWLEDGMENTS

The authors thank the different regional participating centers in DESIR: Pr. Maxime Dougados (Paris–Cochin B), Pr. Andre Kahan (Paris–Cochin A), Pr. Philippe Dieudé (Paris–Bichat), Pr. Bruno Fautrel (Paris–La Pitié-Salpêtrière), Pr. Francis Berenbaum (Paris–Saint-Antoine), Pr. Pascal Claudepierre (Créteil), Pr. Maxime Breban (Boulogne-Billancourt), Dr. Bernadette Saint-Marcoux (Aulnay-sous-Bois), Pr. Philippe Goupille (Tours), Pr. Jean Francis Maillefert (Dijon), Dr. Emmanuelle Dernis (Le Mans), Pr. Daniel Wendling (Besançon), Pr. Bernard Combe (Montpellier), Pr. Liana Euller-Ziegler (Nice), Pr. Pascal Richette (Paris Lariboisière), Pr. Pierre Lafforgue (Marseille), Dr. Patrick Boumier (Amiens), Pr. Martin Soubrier (Clermont-Ferrand), Dr. Nadia Mehsen (Bordeaux), Pr. Damien Loeuille (Nancy), Pr. René-Marc Flipo (Lille), Pr. Alain Saraux (Brest), Pr. Xavier Mariette (Le Kremlin-Bicêtre), Pr. Alain Cantagrel (Toulouse), Pr. Olivier Vittecoq (Rouen). The authors also thank the research nurses, the staff members of the Clinical Research Unit of Paris Center, the staff members of the Biological Resource Center of Bichat Hospital, the staff members of the Department of Statistics of Nîmes and all the investigators and in particular Jerome Allain, Emmanuelle Dernis, Salah Ferkal, Clement Prati, Marie-Agnes Timsit, and Eric Toussirot for active patient recruitment and monitoring.

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. Ramiro 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. Hirano, van Gaalen, van der Heijde, Ramiro.

Analysis and interpretation of data. Hirano, Landewé, van Gaalen, van der Heijde, Gaujoux-Viala, Ramiro.

ROLE OF THE STUDY SPONSOR

Pfizer 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 Pfizer.

REFERENCES

1. Lukas C, Landewé R, Sieper J, Dougados M, Davis J, Braun J, et al. Development of an ASAS-endorsed disease activity score (ASDAS) in patients with ankylosing spondylitis. Ann Rheum Dis 2009;68: 18–24.
2. Van Der Heijde D, Ramiro S, Landewé R, Baraliakos X, Van Den Bosch F, Sepriano A, et al. 2016 update of the ASAS-EULAR management recommendations for axial spondyloarthritis. Ann Rheum Dis 2017;76:978–91.
3. Kievit W, van Hulst L, van Riel P, Fraenkel L. Factors that influence rheumatologists' decisions to escalate care in rheumatoid arthritis: results from a choice-based conjoint analysis. Arthritis Care Res (Hoboken) 2010;62:842–7.
4. Spoorenberg A, van Tubergen A, Landewé R, Dougados M, van der Linden S, Mielants H, et al. Measuring disease activity in ankylosing spondylitis: patient and physician have different perspectives. Rheumatology (Oxford) 2005;44:789–95.
5. Desthieux C, Molto A, Granger B, Saraux A, Fautrel B, Gossec L. Patient-physician discordance in global assessment in early spondyloarthritis and its change over time: the DESIR cohort. Ann Rheum Dis 2016;75:1661–6.
6. Wang CT, Fong W, Kwan YH, Phang JK, Lui NL, Leung YY, et al. A cross-sectional study on factors associated with patient–physician discordance in global assessment of patients with axial spondyloarthritis: an Asian perspective. Int J Rheum Dis 2018;21: 1436–42.
7. Boers M, Beaton DE, Shea BJ, Maxwell LJ, Bartlett SJ, Bingham CO, et al. OMERACT filter 2.1: elaboration of the conceptual framework for outcome measurement in health intervention studies. J Rheumatol 2019;46:1021–7.
8. Dougados M, d'Agostino MA, Benessiano J, Berenbaum F, Breban M, Claudepierre P, 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.
9. Heuft-Dorenbosch L, Spoorenberg A, Van Tubergen A, Landewé R, Van Der Tempel H, Mielants H, et al. Assessment of enthesitis in ankylosing spondylitis. Ann Rheum Dis 2003;62:127–32.
10. Maksymowych WP, Inman RD, Salonen D, Dhillon SS, Krishnananthan R, Stone M, et al. Spondyloarthritis Research Consortium of Canada magnetic resonance imaging index for assessment of spinal inflammation in ankylosing spondylitis. Arthritis Care Res (Hoboken) 2005;53:502–9.
11. Maksymowych WP, Inman RD, Salonen D, Dhillon SS, Williams M, Stone M, et al. Spondyloarthritis Research Consortium of Canada magnetic resonance imaging index for assessment of sacroiliac joint inflammation in ankylosing spondylitis. Arthritis Care Res (Hoboken) 2005;53:703–9.
12. Madari Q, Sepriano A, Ramiro S, Molto A, Claudepierre P, Wendling D, et al. 5-year follow-up of spinal and sacroiliac MRI abnormalities in early axial spondyloarthritis: data from the DESIR cohort. RMD Open 2020;6:e001093.
13. Twisk J. Applied longitudinal data analysis for epidemiology: a practical guide. Cambridge: Cambridge University Press; 2003.
14. Thiese MS, Ronna B, Ott U. P value interpretations and considerations. J Thorac Dis 2016;8:E928–31.
15. Ramiro S, Stolwijk C, Van Tubergen A, Van Der Heijde D, Dougados M, Van Den Bosch F, et al. Evolution of radiographic damage in ankylosing spondylitis: a 12 year prospective follow-up of the OASIS study. Ann Rheum Dis 2015;74:52–9.
16. Salaffi F, De Angelis R, Carotti M, Gutierrez M, Sarzi-Puttini P, Atzeni F. Fibromyalgia in patients with axial spondyloarthritis: epidemiological profile and effect on measures of disease activity. Rheumatol Int 2014; 34:1103–10.
17. Tournadre A, Pereira B, Lhoste A, Dubost JJ, Ristori JM, Claudepierre P, 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.
18. Machado P, Landewé R, Braun J, Hermann KG, Baker D, Van Der Heijde D. Both structural damage and inflammation of the spine contribute to impairment of spinal mobility in patients with ankylosing spondylitis. Ann Rheum Dis 2010;69:1465–70.
19. Ramiro S, van der Heijde D, Sepriano A, van Lunteren M, Moltó A, Feydy A, et al. Spinal radiographic progression in early axial spondyloarthritis: five-year results from the DESIR cohort. Arthritis Care Res (Hoboken) 2019;71:1678–84.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 2, February 2022, pp 274–280
DOI 10.1002/acr.24455

Association of Serum Low-Density Lipoprotein, High-Density Lipoprotein, and Total Cholesterol With Development of Knee Osteoarthritis

Jessica L. Schwager,[1] Michael C. Nevitt,[2] James Torner,[3] Cora E. Lewis,[4] Nirupa R. Matthan,[5] Na Wang,[1] Xianbang Sun,[1] Alice H. Lichtenstein,[5] and David Felson,[6] for the Multicenter Osteoarthritis Study Group

Objective. Studies suggest an association between elevated total serum cholesterol, particularly low-density lipoprotein (LDL), and osteoarthritis (OA). The present study was undertaken to evaluate the association between total cholesterol, LDL, and high-density lipoprotein (HDL) and risk of knee OA.

Methods. We studied participants from the Multicenter Osteoarthritis study (MOST) cohort at risk of developing knee OA. From baseline through 7 years, repeated knee radiographs and magnetic resonance images (MRIs) were obtained, and knee symptoms were queried. From baseline fasting blood samples, lipids and lipoproteins were analyzed using standard assays. After excluding participants with baseline OA, we defined 2 sets of patients: those developing radiographic OA, and those developing symptomatic OA (knee pain and radiographic OA). Controls did not develop these outcomes. Additionally, we examined worsening of cartilage loss and synovitis on MRI and of knee pain using the Western Ontario and McMaster Universities Osteoarthritis Index scale. We carried out logistic regression adjusting for age, sex, body mass index, education, baseline pain, and depressive symptoms, testing total cholesterol and lipoproteins as continuous measures, and we performed sensitivity analyses examining whether commonly used thresholds for high cholesterol, LDL, or low HDL increased risk.

Results. We studied 337 patients with incident symptomatic OA and 283 patients with incident radiographic OA. The mean age at baseline was 62 years (55% women). Neither total cholesterol, LDL, nor HDL showed a significant association with radiographic or symptomatic OA. Additionally, we found no association of these lipid measures with cartilage loss, worsening synovitis, or worsening knee pain.

Conclusion. Our data do not support an association between total cholesterol, LDL, or HDL with OA outcomes.

INTRODUCTION

Osteoarthritis (OA) is the most common form of arthritis and is a leading cause of disability in the US (1). Its prevalence is estimated at ~30 million in the US and 240 million worldwide, with increasing disease burden attributed to the obesity epidemic and an aging population (1,2). Current treatment focuses on management of pain and improvement of function. Thus far, therapies aimed at delaying structural deterioration of the joint have been unsuccessful in both modifying the course of the disease and reducing pain.

OA has been viewed traditionally as a disease caused by excessive mechanical loading, e.g., wear and tear, leading to degeneration of articular cartilage over time. Over the past decade, the understanding of processes underlying the pathology of OA have increasingly included, in addition to loading, a model of synovial inflammation driven by a complex interplay of cytokines, metalloproteinases, and reactive oxygen species causing cartilage degeneration and bone loss (2). Studies evaluating OA have identified a link between obesity and OA of hand joints (3–5), suggesting that OA could be caused, in part, by factors

Supported by the NIH (grants R01-AR-071950 and P30-AR-072571). The MOST study was supported by the NIH (grant U01-AG-18820 to Dr. Felson, grant U01-AG-18947 to Dr. Lewis, grant U01-AG-19069 to Dr. Nevitt, and grant U01-AG-18832 to Dr. Torner). Dr. Felson's work was supported by the NIHR as part of the Manchester Musculoskeletal NIHR Biomedical Research Centre grant.

[1]Jessica L. Schwager, DO, Na Wang, MA, Xianbang Sun, MS: Boston University, Boston, Massachusetts; [2]Michael C. Nevitt, PhD, MPH: University of California, San Francisco; [3]James Torner, PhD, MS: University of Iowa, Iowa City; [4]Cora E. Lewis, MD, MSPH: University of Alabama at Birmingham; [5]Nirupa R. Matthan, PhD, Alice H. Lichtenstein, DSc: Tufts University, Medford, Massachusetts; [6]David Felson, MD, MPH: Boston University, Boston, Massachusetts, and University of Manchester Centre for Epidemiology, and the National Institute for Health Research Manchester Biomedical Research Centre, Manchester University NHS Trust, Manchester, UK.

No potential conflicts of interest relevant to this article were reported.

Address correspondence to Jessica L. Schwager, DO, Rheumatology, E5, Boston University School of Medicine, Boston, MA 02118. Email: jessica.schwager17@gmail.com.

Submitted for publication March 2, 2020; accepted in revised form September 10, 2020.

SIGNIFICANCE & INNOVATIONS

- Recent data have suggested an association between elevated serum cholesterol levels, particularly low-density lipoprotein (LDL), and the development of osteoarthritis (OA); however, there is a paucity of literature directly evaluating the relationship between serum lipid and lipoprotein concentrations and incident OA.
- This is the first study to comprehensively and longitudinally determine whether circulating total cholesterol, LDL, and high-density lipoprotein are associated with the risk of developing radiographic and symptomatic knee OA.

unrelated to mechanical load. These findings support current research focusing on the hypothesis that OA is not only a localized joint disease, but an inflammatory disease involving both metabolic and mechanical factors.

Several experimental studies have suggested an association between lipid levels and the development of OA. Proposed mechanisms by which this may occur include antiinflammatory effects of elevated serum high-density lipoprotein (HDL) and proinflammatory effects of elevated serum low-density lipoprotein (LDL) and oxidized LDL. De Munter et al showed that mice that were fed a cholesterol-rich diet had LDL accumulation in synovial cells, synovitis, and increased ectopic bone formation; the authors proposed a mechanism by which increased levels of oxidized LDL activate synovial macrophages, fibroblasts, and endothelial cells, leading to local inflammation, cartilage loss, and ectopic bone formation (5–7).

Busso et al showed that plasma total cholesterol, HDL, and LDL were highly correlated with synovial fluid levels of these lipoproteins, suggesting that circulating serum lipids are able to freely move into the joint, further strengthening the hypothesis that altered lipid levels in the blood may have a direct effect on joint homeostasis (8).

Given the experimental evidence supporting an association between dyslipidemias and the development of OA, several authors have attempted to evaluate this relationship through observational studies. However, these studies have reported mixed results, highlighting the need for further research in this area (9,10). Few of these reports have focused on serum HDL and LDL concentrations, and to our knowledge, only 3 studies reported longitudinal data. In one longitudinal study, results suggested a decreased risk of hand OA with higher levels of HDL, but the sample size was small and confidence intervals wide. The other studies were not about painful OA per se but focused on magnetic resonance imaging (MRI) bone marrow lesions (BMLs) and their relation to lipids (11,12). While these data serve as a starting point in understanding the role that plasma lipids may play in the development of OA, there exists a paucity of literature directly evaluating the relationship between serum lipid and lipoprotein concentrations and incident OA, particularly incidence of disease in the knee, the site of most disabling OA.

Using longitudinal data from the Multicenter Osteoarthritis study (MOST) cohort, the goal of this study was to comprehensively and longitudinally determine whether circulating total cholesterol, LDL, and HDL were associated with the risk of radiographic and symptomatic knee OA in the MOST cohort.

PATIENTS AND METHODS

Study sample. MOST is a large NIH-funded longitudinal observational study focused on symptomatic and radiographic knee OA in a cohort of community-dwelling older adults with or at high risk for knee OA (13). The study enrolled 3,026 participants ages 50–79 years from 2003–2006 at 2 clinical sites (Iowa City, Iowa, and Birmingham, Alabama). Information on participants' demographic, medical, and lifestyle characteristics, as well as imaging results, were collected at baseline. Participants were followed with repeated examinations at 30, 60, and 84 months.

Weight-bearing, semiflexed posteroanterior and lateral views of the knees were obtained at baseline and at each examination according to the MOST radiograph protocol (14). Two readers interpreted and graded all radiographs according to Kellgren/Lawrence (K/L) grade, and if they disagreed, readings were adjudicated by a panel of 3 readers (15). MRIs of the knee were acquired at each visit using an OrthOne 1.0T magnet (ONI) and a circumferential extremity coil. All images were acquired without contrast. As in previous work (16), we read 1 randomly selected knee MRI per person. This was done for budgetary reasons and because of the high rate of symmetry in knee MRIs. The MRIs were read by 2 experienced musculoskeletal radiologists using the Whole-Organ MRI Score (WORMS) (17). Synovitis and cartilage morphology were scored in MRIs at baseline, 30, and 60 months. There was good interobserver agreement for each of the features reported (18). At each examination, participants completed the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) questionnaire, reporting on the amount of pain experienced during activities in each knee.

Anthropometric measurements (body mass index [BMI]). Weight was measured to the nearest 0.1 kg on a standard medical balance beam scale, and height was measured on full inspiration to the nearest 1 mm with a wall-mounted Harpenden stadiometer by certified MOST personnel following a written protocol. BMI was calculated as weight in kilograms divided by the height in meters squared.

Assessment of OA structure and pain symptoms. We defined 2 primary knee outcomes: incident radiographic OA and incident symptomatic OA, both up to 7 years after baseline, and created 2 separate nested case–control studies. In the first of

these, the outcome was incident radiographic knee OA among the subset of participants who had no radiographic OA in either knee (both knees with K/L grade <2) at baseline. Those participants who developed either radiographic knee OA (K/L grade ≥2) or had a knee arthroplasty in either knee by follow-up were defined as having incident radiographic knee OA. In the second case–control study, we focused on symptomatic OA. It was defined in a person when they had the combination of frequent knee pain (answering yes to the question, "Do you have pain, aching or stiffness in either knee on most days?") and had concurrent radiographic OA in that knee. Individuals with symptomatic OA at baseline in either knee were excluded from both case–control studies, and for each of these studies, we followed participants for 7 years to identify incident cases. For each of these case groups, we used risk set sampling to select controls randomly from eligible participants at baseline who did not become cases during follow-up. One risk set took 30-month follow-up and randomly selected controls who were not cases, then, a second set of cases at 60 months and a third set of cases at 84 months. For incident radiographic OA, we selected 1 control per case. To increase the likelihood that we would detect an association for the clinically important outcome of symptomatic OA, we selected 2 controls per case.

To further investigate potential associations of lipid and lipoprotein levels with outcomes, we assessed several secondary outcomes, including cartilage loss and change in synovitis based on MRI readings (17). These analyses were performed in the combined sample of all cases and controls. Within each of 14 subregions in each knee, cartilage morphology was scored 0–6 using the WORMS scale (17). We defined worsening cartilage morphology by analyzing each subregion and characterizing each as worsening when the score increased by 1 point. Subregions with baseline scores of 6 were excluded. Second, we examined change in synovitis. Each region (infrapatellar, intercondylar, and suprapatellar) was scored 0–3 based on volume at each time point, and the scores were then summed (0–9). We defined worsening synovitis as an increase in the summed score of ≥1, excluding knees with synovitis scores of 9 at baseline (19). We assessed 1 knee pain outcome (change in WOMAC pain) and calculated changes in pain as the difference of WOMAC pain score from baseline to the end of follow-up in each knee.

Lipid and lipoprotein profile. Blood draws were performed at the time of the baseline visit following an overnight fast. Blood samples were allowed to clot at room temperature for 30 minutes, and serum was separated by centrifugation at 1,500 gm at 4°C for 20 minutes. Aliquots were stored at –800°C in the MOST repository. For the determination of lipid profiles, matched case–control samples (n = 994) were shipped overnight on dry ice to the Cardiovascular Nutrition Laboratory at the Jean Mayer United States Department of Agriculture Human Nutrition Research Center of Aging at Tufts University. Serum total cholesterol and HDL concentrations were measured on an AU400e automated analyzer (Beckman Coulter) (intraassay coefficient of variation [CV] <3%; interassay CV <4%) using enzymatic reagents (Beckman Coulter). LDL concentration was calculated using the Friedewald equation (20) except when triglycerides were above 400 mg/dl. For those samples, a direct LDL method was used (AU400e automated analyzer, Beckman Coulter) (intraassay CV <2.4%; interassay CV <3.6%).

Potential confounders. As indicated for each analysis, the data were adjusted for participants' demographic, lifestyle, and medical history reported on the baseline questionnaire, age, sex (men, women), education (college or above; yes versus no), physical activity (Physical Activity Scale for the Elderly; continuous), smoking (never, past, current), BMI (kg/m^2; continuous), and statin use (yes/no). We used an indicator variable to adjust for race (White versus non-White) and clinic site. For all pain outcomes, we included depressive symptom score (Center for Epidemiologic Studies Depression scale score >16; yes versus no as a covariate). For knee pain analyses, we adjusted for baseline WOMAC pain score (continuous).

Statistical analyses. Our analytic sample consisted of MOST participants who were either selected as cases or controls

Table 1. Description of study participants according to case–control study*

Characteristic	Incident radiographic OA		Incident symptomatic OA	
	Cases (n = 285)	Controls (n = 329)	Cases (n = 338)	Controls (n = 560)
Age, years	60.8 ± 8.0	59.9 ± 7.5	62.4 ± 8.1	61.1 ± 7.8
Female, %	62	59	63	57
BMI, kg/m^2	30.5 ± 5.3	28.7 ± 4.5	30.9 ± 5.6	29.3 ± 4.8
Some college education, %	79	77	74	78
Cholesterol	229 ± 4.8	230 ± 51.2	225 ± 47.1	229 ± 48.9
LDL	136 ± 38.6	138 ± 42.9	133 ± 38.2	138 ± 4.03
HDL	62 ± 16.8	62 ± 16.7	62 ± 17.1	62 ± 16.5
Statin use, %	22	23	27	24

* Values are the mean ± SD unless indicated otherwise. BMI = body mass index; HDL = high-density lipoprotein; LDL = low-density lipoprotein; OA = osteoarthritis.

Table 2. Association of cholesterol, low-density lipoprotein (LDL), and high-density lipoprotein (HDL) with incident knee osteoarthritis (OA) according to results of logistic regression analyses*

	Mean (mg/dl)		OR_{adj} 95% CI (per SD)	*P*
	Cases	Controls		
Incident knee radiographic OA (cases/controls = 283/329)				
Total cholesterol	229	229	1.00 (0.88–1.14)	0.97
LDL	136	138	0.98 (0.87–1.12)	0.80
HDL	62	61	1.11 (0.96–1.29)	0.16
Incident symptomatic knee OA (cases/controls = 336/559)				
Total cholesterol	224	229	0.91 (0.80–1.03)	0.12
LDL	132	137	0.89 (0.79–1.01)	0.06
HDL	62	61	1.09 (0.95–1.24)	0.24

* Logistic regressions adjusted for the following covariates: age, sex, body mass index, educational attainment, race, clinic site, and Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain outcome score; adjusted also for baseline WOMAC pain score and depressive symptoms (yes/no). 95% CI = 95% confidence interval; OR_{adj} = adjusted odds ratio.

in 1 of our case–control studies (incident radiographic OA or incident symptomatic OA) and who had an archived baseline fasting serum sample. They had to be followed until at least the second MOST examination at 30 months.

Our analyses looked at the range of lipids on OA outcomes, testing each on a continuous basis and examining risk per SD increase. To avoid missing potential associations between total cholesterol, LDL, or HDL on OA outcomes, we tested commonly used thresholds for high total cholesterol (≥200 mg/dl), high LDL (≥130 mg/dl), and low HDL (≤60 mg/dl).

Analyses of radiographic OA and incident symptomatic OA were at the level of the person. For each case–control study, we used logistic regression to analyze the association of the lipid or lipoprotein level at baseline with the OA outcome. The dependent variable for each of these analyses was cumulative incidence of the OA outcome over 7 years. For analyses of MRI findings and of WOMAC pain, we combined data from all cases and controls from the primary case–control studies, creating 1 large sample. Analyses of MRI findings and of WOMAC pain were at the level of the knee or knee subregion for cartilage loss; to adjust for the correlation between knees (or subregions of knees), we performed generalized estimating equations.

We carried out several sensitivity analyses. First, because of concerns that baseline levels of lipids and lipoproteins might not accurately reflect levels up to 7 years later, we carried out analyses limiting incidence to 5 years. Second, some incident OA is caused by injury, which would tend to cause unilateral disease. We wanted to focus on those with systemic factors affecting disease; so, in another sensitivity analysis, we defined cases as those who during follow-up developed incidence in both knees, either contemporaneously (e.g., both at 30 months) or sequentially (e.g., 1 knee at 30 months, the other at 60 months). In addition, we examined quartiles of cholesterol, HDL, and LDL to see if high (or low) levels affected risk of disease. Last, we added visceral adiposity (21) as a covariate in our analysis to see if the relationships of lipids with OA outcomes were altered. Analyses were carried out using SAS, version 9.4.

Institutional review board approvals were obtained from the University of California, San Francisco, Boston University, the University of Alabama at Birmingham, and the University of Iowa. All participants provided written consent for study participation.

RESULTS

Mean age at baseline was 62 years, 55% of participants were women, and >20% received statins (see Table 1 for a description of the study participants), with 614 participants in the case–control study of radiographic OA and 898 in the study of symptomatic OA. There were no associations of incident radiographic OA and incident symptomatic OA with total cholesterol, LDL, and HDL cholesterol levels (Table 2).

Table 3. Associations of cholesterol, high-density lipoprotein (HDL), and low-density lipoprotein (LDL) levels with change in worsening synovitis or pain and with cartilage loss*

	Worsening synovitis (218 of 711)		WOMAC worsening (300 of 1,654)		Cartilage loss (4,268 of 10,297)	
	OR_{adj} (95% CI)	*P*	OR_{adj} (95% CI)	*P*	OR_{adj} (95% CI)	*P*
Total cholesterol	1.03 (0.70–1.52)	0.87	1.02 (0.88–1.20)	0.78	1.02 (0.96–1.10)	0.47
HDL	0.98 (0.81–1.18)	0.81	0.97 (0.83–1.14)	0.73	1.09 (1.02–1.17)	0.02
LDL	0.97 (0.80–1.18)	0.77	1.03 (0.88–1.21)	0.69	1.02 (0.95–1.09)	0.64

* Logistic regressions adjusted for the following covariates: age, sex, body mass index, educational attainment, race, clinic site, and for Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain outcome; adjusted also for baseline WOMAC pain score and depressive symptoms (yes/no). 95% CI = 95% confidence interval; OR_{adj} = adjusted odds ratio.

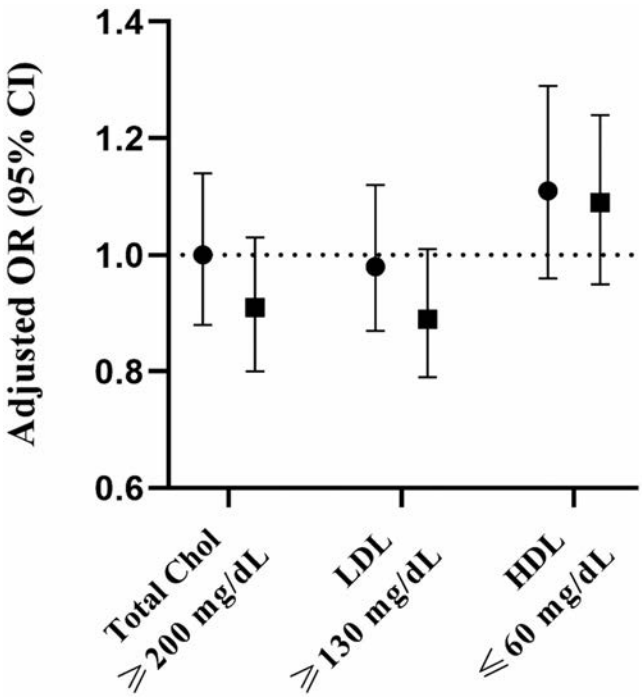

Figure 1. Association of high levels of cholesterol (chol) and low-density lipoprotein (LDL) and low levels of high-density lipoprotein (HDL) with incident osteoarthritis (OA) outcomes for radiographic knee OA (circles) and symptomatic knee OA (squares). 95% CI = 95% confidence interval; OR = odds ratio.

We also found no association of total cholesterol, LDL, and HDL with cartilage loss, worsening synovitis or worsening knee pain (Table 3). Last, we examined whether individuals with high total cholesterol and LDL or low HDL had an increased risk of OA and found no suggestive associations (Figure 1).

In sensitivity analyses, we found no associations of cholesterol, LDL, or HDL with incident OA if we limited incidence to the first 5 years after baseline, and we also found no association with lipids or lipoproteins when we restricted cases to those who developed bilateral disease. In analyses examining quartiles of cholesterol, HDL, and LDL, we similarly found no associations with incident radiographic or symptomatic OA. Last, adjusting for visceral adiposity in our main analyses (Table 2) did not affect results of these analyses (data not shown).

DISCUSSION

In this longitudinal nested case–control study, we examined the relationship between lipid and lipoprotein levels and OA within the MOST cohort. Although there is a growing body of experimental and epidemiologic evidence that suggest an association between elevated serum total cholesterol and LDL and low HDL with the development of OA, our data did not support such an association. We also did not find an association between serum lipid and lipoprotein levels with cartilage loss, worsening synovitis, or worsening knee pain.

Advances in the understanding of the pathogenesis of OA indicate that it involves not only wear and tear related to age and mechanical loading, but also synovial inflammation. This knowledge has led to a growing body of research evaluating the mechanisms by which metabolic factors that affect inflammation may contribute to OA. Several experimental studies evaluating the relationship of dyslipidemia and OA suggest that alterations in lipids play a role in the development of OA. De Munter et al showed that in mice with increased levels of LDL via cholesterol-rich diet or apolipoprotein E deficiency, there was increased synovial thickening and ectopic bone formation (6,7). Additionally, Triantaphyllidou et al (22) showed that mice with low HDL levels and high LDL levels (based on a lecithin-cholesterol acyltransferase knockout and apolipoprotein A-I knockout) that were placed on a Western diet developed OA in their joints, whereas control mice did not. In these mice, the Western diet activated enzymes that break down cartilage (5). Busso et al showed that total cholesterol, LDL, and HDL levels were all similar in plasma and synovial fluid, suggesting easy transit between compartments (23).

Another factor that has been implicated in leading to OA is oxidative stress and reactive oxygen species within the joint, particularly oxidized LDL. The complex mechanisms by which oxidative stress contribute to the pathogenesis of OA were recently reviewed by Lepetsos and Papavassiliou (24) and include overproduction of reactive oxygen species within chondrocytes leading to increased chondrocyte apoptosis and synovial inflammation, cartilage degradation, and subchondral bone dysfunction (24–27). Oxidized LDL, a lipid peroxidation product produced mainly by reactive oxygen species, has been shown to play a role in the pathogenesis of OA. Shen et al found increased levels of oxidized LDL in the synovial fluid as well as increased lectin-like oxidized LDL receptor 1 expression in the cartilage of OA patients in comparison to controls (28). Erturk et al investigated the relationship of paraoxonase 1 (PON-1), an enzyme that protects LDL and HDL from oxidative damage, with oxidized LDL and oxidative stress in OA (29). They found increased levels of oxidized LDL, higher markers of oxidative stress, and significantly lower levels of PON-1 in the sera of participants with OA in comparison to controls. There was also a correlation between serum oxidized LDL and knee OA grade utilizing the K/L scoring system and pain via WOMAC score.

The results of these experimental studies suggest that elevated levels of lipids and/or lipoproteins are major contributors to the pathogenesis of OA. Despite this strong experimental evidence, this study showed a null association with incident symptomatic and radiographic knee OA and elevated serum LDL levels.

Although there have been cross-sectional studies reporting possible associations between lipid levels and OA as recently evaluated in the systematic review and meta-analysis by Baudart et al (10), caution must be used when drawing conclusions regarding causation based on the results of cross-sectional studies.

To date, only 3 longitudinal studies have specifically assessed the relationship between certain features of OA (cartilage loss) and lipid and lipoprotein profiles. The results have been mixed. Garcia-Gil et al evaluated whether serum lipid levels were associated with incident radiographic hand OA in 277 participants without OA at baseline (9). Hand radiographs were repeated 11 years after baseline. No statistically significant associations between serum lipids and radiographic hand OA were observed, but a trend toward high HDL levels being associated with a lower risk of hand OA was reported. Additionally, the patient population included only women who were younger (mean age 50 years) with lower levels of obesity than the average US population.

Davies-Tuck et al evaluated 148 female participants without a history of OA for the development of BMLs of the knee; the authors obtained baseline lipids and followed up with knee MRI 2 years later (11). Results showed an association between BML incidence and higher total cholesterol levels, but no association with LDL or HDL. They did not find an association between presence of BMLs at baseline with any lipid parameter, nor with change in cartilage volume over 2 years. This study's results were at odds with a longitudinal study published 3 years later by Dore et al, who evaluated the development of BMLs of the knee in 394 male and female participants (12). Baseline lipids and knee MRI findings were evaluated, and BMLs were assessed by MRI 2.7 years later. They found no significant cross-sectional or longitudinal association between total cholesterol and BMLs, but they did report an inverse association of BML change and HDL.

In this study, we did not find any significant associations between incident symptomatic knee OA and lipid levels. It is unlikely that this null association was attributable to sample size, as our confidence bounds were narrow. For example, increasing levels of LDL were associated with a modest reduction in risk of OA, but the upper border of that confidence bound stretched only to 1.12, consistent with a 12% increase in the odds of disease.

If an association does exist, there may be other factors that have not yet been identified contributing to the correlation between lipids and OA. Among these other factors may be metabolic syndrome and a component of it, visceral adiposity. A recent review has strongly suggested that metabolic syndrome is not associated with knee OA, especially after adjustment for BMI (30). We carried out additional analyses testing visceral adiposity as a confounder, and it did not affect the associations that we report, suggesting it does not account for a spurious null association. Despite the strong evidence that lipids enter the joint freely and intraarticular oxidized LDL leads to joint damage, perhaps another mechanism exists within the joint that disturbs this relationship, making measured serum lipids unrelated to intraarticular oxidized LDL. Last, although our data showed a small protective association of elevated total cholesterol and elevated LDL on symptomatic OA, this was not statistically significant and was likely due to chance.

This study has several strengths. This is one of the few studies evaluating the temporal relationship between serum lipid levels and incident knee OA using a longitudinal design. We evaluated participants using indices that account for both symptoms (pain) and radiographic changes, whereas previous longitudinal studies have assessed only radiographic changes; therefore, we are able to identify participants who meet American College of Rheumatology criteria for knee OA, which includes the presence of pain (31). The nested case–control design of this study reduces selection bias because cases and controls are selected from the same population. Finally, we tested cholesterol as both a continuous measure and did sensitivity analyses examining whether commonly used lipid thresholds increased risk, thereby evaluating the data from several angles to ensure robustness.

One limitation of this study is that participants within the MOST cohort are primarily older, White Americans, and therefore our results may not be generalizable to a more diverse population. Additionally, we focused on incident knee OA, and it is possible that the metabolic factors affecting non–weight bearing joints such as the hand may differ from those affecting large, weight-bearing joints such as the knee. Also, our analyses of secondary outcomes such as cartilage loss may be affected by selection bias in that these outcomes were correlated with case status. While we adjusted in analyses for statin use at baseline, statin use may take decades to influence the occurrence of OA, and we did not know the duration of statin use in our participants. Further, we did not examine the effect of other lipid-lowering agents that were rarely used by MOST participants. Finally, lipid levels were taken at a single point in time, and therefore, we cannot rule out whether additional factors such as lifestyle modifications affected the lipid levels later, thus modifying OA risk.

In conclusion, we did not find an association between total cholesterol, LDL, or HDL levels with incident OA or other OA outcomes. While LDL may have local deleterious effects on joint structure, elevated systemic levels probably do not confer risk of disease.

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. Schwager 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. Felson.

Acquisition of data. Nevitt, Torner, Lewis, Matthan, Lichtenstein.

Analysis and interpretation of data. Schwager, Wang, Sun.

REFERENCES

1. Neogi T, Zhang Y. Epidemiology of osteoarthritis. Rheum Dis Clin North Am 2013;39:1–19.
2. Mathiessen A, Conaghan PG. Synovitis in osteoarthritis: current understanding with therapeutic implications. Arthritis Res Ther 2017; 19:18.

3. Yusuf E, Nelissen RG, Ioan-Facsinay A, Stojanovic-Susulic V, DeGroot J, van Osch G, et al. Association between weight or body mass index and hand osteoarthritis: a systematic review. Ann Rheum Dis 2010;69:761–5.
4. Gkretsi V, Simopoulou T, Tsezou A. Lipid metabolism and osteoarthritis: lessons from atherosclerosis. Prog Lipid Res 2011;50:133–40.
5. De Munter W, van der Kraan PM, van den Berg WB, van Lent PL. High systemic levels of low-density lipoprotein cholesterol: fuel to the flames in inflammatory osteoarthritis? Rheumatology (Oxford) 2016; 55:16–24.
6. De Munter W, Blom AB, Helsen MM, Walgreen B, van der Kraan PM, Joosten LA, et al. Cholesterol accumulation caused by low density lipoprotein receptor deficiency or a cholesterol-rich diet results in ectopic bone formation during experimental osteoarthritis. Arthritis Res Ther 2013;15:R178.
7. De Munter W, van den Bosch MH, Sloetjes AW, Croce KJ, Vogl T, Roth J, et al. High LDL levels lead to increased synovial inflammation and accelerated ectopic bone formation during experimental osteoarthritis. Osteoarthritis Cartilage 2016;24:844–55.
8. Busso N, Dudler J, Salvi R, Peclat V, Lenain V, Marcovina S, et al. Plasma apolipoprotein(a) co-deposits with fibrin in inflammatory arthritic joints. Am J Pathol 2001;159:1445–53.
9. Garcia-Gil M, Reyes C, Ramos R, Sanchez-Santos MT, Prieto-Alhambra D, Spector TD, et al. Serum lipid levels and risk of hand osteoarthritis: the Chingford Prospective Cohort Study. Sci Rep 2017;7:3147.
10. Baudart P, Louati K, Marcelli C, Berenbaum F, Sellam J. Association between osteoarthritis and dyslipidaemia: a systematic literature review and meta-analysis. RMD Open 2017;3:e000442.
11. Davies-Tuck ML, Hanna F, Davis SR, Bell RJ, Davison SL, Wluka AE, et al. Total cholesterol and triglycerides are associated with the development of new bone marrow lesions in asymptomatic middle-aged women - a prospective cohort study. Arthritis Res Ther 2009;11: R181.
12. Dore D, de Hoog J, Giles G, Ding C, Cicuttini F, Jones G. A longitudinal study of the association between dietary factors, serum lipids, and bone marrow lesions of the knee. Arthritis Res Ther 2012; 14:R13.
13. Segal NA, Nevitt MC, Gross KD, Hietpas J, Glass NA, Lewis CE, et al. The Multicenter Osteoarthritis Study (MOST): opportunities for rehabilitation research. PM R 2013;5:647–54.
14. Sheehy L, Culham E, McLean L, Niu J, Lynch J, Segal NA, et al. Validity and sensitivity to change of three scales for the radiographic assessment of knee osteoarthritis using images from the Multicenter Osteoarthritis Study (MOST). Osteoarthritis Cartilage 2015;23:1491–8.
15. Kellgren JH, Lawrence JS. Radiological assessment of osteoarthrosis. Ann Rheum Dis 1957;16:494–502.
16. Felson DT, Niu J, Neogi T, Goggins J, Nevitt MC, Roemer F, et al. Synovitis and the risk of knee osteoarthritis: the MOST Study. Osteoarthritis Cartilage 2016;24:458–64.
17. Peterfy CG, Guermazi A, Zaim S, Tirman PF, Miaux Y, White D, et al. Whole-Organ Magnetic Resonance Imaging Score (WORMS) of the knee in osteoarthritis. Osteoarthritis Cartilage 2004;12:177–90.
18. Lynch JA, Roemer FW, Nevitt MC, Felson DT, Niu J, Eaton CB, et al. Comparison of BLOKS and WORMS scoring systems part I. Cross sectional comparison of methods to assess cartilage morphology, meniscal damage and bone marrow lesions on knee MRI: data from the osteoarthritis initiative. Osteoarthritis Cartilage 2010;18:1393–401.
19. Guermazi A, Eckstein F, Hayashi D, Roemer FW, Wirth W, Yang T, et al. Baseline radiographic osteoarthritis and semi-quantitatively assessed meniscal damage and extrusion and cartilage damage on MRI is related to quantitatively defined cartilage thickness loss in knee osteoarthritis: the Multicenter Osteoarthritis Study. Osteoarthritis Cartilage 2015;23:2191–8.
20. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. Clin Chem 1972;18:499–502.
21. Li S, Schwartz AV, Lavalley MP, Wang N, Desai N, Sun X, et al. Visceral adiposity is associated with pain but not structural osteoarthritis. Arthritis Rheumatol 2020;72:1103–10.
22. Triantaphyllidou IE, Kalyvioti E, Karavia E, Lilis I, Kypreos KE, Papachristou DJ. Perturbations in the HDL metabolic pathway predispose to the development of osteoarthritis in mice following long-term exposure to Western-type diet. Osteoarthritis Cartilage 2013;21: 322–30.
23. Busso N, So A, Chobaz-Peclat V, Morard C, Martinez-Soria E, Talabot-Ayer D, et al. Leptin signaling deficiency impairs humoral and cellular immune responses and attenuates experimental arthritis. J Immunol 2002;168:875–82.
24. Lepetsos P, Papavassiliou AG. ROS/oxidative stress signaling in osteoarthritis. Biochim Biophys Acta 2016;1862:576–91.
25. Kouri JB, Aguilera JM, Reyes J, Lozoya KA, Gonzalez S. Apoptotic chondrocytes from osteoarthrotic human articular cartilage and abnormal calcification of subchondral bone. J Rheumatol 2000;27: 1005–19.
26. Cake MA, Read RA, Appleyard RC, Hwa SY, Ghosh P. The nitric oxide donor glyceryl trinitrate increases subchondral bone sclerosis and cartilage degeneration following ovine meniscectomy. Osteoarthritis Cartilage 2004;12:974–81.
27. Charlier E, Relic B, Deroyer C, Malaise O, Neuville S, Collee J, et al. Insights on molecular mechanisms of chondrocytes death in osteoarthritis. Int J Mol Sci 2016;17.
28. Shen P, Zhu Y, Zhu L, Weng F, Li X, Xu Y. Oxidized low density lipoprotein facilitates tumor necrosis factoralpha mediated chondrocyte death via autophagy pathway. Mol Med Rep 2017;16: 9449–56.
29. Erturk C, Altay MA, Bilge A, Celik H. Is there a relationship between serum ox-LDL, oxidative stress, and PON1 in knee osteoarthritis? Clin Rheumatol 2017;36:2775–80.
30. Li S, Felson DT. What is the evidence to support the association between metabolic syndrome and osteoarthritis? A systematic review. Arthritis Care Res (Hoboken) 2019;71:875–84.
31. Altman R, Asch E, Bloch D, Bole G, Borenstein D, Brandt K, et al. Development of criteria for the classification and reporting of osteoarthritis: classification of osteoarthritis of the knee. Arthritis Rheum 1986;29:1039–49.

Arthritis Care & Research
Vol. 74, No. 2, February 2022, pp 281–290
DOI 10.1002/acr.24460

American College of Rheumatology
Empowering Rheumatology Professionals

Qualitative Exploration of Dyadic Influence on Physical Activity Between Latina Patients With Osteoarthritis and a Supporter of Their Physical Activity

Sandra H. Soto,[1] Diane C. Berry,[2] and Leigh F. Callahan[1]

Objective. Research indicates that social support may promote physical activity; however, most Latina individuals with osteoarthritis (OA) are not sufficiently active. The purpose of this qualitative dyadic study was to explore how Latina patients with OA and a self-selected physical activity "supporter" motivate each other to be more active. Furthermore, perceptions of how OA symptoms impact support and physical activity were examined.

Methods. Semistructured dyadic interviews were conducted with Latina patients with OA and a member of their social network age ≥16 years who supports their physical activity (n = 14 dyads). We used framework analysis to reduce qualitative data to themes and subthemes.

Results. Daughters (n = 5), spouses (n = 4), sons (n = 2), a granddaughter (n = 1), a nephew (n = 1), and a friend (n = 1) provided support for the target behavior. In many cases, members of dyads said the motivation to engage in physical activity was reciprocated rather than focused solely on Latina patients with OA. Support was often reciprocated by engaging in physical activity together, using pressure, talking about being active, modeling physical activity, and helping with household responsibilities. Although participants agreed that physical activity was beneficial and Latina patients desired additional support when experiencing OA symptoms, there was concern about the safety of activity in the presence of symptoms. Several adult daughters indicated that their mothers' OA symptoms motivated their own physical activity.

Conclusion. Dyadic strategies for promoting physical activity among Latina patients with OA and how support may be reciprocated were identified.

INTRODUCTION

More than 1 in 5 Latino adults in the US has osteoarthritis (OA) (1). Latino individuals with arthritis have higher odds of severe joint pain (1.9 times higher), functional limitations (1.3 times higher), and work limitations (1.6 times higher) than non-Latino White patients (2,3). OA symptoms are made worse with physical inactivity (4). Unfortunately, among individuals with OA, Latino patients are less likely to report engaging in any exercise (50%) than non-Latino White patients (65%) (5).

Social influences may directly or indirectly affect physical activity (6). Direct influences occur when resources are shared to enable physical activity (7) and by offering physical activity companionship (7,8). Indirect influences occur when social modeling (9) or social control (e.g., pressure to exercise) (10) promote self-efficacy, when others' inactivity or activity creates physical activity norms (11), and when social support motivates self-regulation (12). Finally, Social Cognitive Theory posits that behaviors, individual factors (e.g., cognitions), and environmental factors (e.g., social influences) determine one another (13). For example, dyadic analyses among couples showed a bidirectional association between self-efficacy and social support provided by one's partner (14).

Traditional Latino culture emphasizes interdependence and reliance on others to regulate behaviors (15). Interdependence may be especially crucial to the promotion of physical activity among Latina individuals who often rely on social support and companionship to engage in physical activity (16–18). In a physical activity intervention with 266 Latina individuals a

Supported by the National Center for Advancing Translational Sciences, the North Carolina Translational and Clinical Sciences Institute (UL1-TR-002489), and the Thurston Arthritis Research Center, with funding from the National Institute of Arthritis and Musculoskeletal and Skin Diseases (P30-AR-072520-01).

[1]Sandra H. Soto, PhD, MPH, BSN, Leigh F. Callahan, PhD: Thurston Arthritis Research Center and University of North Carolina, Chapel Hill; [2]Diane C. Berry, PhD, ANP-BC, FAANP, FAAN: University of North Carolina, Chapel Hill.

No potential conflicts of interest relevant to this article were reported.

Address correspondence to Sandra H. Soto, PhD, MPH, BSN, School of Nursing, Carrington Hall, Campus Box 7460, University of North Carolina, Chapel Hill, NC 27599-7460. Email: shsoto@live.unc.edu.

Submitted for publication February 19, 2020; accepted in revised form September 15, 2020.

SIGNIFICANCE & INNOVATIONS

- Social support for physical activity was often reciprocated between Latina patients with osteoarthritis (OA) and their supporters.
- Engaging in physical activity together was a common and often preferred way of receiving support for physical activity among dyads.
- Some participants were unsure whether it was safe to engage in physical activity when experiencing severe OA symptoms.
- OA symptoms prompted adult children of Latina patients with OA to become more active.

greater increase in physical activity was found among women who were active with others versus alone (8). Qualitative research shows that expectations for social support from one's social network are strong, especially among Latina patients with arthritis (19,20). One of the few studies examining the social support of Latino individuals with arthritis and other rheumatic diseases found that 87% of participants reported receiving social support to cope with their symptoms (21). However, these few studies of social support for Latina patients largely focus on general social support to cope with arthritis pain (20–22) and the psychological well-being gained from receiving support (19,23). What is unknown is how others influence the physical activity of Latina patients with OA. Understanding the perceived social influences on Latina individuals' physical activity is imperative, given the interpersonal nature of Latino culture (15) and the general insufficiency of physical activity among Latina patients with OA (24). Thus, the purpose of this study was to explore the social influences of physical activity between Latina patients with OA and a supporter of their physical activity.

MATERIALS AND METHODS

We conducted semistructured, dyadic interviews in the preferred language of participants (English or Spanish), which were digitally recorded, transcribed verbatim, and coded in their original language. Quotes used in the presentation of the results were translated to English when necessary. All procedures were approved by the University of North Carolina (UNC) Institutional Review Board in accordance with their ethical standards and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Written informed consent was obtained by all participants prior to data collection.

Participants and sampling. Our research team used purposive sampling (25) to recruit participants who met predefined criteria: Latina patients with OA of the knee/hip and a supporter of their physical activity. Latina patients with OA were recruited in person through the Rheumatology Clinic at UNC, at a local senior center, and via mailed letters to individuals identified through the Carolina Data Warehouse for Health (a clinical repository of UNC Health Care data). Interested Latina patients were called by the principal investigator (SHS) or a trained interviewer, given an overview of the study, and if interested in participating, screened for eligibility. Eligibility criteria included a clinical diagnosis of hip and/or knee OA or probable OA (26) and the ability to nominate a person who supports their physical activity ("supporter"). Participants then provided the contact information of supporters, who were contacted via telephone and screened for eligibility (age ≥16 years and willing and able to attend the interview).

Procedure for data collection. Data were collected in the location of participants' choice (home [n = 11] or private community space [n = 3]). Conducting the interviews and collecting demographic information from participants lasted 60–90 minutes.

Table 1. Interview guide*

Topics	Prompts
Physical activity	How do you feel about physical activity in general? Are there physical activities that you enjoy?
Physical activity and social support	Are there certain types of support that help you be physically active? Are there ways you support others in their physical activity? How do you ask each other for help or motivation for physical activity? How have you helped each other be more physically active? How do you prefer to receive motivation to help you to be more physically active? What things do you wish the other person could do or say to motivate and help you be more physically active?
Osteoarthritis and physical activity	How does your arthritis or joint pain affect your physical activity? (directed at Latina patients with OA) How does the other person's arthritis or joint pain impact your physical activity in any way? (directed at supporters)
Osteoarthritis and social support	Does your arthritis or joint pain affect how you prefer to receive motivation or help to be physically active? (directed at Latina patients with OA) How does the other person's arthritis or joint pain affect the way you provide motivation or help for their physical activity? (directed at supporters)

* Unless otherwise stated, prompts were directed at both participants. OA = osteoarthritis.

Participants were given $20 each for their time. The principal investigator (SHS) and a trained bilingual interviewer conducted the interviews.

We developed a semistructured interview guide (Table 1) based on a literature review and prior experience of important social influences on physical activity. The guide was pilot tested with 1 dyad consisting of sisters-in-law (not included in the study sample), which demonstrated redundancy in some of the items. Most interviews were conducted in Spanish, 3 in English, and 1 in both languages. Using open-ended questions, the semistructured interview guide allowed for participants to stray from the questions to discuss related topics (27). Interviews were conducted until saturation of the data had been reached with respect to themes related to social influences on physical activity among dyads (28).

Data analysis. An initial codebook was created based on the interview guide and then adjusted after 2 bilingual coders separately coded 4 interviews that varied by language and dyad type. To analyze the coded data, framework analysis (29), which uses a systematic approach to identify themes within and between dyads, was implemented using ATLAS.ti (Mac version 1.6.0, Scientific Software Development). The transcripts were coded and hierarchically grouped to reduce the data to themes and subthemes for interpretation. Themes/subthemes were initially examined using all the data. Then the data were categorized by participant type (Latina patients with OA versus supporters), and dyad type (e.g., mother–child), and themes/subthemes were examined across the various groups.

We used the concepts of trustworthiness and authenticity outlined by Lincoln and Gupta (30) to ensure the rigor of our methods (see reference 31 for a recent description among clinical populations). Trustworthiness was established through 1) credibility, by interviewing dyads to obtain data from multiple sources and to develop a more comprehensive understanding of the themes (i.e., triangulation); furthermore, investigator triangulation was used to provide alternative explanations from investigators of Latino and non-Latino backgrounds; 2) transferability, by providing illustrative "thick" quotes connected to the themes/subthemes; 3) dependability, by carefully documenting the evolution of the codebook and crosschecking with previously coded interviews to safeguard the use of stable codes across interviews; and 4) confirmability, by using a second coder and discussing discrepancies to reach an agreement on codes and code definitions. Authenticity (30) was established through fairness (eliciting different perspectives) by interviewing a range of supporters and relationship types.

RESULTS

Participant characteristics. There were a variety of dyads represented in the sample: Latina patients with OA and their daughters (n = 5), spouses/partners (n = 4), sons (n = 2), and others, including a granddaughter, a nephew, and a friend. Except for the friend and 1 spouse, all supporters identified as Latino (Table 2).

Table 2. Participant characteristics (n = 14 dyads)*

	Patients with OA	Supporters
Age, mean ± SD years	66 ± 18	48 ± 23
Married or cohabitating	7 (50)	10 (71)
≥ High school education	7 (50)	13 (93)
Retired or homemakers	9 (65)	5 (36)
Monthly household income <$1,500	5 (36)	3 (23)
Born outside the US	13 (93)	5 (36)
Female	14 (100)	7 (50)
Supporter relation		
Daughter	–	5 (36)
Spouse/partner	–	4 (29)
Son	–	2 (14)
Granddaughter	–	1 (7)
Nephew	–	1 (7)
Friend	–	1 (7)

* Values are the number (%) unless indicated otherwise. OA = osteoarthritis.

Themes and subthemes. Interviews with dyads revealed 4 overarching themes and corresponding subthemes (Figure 1). The following sections describe the themes/subthemes with illustrative quotes. An outline of the themes/subthemes, along with more extensive quotes, is provided in Table 3.

How participants support one another's physical activity. *Reciprocated motivation.* Motivation and support were not always unidirectional, from supporters to Latina individuals. Eight dyads reported that support for physical activity was reciprocated, flowing back and forth between them. A mother (dyad 6) explained, "In fact, we motivate each other, her and I." These dyads described relying on each other for support and providing additional support when the motivation was lacking in 1 person to ensure their continued efforts to be physically active.

"Let's go for a walk." The most common way that dyads said they provided and preferred to receive support was by engaging in physical activity together (e.g., walking; 10 dyads). The use of this method was equally discussed by Latina patients with OA and by their supporters. For example, 1 daughter (dyad 2) said, "Usually my mom is the one who tells me, 'Let's go for a walk'...We try to go once a week." Across dyads, the commitment to exercise together varied, with some having a recurring appointment and others creating opportunities to exercise together on occasion.

Pressure. The use of pressure was rare among couples and typically given by Latina individuals rather than by their spouses. Five Latina patients and 8 supporters remarked that their counterpart did not use pressure and would be discouraged if they did. Three Latina patients with OA said their reaction depended on the type of pressure and from whom. Four Latina patients and 3 supporters said that receiving pressure was effective in

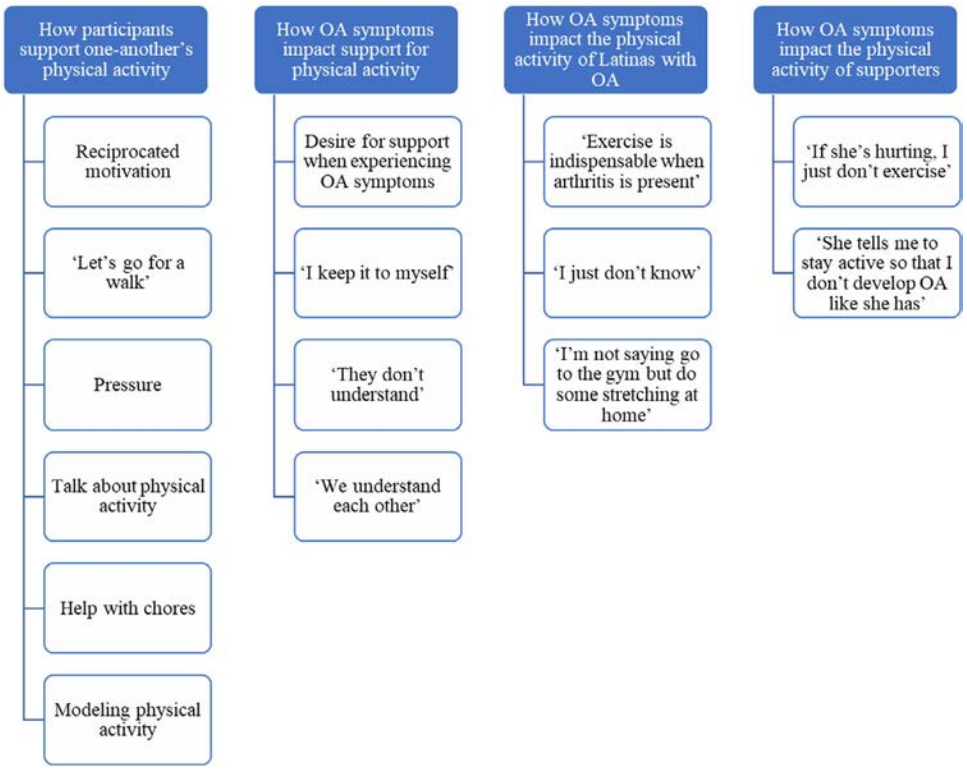

Figure 1. Themes and subthemes.

motivating their physical activity. One mother (dyad 5) said, "[Son] tells me, 'Go to the gym. Go do that. Did you go to the gym? Have you done the exercises prescribed by your doctor?' 'Yes,' I tell him, 'I've done them.' He motivates me a lot." In other words, although pressure was not universally embraced, it was welcomed and deemed useful among those who received this type of social influence.

Talk about physical activity. Three Latina patients with OA and 4 supporters said that they liked to talk about and receive verbal encouragement for physical activity. Topics of discussion often focused on using physical activity to prevent or manage long-term diseases, overcome barriers to physical activity, or to engage in general conversation around exercise. For example, one mother (dyad 10) explained, "Yes, sometimes when I see her after [she exercises], I ask her, 'How was your day?' or 'How's the exercise? What did you do? What did you like? What didn't you like?'"

Help with chores. Three Latina patients with OA and 2 daughters said that help with household responsibilities made it possible for them to be active. One daughter (dyad 6) said, "Sometimes I'll tell [my mom], 'Can you watch my dog a little while longer, I'm going to my gym class.'"

Modeling. Modeling physical activity was a motivator for only 2 Latina patients with OA: an aunt of a nephew supporter who works in a gym and a grandmother of a granddaughter supporter who is a dance teacher. For example, the grandmother (dyad 1) said, "Watching her [dance] motivates me."

How OA symptoms impact support for physical activity. *Desire for support when experiencing OA symptoms.* Six Latina patients expressed a desire to receive support for physical activity when they experienced OA-related pain and stiffness. A Latina individual interviewed with her spouse (dyad 11) said, "There are moments when you're in a lot of pain and you don't want to even move. In that moment, I do need someone to tell me, 'Do it so that the pain will ease.'" They believed physical activity could ease their symptoms, but their pain and stiffness impeded their motivation to be active, hence the desire for additional support.

"I keep it to myself." Three Latina patients with OA said they did not discuss their OA symptoms with their supporters because they did not want to complain and burden their loved ones, especially if they believed that there is nothing that can be done for OA symptoms. For instance, one mother (dyad 12) said, "No, no, I don't talk about pain, never." By not disclosing their symptoms, however, participants with OA did not let supporters know whether they should modify or increase their support to accommodate their loved one's physical and emotional state.

"They don't understand." The inability to recognize OA-related symptoms in their mothers resulted in misunderstandings in 2 mother-daughter dyads. This problem was illustrated by one mother (dyad 10): "She looks at me and says, 'Are you angry?' She often asks me, 'Why are you mad?' or 'Are you mad?' and sometimes I tell her, 'No, daughter, I'm just tired, in pain.'"

Table 3. Themes, subthemes, and sample quotes*

Theme and subtheme	Sample quote
How participants support one another's physical activity	
Reciprocated and bidirectional motivation	Daughter: For example, if I know that my mom is going dancing, then I will probably want to go with her too. Mother: Me too. In fact, we motivate each other, her and I. She motivates me with some things like, "Mom, you should try this, Mom, you should try that." And I with her, like, "Look, I saw this thing, why don't we go do it?", "I'm going to this thing, want to come?" (Dyad 6)
"Let's go for a walk"	What motivated me personally is that he helps me. He tells me let's go do this. I feel supported. We're going to do this together, that way we work as a team, him and I. That motivates me. (Spouse/partner of supporter: dyad 9) Usually my mom is the one who tells me, "Let's go for a walk" We try to go once a week. Sometimes we do it, sometimes because of our schedules or because the kids need help with their homework, but our goal is to go out together for a walk. (Daughter: dyad 2)
Pressure	[Son] tells me, "Go to the gym. Go do that. Did you go to the gym? Have you done the exercises prescribed by your doctor?" "Yes' I tell him, I've done them." He motivates me a lot. (Mother of son supporter: dyad 5) I started going to the gym class about 2 months ago because she was going to it before. I used to hear it every day, "You have to go. You have to do this." I have to give her credit for that. Persistence and perseverance paid off. (Spouse/partner: dyad 4)
Talk about physical activity	Yes, sometimes when I see her after [she exercises], I ask her, "How was your day?" or "How's the exercise? What did you do? What did you like? What didn't you like?" And we'll discuss it. Other times, she'll tell me, "I want to go to a new class" and I'll say, "You should do it. If you don't have a ride, I can take you." (Mother of daughter supporter: dyad 10)
Help with chores	The support is: I have to go walk and he has to take care of [his sister] or vice versa. Or if it's a little more cold and our grandson is here, "You stay with him but I'm going to go [for a walk]. Then you can go." (Spouse/partner of supporter: dyad 9) Sometimes I'll tell her, "Can you watch my dog a little while longer, I'm going to my gym class." It's the only way that I sometimes ask for help because the class has a specific schedule. Leaving work at 6:30 pm leaves me time to go to the last class at 7:15 pm. From 7:15 pm to 8:15 pm is the time of the class and I say, "Please, I just want to quickly go to the gym. One hour. I'll be there at 8:30 pm." She tells me, "OK, it's fine" or "I have to go out, so hurry," that kind of thing. (Daughter: dyad 6)
Modeling physical activity	Watching her [dance] motivates me. (Grandmother of a granddaughter supporter: dyad 1)
How OA symptoms impact support for physical activity	
Desire for support when experiencing OA symptoms	There are moments when you're in a lot of pain and you don't want to even move. In that moment, I do need someone to tell me, "Do it so that the pain will ease." (Spouse/partner of supporter: dyad 11)
"I keep it to myself"	Daughter: I see in her face that she's not feeling well because her face changes. What I do is I ask her, "Mom, what's wrong?" and she tells me, "Nothing." But I notice her face and I keep asking her and she gets angry. I prefer to not ask her anything and not say anything. Mother: Yes, because what I have in my head about these problems, based on what I've read, is that there is no cure. I don't know if it's true or if it's a lie. What I understand is that I have to live with this problem. But if I have to live with it, I'm not going to internalize it. I try to think of it as something that will pass. (Dyad 13) No, no, I don't talk about my pain, never. (Mother of daughter supporter: dyad 12)

(Continued)

Table 3. (*Cont'd*)

Theme and subtheme	Sample quote
"They don't understand"	She looks at me and says, "Are you angry?" She often asks me, "Why are you mad?" or "Are you mad?" and sometimes I tell her, "No, daughter, I'm just tired, in pain," but it's like those words don't work, like she's thinking that I'm mad at her or at something that happened in the moment but that's not it. The pain sometimes puts me in a bad mood, and I'll go, I'll agree and go [to the gym] but sometimes, I tell her "No, you go" because I don't feel up to it. She also knows that there's a point at which she'll say, "OK mom, go rest." (Mother of daughter supporter: dyad 10)
	We have gotten in the habit of walking [by the river] on Saturday mornings. All of us except my mom because she doesn't like to go because of the pain. What we do is, I try to insist, but she tells me that I don't know the pain that she feels, so she prefers to not come with us when we go for a walk. I go with my daughters and my husband. (Daughter: dyad 6)
"We understand each other"	[Arthritis] is a part of me. When something happens to her related to the arthritis, I also feel it. Why? Because I have also felt the same pain and so we motivate each other to do exercise. (Spouse/partner: dyad 11)
How OA symptoms impact the physical activity of Latina patients with OA	
"Exercise is indispensable when arthritis is present"	Yes. I was telling her, maybe if you get to go to exercise and do lots of activities of all of your muscles, your leg muscles maybe need to reinforce it. Myself I'm less for surgery, I'm more proactive with the exercise. (Friend: dyad 3)
	When I'm in pain I tell him, "My knee hurts" and he tells me, "That's because you haven't been to the gym. You haven't been walking. You have to walk." He tells me, "Don't lay down to sleep or rest. No, it's because you haven't done anything. Let's go." (Aunt of nephew supporter: dyad 8)
"I just don't know"	I think, maybe I'm not informed, but when I see her suffering, I'm like, "Don't do anything," more as a precaution, to prevent her from wearing herself out more, that's how I think of it. Surely, I'm wrong, I don't know. (Granddaughter: dyad 1)
"I'm not saying go to the gym but do some stretching at home"	I think, [exercise when in pain] depends on the pain, because a lot of times if you're stiff and in pain, throughout the day you warm up and feel better. Maybe an activity that's very light, but not to stop moving. For example, any extreme activity, for example, doing squats or something with pain, you're going to do it wrong and you can hurt yourself. I think there has to be a middle ground, everything depends on the activity. My mom has never tried swimming, but I feel like that would be very gentle with her joints and it's good exercise. (Daughter: dyad 6)
How OA symptoms impact the physical activity of supporters	
"If she's hurting, I just don't exercise"	The times that she's like not feeling it and she doesn't like it, she doesn't want to do anything, it's a barrier because like I said earlier, I don't like doing exercise by myself because I guess if I'm not seeing someone else to do with it, it's like why should I do it? (Spouse/partner: dyad 7)
	If the pain is strong, it's best to keep still. (Mother of son supporter: dyad 14)
"She tells me to stay active so that I don't develop OA like she has"	Seeing her like that in that moment is motivation for me to not end up like that. It's exactly the reason why it's important to keep moving. Laws of physics, an object in motion stays in motion. (Daughter: dyad 5)

* OA = osteoarthritis.

This lack of understanding led to mothers' expressed frustration that their daughters did not grasp the extent of their pain experience. In these cases, mothers tended to decline invitations to be active when experiencing OA symptoms.

"We understand each other." Conversely, negative situations were avoided when both participants had OA and could empathize with each other's symptoms (2 dyads). For example, one spouse (dyad 6) explained, "When something happens to

her related to the arthritis, I also feel it. Why? Because I have also felt the same pain and so we motivate each other to do exercise."

How OA symptoms impact the physical activity of Latina patients with OA. *Exercise is indispensable when arthritis is present.* Overall, dyads believed it was a good idea to be physically active when Latina patients were experiencing OA symptoms (8 Latina patients with OA and 9 supporters). Reasons for this idea centered around the belief that movement could alleviate pain and stiffness and is a preferable alternative to surgery as a treatment of arthritis. This subtheme was explained by a friend (dyad 3), "Yes. I was telling her, maybe if you get to go to exercise and do lots of activities of all of your muscles, your leg muscles maybe need to reinforce it. Myself I'm less for surgery, I'm more proactive with the exercise."

"I just don't know." Two Latina patients with OA and 5 supporters either did not believe it was a good idea to be active while experiencing OA symptoms or were unsure. A granddaughter from dyad 1 said, "I think, maybe I'm not informed, but when I see her suffering, I'm like, 'Don't do anything,' more as a precaution, to prevent her from wearing herself out more, that's how I think of it. Surely, I'm wrong, I don't know." Although they believed physical activity was beneficial to OA, these participants expressed concern that it could cause harm to the joints when symptoms were present.

"I'm not saying go to the gym but do some stretching at home." Two Latina patients with OA and 1 supporter reconciled their ambivalence about the safety of physical activity by deciding that it was a good idea unless those symptoms were severe, in which case, taking it slow by doing gentle stretching/yoga, walking, swimming, or resting was the best approach. One daughter (dyad 6) said, "I think [exercise when in pain] depends on the pain, because a lot of times if you're stiff and in pain, throughout the day you warm up and feel better. Maybe an activity that's very light, but not to stop moving."

How OA symptoms impact the physical activity of supporters. *"If she's hurting, I just don't exercise."* The physical activity of supporters was sometimes altered by Latina patients' OA symptoms. Three supporters, typically spouses, explained that they were less likely to engage in physical activity when their partner experienced OA symptoms because they no longer received the same level of physical activity support from their partner. For example, one spouse (dyad 7) said, "It's a barrier because like I said earlier, I don't like doing exercise by myself." Supporters also explained that they were less likely to engage in physical activity when Latina patients were experiencing OA symptoms because they would opt to stay home to provide care rather than leaving them to be active.

"She tells me to stay active so that I don't develop OA like she has." Four supporters explained that they became more active when Latina patients experienced OA symptoms. A few said that this increase in activity was because they took on more household chores and thus engaged in more activity. Others, typically daughters, said that seeing their mothers in pain motivated them to become more active to prevent the onset of OA. One daughter (dyad 5) explained that "Seeing her like that in that moment is motivation for me to not end up like that. It's exactly the reason why it's important to keep moving." Similarly, mothers used their pain experience as an example to encourage their daughters to become more physically active.

DISCUSSION

This dyadic interview study explored support for physical activity between Latina patients with OA and a family member/friend. While other researchers have observed Latino individuals seeking social support to cope with arthritis pain (20–22), our study explored how Latina patients are seeking support for physical activity, a pain-coping behavior.

Supporters were primarily adult children, mostly daughters, and to a lesser degree, spouses and others, consistent with prior research showing that Latina individuals typically rely on daughters as a source of social support (7,22). However, interventions on the management of chronic illnesses have generally focused on couples (32), underscoring a gap in intervention research that is responsive to the needs of Latino communities. Thus, an essential next step for research is designing dyadic interventions that include nonspouse supporters, especially because the efficacy of these interventions does not appear to depend on the composition of the dyad (33).

The dyadic design of our study enabled us to identify examples of reciprocity between participants and to respond to limitations of prior social support literature that describe support recipients as passive receivers (34). We found that social support for physical activity is received *and* given by Latina patients with OA, especially when both participants are inactive. When physical activity is not yet a habit, providing support may promote physical activity above and beyond receiving support (34), thus reciprocated support may explain why including dyads in physical activity interventions may be more effective than enrolling target participants alone (33). Furthermore, equal intervention participation by both participants likely has a greater effect on physical activity than does unequal participation (35,36).

Participants supported one another's physical activity in a variety of ways, primarily by engaging in physical activity together. This echoes prior research showing that engaging in physical activity with others (i.e., activity partners) is highly influential among women (7,16) because of the accountability and enjoyment created by their partners (37). Our finding supports social support theory (6), which posits the direct influence of activity partners on physical activity, and adds to prior literature demonstrating these effects among Latina individuals (7,8). Walking was a common method of engaging in physical activity together in the current study. In addition to being accessible and easily

modifiable, walking is deemed safe and recommended as an effective self-management method of reducing OA symptoms (38). Evidence-based walking interventions for OA that target individuals (39) may be more effective among Latina individuals if they are adapted to dyads.

As in previous studies, some of our participants said receiving pressure-like tactics would backfire and cause them to resist being physically active (40). However, other women, with and without OA, welcomed the use of pressure and deemed it an effective motivator. In a study among couples, receiving pressure-like tactics resulted in greater self-efficacy (10), possibly due to a defiant response in one's confidence to be more active. In our study, however, participants were pleased with the pressure they received. Although the examples of pressure provided by participants sound negative: "Have you done the exercises prescribed by your doctor?", "You're not going to fit in your bikini sitting on the couch," the tone of the messages may be more persuasive (e.g., a concerned tone in the former quote and a jovial tone in the latter) (41). The tone, context, and the perceptions of the receiver may be more important than the content of the message (40). To date, most studies have investigated the use of pressure among non-Latino White couples (10,40,42,43), but the couples in our study rarely used this tactic, suggesting possible differences in methods used by spouses across ethnicities. Further research can address this gap and clarify associations among pressure, mediating theoretical constructs, and physical activity in noncouple dyads.

Our findings showed that most Latina patients wanted additional support for physical activity when they experienced more severe OA symptoms and needed help overcoming physical activity barriers. This finding was also seen in a study of 588 Latino patients with arthritis (44). Those with severe pain had stronger motivation to exercise compared to those with less severe pain. However, individuals with more pain also experienced more barriers to physical activity. Unfortunately, many supporters in our study were unaware of how to provide physical activity support when their counterpart experienced OA symptoms.

Some Latina patients said their supporters did not sympathize with their pain experience and described interpersonal conflict, misunderstanding, and frustration, making it difficult to receive support. For dyads who face conflict during increased OA symptom severity, interventions can build skills in support provision and develop interpersonal communication (45,46). Interventions have demonstrated effectiveness in increasing support for coping with pain among couples (47), but the dyads in our study who discussed interpersonal difficulty typically consisted of mothers and children. Again, research that focuses on noncouple dyads is especially needed among Latina individuals who frequently rely on their adult children for social support (7,22).

Although most participants agreed that physical activity was beneficial to the management of OA, some were uncertain about its safety during episodes of severe OA symptoms. This uncertainty has also been documented among individuals with rheumatoid arthritis living in the UK, where 44% said they worry that physical activity could cause harm to their joints, and 34% were not sure (48). Furthermore, 39% said they did not know what type of exercise they should be doing.

Supporters' physical activity was influenced by the presence of OA symptoms in Latina individuals. In some cases, supporters engaged in less activity either because they depended on their counterpart to motivate them or because caregiving superseded time to exercise. Conversely, some supporters, especially daughters, noticed an increase in their physical activity when their mothers experienced OA symptoms, primarily because they were reminded of their own risk for OA. A population-based study in the US among adults between ages 25 and 44 years showed a higher perceived risk of developing OA in those with a family history versus no family history of OA (49). Thus, adult children and their mothers may have an awareness of children's elevated risk for OA and may become prompted to engage in healthy behaviors in the presence of OA symptoms. Given the potentially reciprocal nature of support and physical activity between dyads, the physical activity of supporters warrants further investigation.

It is essential to consider our study's limitations when interpreting its findings. First, due to the qualitative design, our results may not generalize to other populations. Specifically, our sample consisted of Latina patients with OA and a supporter of their physical activity recruited from an urban area of North Carolina. Findings cannot speak to Latino men with OA or support provided by individuals not represented here (e.g., health care providers). Furthermore, supporters were eligible if they were willing and able to attend the interview. Thus, we did not include supporters who did not live locally but provide support remotely. Notably, data collection and interpretation were primarily conducted by the first author (SHS), a bicultural/bilingual Latina. The perspectives of SHS helped to attach cultural meaning to the data. The second coder and coauthors are non-Latino White and helped to offer alternative explanations for the data.

This study provides the foundation for future investigation and potential dyadic strategies to promote physical activity among Latina patients with OA. Further research will help to generalize findings to other Latino patients with OA and provide valuable insight into the development of interventions for this 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. Soto 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. Soto, Berry, Callahan.
Acquisition of data. Soto.
Analysis and interpretation of data. Soto.

REFERENCES

1. Blackwell D, Villarroel M. Tables of summary health statistics for US adults: National Health Interview Survey, 2015. Hyattsville (MD): National Center for Health Statistics; 2015.

2. Bolen J, Schieb L, Hootman JM, Helmick CG, Theis K, Murphy LB, et al. Differences in the prevalence and severity of arthritis among racial/ethnic groups in the United States: National Health Interview Survey, 2002, 2003, and 2006. Prev Chronic Dis 2010;7:A64.
3. Barbour KE, Helmick CG, Boring M, Brady TJ. Vital signs: prevalence of doctor-diagnosed arthritis and arthritis-attributable activity limitation. United States, 2013-2015. MMWR Morb Mortal Wkly Rep 2017;66:246–53.
4. Kolasinski SL, Neogi T, Hochberg MC, Oatis C, Guyatt G, Block J, 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.
5. Vina ER, Hannon MJ, Hausmann LR, Ibrahim SA, Dagnino J, Arellano A, et al. Modifiable determinants of exercise use in a diverse ethnic population with osteoarthritis. Arthritis Care Res (Hoboken) 2019;71:1495–503.
6. Scarapicchia TM, Amireault S, Faulkner G, Sabiston CM. Social support and physical activity participation among healthy adults: a systematic review of prospective studies. Int Rev Sport Exerc Psychol 2017;10:50–83.
7. Soto SH, Arredondo EM, Haughton J, Shakya H. Leisure-time physical activity and characteristics of social network support for exercise among Latinas. Am J Health Promot 2018;32:432–9.
8. Marquez B, Dunsiger SI, Pekmezi D, Larsen BA, Marcus BH. Social support and physical activity change in Latinas: results from the Seamos Saludables trial. Health Psychol 2016;35:1392.
9. Berli C, Bolger N, Shrout PE, Stadler G, Scholz U. Interpersonal processes of couples' daily support for goal pursuit: the example of physical activity. Pers Soc Psychol Bull 2018;44:332–44.
10. Hohl DH, Lüscher J, Keller J, Heuse S, Scholz U, Luszczynska A, et al. Inter-relations among negative social control, self-efficacy, and physical activity in healthy couples. Br J Health Psychol 2018;23:580–96.
11. Shakya HB, Christakis NA, Fowler JH. Self-comparisons as motivators for healthy behavior. Obesity 2015;23:2477–84.
12. Paech J, Luszczynska A, Lippke S. A rolling stone gathers no moss: the long way from good intentions to physical activity mediated by planning, social support, and self-regulation. Front Psychol 2016;7:1024.
13. Bandura A. The self system in reciprocal determinism. Am Psychol 1978;33:344.
14. Hohl DH, Schultze M, Keller J, Heuse S, Luszczynska A, Knoll N. Inter-relations between partner-provided support and self-efficacy: a dyadic longitudinal analysis. Appl Psychol Health Well Being 2019;11:522–42.
15. Campos B, Kim HS. Incorporating the cultural diversity of family and close relationships into the study of health. Am Psychol 2017;72:543.
16. Vrazel J, Saunders RP, Wilcox S. An overview and proposed framework of social-environmental influences on the physical-activity behavior of women. Am J Health Promot 2008;23:2–12.
17. Im E, Lee B, Hwang H, Yoo KH, Chee W, Stuifbergen A, et al. "A waste of time": Hispanic women's attitudes toward physical activity. Women Health 2010;50:563–79.
18. Marquez DX, Aguiñaga S, Vásquez P, Marques IG, Martinez M. Physical activity among Latinos. In: Bopp M, editor. Physical activity in diverse populations. Milton Park (UK): Routledge; 2017. p. 62–83.
19. Abraído-Lanza AF. Latinas with arthritis: effects of illness, role identity, and competence on psychological well-being. Am J Community Psychol 1997;25:601–27.
20. Niu NN, Davis AM, Bogart LM, Thornhill TS, Abreu LA, Ghazinouri R, et al. Patient disease perceptions and coping strategies for arthritis in a developing nation: a qualitative study. BMC Musculoskelet Disord 2011;12:228.
21. Brooks AT, Andrade RE, Middleton KR, Wallen GR. Social support: a key variable for health promotion and chronic disease management in Hispanic patients with rheumatic diseases. Clin Med Insights Arthritis Musculoskelet Disord 2014;7:21–6.
22. Abraído-Lanza AF, Revenson TA. Coping and social support resources among Latinas with arthritis. Arthritis Care Res (Hoboken) 1996;9:501–8.
23. Abraído-Lanza AF. Social support and psychological adjustment among Latinas with arthritis: a test of a theoretical model. Ann Behav Med 2004;27:162–71.
24. Stubbs B, Hurley M, Smith T. What are the factors that influence physical activity participation in adults with knee and hip osteoarthritis? A systematic review of physical activity correlates. Clin Rehabil 2015;29:80–94.
25. Palinkas LA, Horwitz SM, Green CA, Wisdom JP, Duan N, Hoagwood K. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. Adm Policy Ment Health 2015;42:533–44.
26. Cisternas MG, Murphy L, Sacks JJ, Solomon DH, Pasta DJ, Helmick CG. Alternative methods for defining osteoarthritis and the impact on estimating prevalence in a US population-based survey. Arthritis Care Res (Hoboken) 2016;68:574–80.
27. Bernard H. Interviewing: unstructured and semistructured. In: Research methods in anthropology: qualitative and quantitative approaches. 4th ed. New York: Alta Mira Press; 2006. p. 210–50.
28. Charmaz K. Constructing grounded theory. Thousand Oaks (CA): Sage; 2014.
29. Ritchie J, Lewis J, Nicholls CM, Ormston R. Qualitative research practice: a guide for social science students and researchers. Thousand Oaks (CA): Sage; 2013.
30. Lincoln YS, Gupta EG. But is it rigorous? Trustworthiness and authenticity in naturalistic evaluation. New directions for program evaluation 1986;30:73–84.
31. Amin ME, Nørgaard LS, Cavaco AM, Witry MJ, Hillman L, Cernasev A, et al. Establishing trustworthiness and authenticity in qualitative pharmacy research. Res Social Adm Pharm 2020;16:1472–82.
32. Martire LM, Helgeson VS. Close relationships and the management of chronic illness: associations and interventions. Am Psychol 2017;72:601.
33. Carr R, Prestwich A, Kwasnicka D, Thøgersen-Ntoumani C, Gucciardi D, Quested E, et al. Dyadic interventions to promote physical activity and reduce sedentary behaviour: systematic review and meta-analysis. Health Psychol Rev 2019;13:91–109.
34. Feeney BC, Collins NL. A new look at social support: a theoretical perspective on thriving through relationships. Pers Soc Psychol Rev 2015;19:113–47.
35. Arden-Close E, McGrath N. Health behaviour change interventions for couples: a systematic review. Br J Health Psychol 2017;22:215–37.
36. Gellert P, Ziegelmann JP, Warner LM, Schwarzer R. Physical activity intervention in older adults: does a participating partner make a difference? Eur J Ageing 2011;8:211.
37. Hoebeke R. Low-income women's perceived barriers to physical activity: focus group results. Appl Nurs Res 2008;21:60–5.
38. White DK, Tudor-Locke C, Zhang Y, Fielding R, LaValley M, Felson DT, et al. Daily walking and the risk of incident functional limitation in knee osteoarthritis: an observational study. Arthritis Care Res (Hoboken) 2014;66:1328–36.
39. Callahan LF, Rivadeneira A, Altpeter M, Vilen L, Cleveland RJ, Sepulveda VE, et al. Evaluation of the Arthritis Foundation's Camine Con Gusto program for Hispanic adults with arthritis. Hisp Health Care Int 2016;14:132–40.
40. Craddock E, vanDellen MR, Novak SA, Ranby KW. Influence in relationships: a meta-analysis on health-related social control. Basic Appl Soc Psych 2015;37:118–30.

41. Fekete E, Geaghan TR, Druley JA. Affective and behavioural reactions to positive and negative health-related social control in HIV men. Psychol Health 2009;24:501–15.
42. Martire LM, Stephens MA, Mogle J, Schulz R, Brach J, Keefe FJ. Daily spousal influence on physical activity in knee osteoarthritis. Ann Behav Med 2013;45:213–23.
43. Ranby KW, Aiken LS. Incorporating husband influences into a model of physical activity among older women. Br J Health Psychol 2016; 21:677–93.
44. Cheriel C, Huguet N, Gupta S, McClure H, Leman RF, Ngo DL. Arthritic pain among Latinos: results from a community-based survey. Arthritis Care Res (Hoboken) 2009;61:1491–6.
45. Luger T, Cotter KA, Sherman AM. It's all in how you view it: pessimism, social relations, and life satisfaction in older adults with osteoarthritis. Aging Ment Health 2009;13:635–47.
46. Carriere JS, Lazaridou A, Martel MO, Cornelius M, Campbell C, Smith M, et al. The moderating role of pain catastrophizing on the relationship between partner support and pain intensity: a daily diary study in patients with knee osteoarthritis. J Behav Med 2020;43: 807–16.
47. Martire LM, Schulz R, Helgeson VS, Small BJ, Saghafi EM. Review and meta-analysis of couple-oriented interventions for chronic illness. Ann Behav Med 2010;40:325–42.
48. Law R, Markland DA, Jones JG, Maddison PJ, Thom JM. Perceptions of issues relating to exercise and joint health in rheumatoid arthritis: a UK-based questionnaire study. Musculoskeletal Care 2013;11: 147–58.
49. Michl GL, Katz JN, Losina E. Risk and risk perception of knee osteoarthritis in the US: a population-based study. Osteoarthritis Cartilage 2016;24:593–6.

Arthritis Care & Research
Vol. 74, No. 2, February 2022, pp 291–300
DOI 10.1002/acr.24478

American College of Rheumatology
Empowering Rheumatology Professionals

Systematic Review and Meta-Analysis of Health State Utility Values for Osteoarthritis-Related Conditions

Ting Zhao,[1] Tania Winzenberg,[1] Barbara de Graaff,[1] Dawn Aitken,[1] Hasnat Ahmad,[1] and Andrew J. Palmer[2]

Objective. Health state utility values (HSUVs) are a key input in health economic modeling, but HSUVs of people with osteoarthritis (OA)–related conditions have not been systematically reviewed and meta-analyzed. Our objective was to systematically review and meta-analyze the HSUVs for people with OA.

Methods. Searches within health economic/biomedical databases were performed to identify eligible studies reporting OA-related HSUVs. Data on study design, participant characteristics, affected OA joint sites, treatment type, HSUV elicitation method, considered health states, and the reported HSUVs were extracted. HSUVs for people with knee, hip, and mixed OA in pre- and posttreatment populations were meta-analyzed using random effects models.

Results. A total of 151 studies were included in the systematic review, and 88 in meta-analyses. Of 151 studies, 56% were conducted in Europe, 75% were in people with knee and/or hip OA, and 79% were based on the EuroQoL 5-dimension instrument. The pooled mean baseline HSUVs for knee OA core interventions, medication, injection, and primary surgery treatments were 0.64 (95% confidence interval [95% CI] 0.61–0.66), 0.56 (95% CI 0.45−0.68), 0.58 (95% CI 0.50–0.66), and 0.52 (95% CI 0.49–0.55), respectively. These were 0.71 (95% CI 0.59–0.84) for hip OA core interventions and 0.52 (95% CI 0.49–0.56) for hip OA primary surgery. For all knee OA treatments and hip OA primary surgery, pooled HSUVs were significantly higher in the post- than the pretreatment populations.

Conclusion. This study provides a comprehensive summary of OA-related HSUVs and generates an HSUV database for people with different affected OA joint sites undergoing different treatments to guide HSUV choices in future health economic modeling of OA interventions.

INTRODUCTION

Osteoarthritis (OA) is one of the most common chronic joint diseases. It mostly affects knees, hips, and small joints of hands. OA is characterized by joint pain, stiffness, swelling, loss of function, and disability, which in turn negatively impact individuals' health-related quality of life (HRQoL) (1) and incur a substantial socioeconomic burden (2,3). Currently, there is no cure for OA, but many treatments and approaches, including lifestyle, medications, injections, and surgery, are available to help relieve disease syptoms.

Health state utility values (HSUVs) are typically used to reflect HRQoL and to calculate quality-adjusted life years, a preferred measure of clinical effectiveness in cost utility (CUA)/clinical effectiveness analyses (CEA) (4). HSUVs measure the strength of a preference for a particular health state, represented as a number between 0 (death) and 1 (optimal health). Health states worse than death may exist, with negative HSUVs assigned (5). HSUVs can be obtained through several methods (6). Direct methods ask individuals to describe and assess health states and place weights on them, using valuation techniques such as the standard gamble, time trade-off, and rating scales (6). Indirect methods involve the use of preference-based multi-attribute utility instruments (MAUIs), where patients answer questions relating to multiple dimensions of their current health state, and the responses are then scored using a value set obtained from respective general populations. Commonly used MAUIs include the EuroQoL 5-dimension (EQ-5D), the Health Utility Index, the Short Form 6-dimension (SF-6D), and the Assessment of Quality of Life (AQoL) instruments (7). Finally, mapping techniques are used to transform nonpreferenced-based HRQoL measures into HSUVs.

[1]Ting Zhao, PhD, Tania Winzenberg, PhD, Barbara de Graaff, PhD, Dawn Aitken, PhD, Hasnat Ahmad, PhD: University of Tasmania, Hobart, Tasmania, Australia; [2]Andrew J. Palmer, MBBS: University of Tasmania, Hobart, Tasmania, and The University of Melbourne, Melbourne, Victoria, Australia.

No potential conflicts of interest relevant to this article were reported.

Address correspondence to Ting Zhao, PhD, Menzies Research Institute Tasmania, University of Tasmania, Medical Science 1 Building, 17 Liverpool Street, Hobart, Tasmania 7000, Australia (email: ting.zhao@utas.edu.au); or to Andrew J. Palmer, MBBS, Menzies Research Institute Tasmania, University of Tasmania, Medical Science 1 Building, 17 Liverpool Street, Hobart, Tasmania 7000, Australia (email: Andrew.palmer@utas.edu.au).

Submitted for publication May 12, 2020; accepted in revised form October 1, 2020.

SIGNIFICANCE & INNOVATIONS

- This is the first study comprehensively reviewing osteoarthritis (OA)–related health state utility values (HSUVs).
- It identified important areas where the current evidence is lacking, including underrepresented geographical locations and ethnicities, OA joint sites, treatment options, and multi-attribute utility instruments.
- This study is the first to meta-analyze the OA-related HSUVs for different affected joint sites before and after various treatments.
- It generated an HSUV database for OA patients with different affected joint sites undergoing different treatments to guide HSUV choices in future health economic modeling of OA interventions.

As the stated preference data for a set of health states for an appropriate population are not always available, HSUVs obtained from the literature are widely used in economic evaluations (4). These HSUV estimates may differ from each other due to several factors, including differences in the utility elicitation techniques, MAUIs, the choice of respondent, sample size, and the quality of studies (4). With an increasingly growing literature of HSUVs, the selection of which values to use in economic evaluations becomes challenging. The correct choice of HSUVs is important to accurately calculate quality-adjusted life years and other CUA outcomes. To obtain the best estimate for a decision-analytic model from the literature, the methods of identification of the data should be systematic and transparent. To date, there is no systematic review and meta-analysis that summarizes estimates of OA-related HSUVs and evaluates the extent of differences between various subgroups of patients based on affected OA joint sites, treatments, and utility measurements. Our systematic review and meta-analysis aimed to generate a database of OA-related HSUVs to address this gap.

MATERIALS AND METHODS

Protocol registration. The study protocol was registered on April 17, 2019 at PROSPERO (#CRD42019129408). Our systematic review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines (8).

Literature search. Based on previous recommendations (9), 4 databases were searched from their inception up to March 2019: Embase, Health Technology Assessment database, Medline, and Scopus. This search was supplemented by hand searching the bibliography lists of all included articles and relevant reviews. The search strategy was developed in consultation with co-authors based on the previous literature (10,11). Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24478/abstract, provided the search strategy used for Embase, which was also revised to suit other databases.

Screening criteria. Title/abstract screening and full-text screening were conducted in Covidence (12) (an online systematic review program to manage and facilitate the selection of studies) by 2 reviewers (TZ and HA) independently based on predefined criteria. Any disagreements were discussed between the 2 reviewers, and a third reviewer (AJP) was consulted in cases of no consensus. Studies were included if they involved humans, reported OA-related HSUV estimates (excluding those based on mapping techniques), and were published in English, Chinese, or German. Conference abstracts were included when adequate data were available for extraction. If the OA patients were part of a broader study population, we included studies reporting on a cohort with ≥80% OA representation. Health economic modeling studies based on HSUVs reported elsewhere and those based on systematic reviews or meta-analyses were excluded. Review reports, books, and case reports were excluded.

Data extraction. A predefined Microsoft Excel spreadsheet was piloted to extract data from 20% of studies by the first author (TZ). Adjustments and improvements were made to the initial spreadsheet where necessary, and the improved spreadsheet was then used to extract data independently by TZ and HA. Discrepancies were resolved by consensus, and an additional reviewer (AJP) was consulted to reach an agreement in cases of no consensus. The following data were extracted: authors' names, year of publication, study setting, study design (e.g., trial, observational), sample size, characteristics of the patients (e.g., age, sex, body mass index), affected OA joint sites, treatment type, utility elicitation method, the health states considered, and the reported HSUVs (mean ± SD/SE, 95% confidence intervals [95% CIs], median, minimum/maximum, quartile) (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24478/abstract).

Meta-analyses. Based on data availability, the selection of studies for meta-analyses included studies related to knee, hip, and mixed OA (including a variety of OA patients without specifying their affected OA joint site), and studies of core intervention, medication, intraarticular injection, and primary surgery treatments. We followed OA management guidelines (13) to group the included interventions under 1 of these 4 categories of treatment. The core intervention category included exercise, weight management, and education/programs related to exercise and weight management. Medications included all drugs used to decrease pain and improve function in patients with OA. Intraarticular injections included corticosteroids, viscosupplements, and blood-derived products. Finally, primary surgery

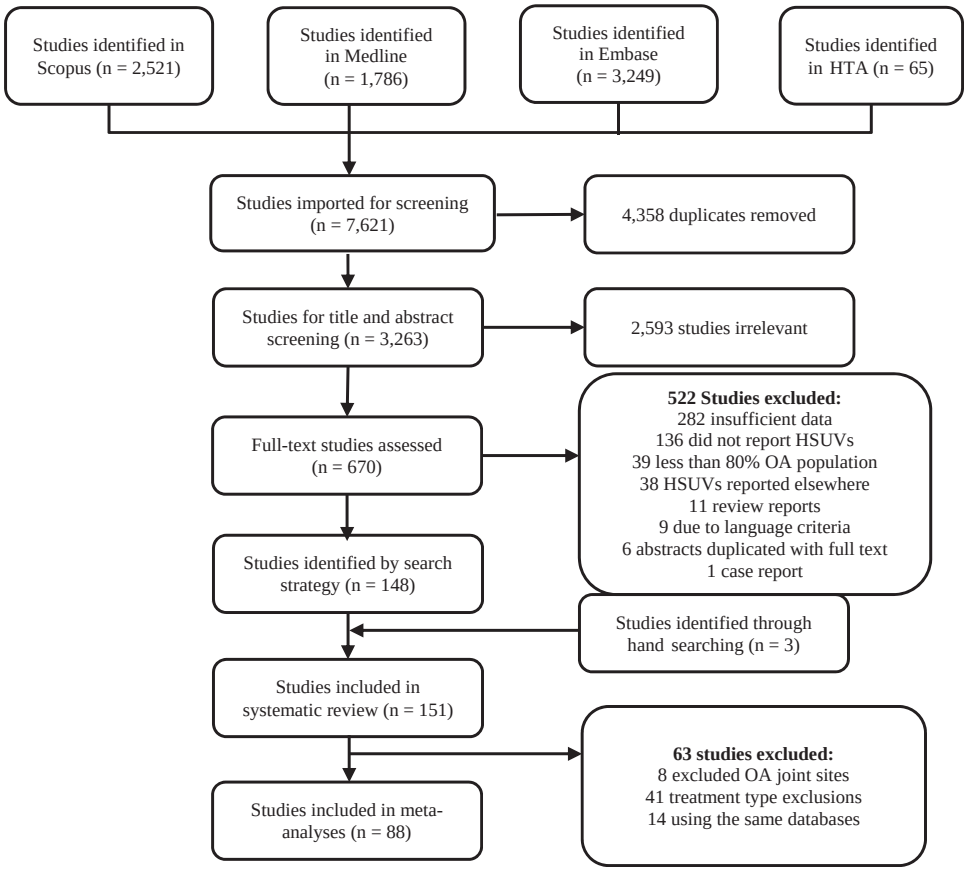

Figure 1. Flow chart results of study search based on Preferred Reporting Items for Systematic Reviews and Meta-Analyses methodology. The exclusions by osteoarthritis (OA) joint sites and treatment type were because of the small numbers of studies in these joint sites and treatments, which meant that meta-analysis was not feasible. Eight exclusions by OA joint sites involved 2 shoulder and 6 hand OA–related studies; 41 exclusions by treatment type involved studies of massage, foot insoles, brace, mud therapy, balneotherapy, spa therapy, revision surgery, and observational studies that did not focus on any treatment. HTA = Health Assessment Technology database; HSUV = health state utility value.

included joint resurfacing and primary joint replacement. Supplementary Appendix B, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24478/abstract, provides the full list of included interventions under each category of treatment for knee, hip, and mixed OA from studies included in meta-analyses. Observational studies that did not include delivery of an intervention were excluded from the meta-analysis. HSUVs were summarized by key OA affected joint sites (knee, hip, and mixed OA) for baseline (pretreatment) and at the most commonly available posttreatment time points (i.e., 3, 6, 12, and 24 months). When more than 1 HSUV study was based on the same data, the study with the highest number of participants was included in the meta-analyses. Subgroup meta-analyses by utility elicitation methods were also conducted, where possible.

The meta-analyses were programmed in Stata software, version 15.1, using the "metan" command that required mean and SD/SE as meta-analytical inputs (14). Therefore, when the mean values and SD/SE were not reported, we used 95% CIs, median, minimum/maximum, first quartile, and third quartile values to estimate these parameters (15–17). HSUVs at baseline in observational studies and in both control and intervention groups of trials were pooled (termed pretreatment HSUVs). Posttreatment HSUVs were calculated by pooling HSUVs from longitudinal observational studies of interventions and intervention arms of trials (including active treatment groups but not control groups), for each time point. Heterogeneity among the pooled studies was assessed using the I^2 statistic (where an $I^2 \geq 50\%$ indicated substantial heterogeneity) (17). To account for within-study and between-study heterogeneity, random-effects models were estimated.

RESULTS

Eligible studies. Initially 7,621 potential references were identified (Figure 1). After we removed duplicates (n = 4,358), 3,263 were left for title and abstract screening. We excluded 2,593 during title and abstract screening, leaving 670 for the full-text assessment. Of those, 522 were excluded due to not meeting the inclusion criteria (Figure 1). Three additional studies

identified through reference hand-searching were subsequently included, resulting in a final total of 151 (including 7 abstracts) being included in the systematic review. Eighty-eight of these studies were included in meta-analyses (including 4 conference abstracts).

Results of systematic review. The majority of included studies (n = 131, 87%) were published after 2010 (Figure 2A). More than half (n = 86, 57%) were conducted in Europe, followed by Asia (n = 20, 13%) and the Americas (n = 16, 11%). Four studies focused on Australians with OA, 1 study was conducted in multiple countries, and 24 studies did not report the study setting (Figure 2B). Fifty-eight included studies (38%) were trials, 65 (43%) were observational studies of interventions, and 28 (19%) were observational studies that did not have an intervention component.

Fifty-nine studies (39%) focused on knee OA, and 41 (27%) focused on hip OA. Thirteen studies (9%) focused on both knee and hip OA and reported HSUVs separately. Two studies (1%) and 6 (4%) were focused on shoulder and hand OA, respectively. Thirty studies (20%) focused on mixed OA (Figure 2C).

Of the 72 knee OA–related studies, 10 (14%) focused on core interventions, 6 (8%) and 5 (7%) focused on medication and injection treatments, respectively. Thirty-two studies (44%) focused on surgical treatments. Seven investigated other treatments such as massage, foot insoles, brace, and mud therapy. Twelve (17%) reporting the cross-sectional HSUVs of knee OA did not focus on any specific treatment (Table 1).

Of the 54 hip OA–related studies, the majority (n = 46, 85%) focused on surgical treatments. Two studies (4%) focused on core interventions, 1 investigated balneotherapy, and 5 (9%) reporting the cross-sectional HSUVs of hip OA did not focus on any specific treatment. There were no studies reporting the HSUVs related to hip OA medication and injection treatments (Table 1).

Two shoulder OA–related studies focused on surgical treatments. Among 6 hand OA–related studies, 2 reported the cross-sectional HSUVs of hand OA populations, and 1 study each focused on spa, mud, a core intervention, and surgery treatment.

Of the 30 mixed OA–related studies, 14 (47%) focused on core interventions, and 12 (40%) reported the cross-sectional HSUVs of an OA population without specifying any treatment type. Two studies focused on surgical treatments, 1 focused on medication, and 1 focused on spa therapy (Table 1).

Nine HSUV measures were used in the included studies, with most studies (n = 120, 79%) using the EQ-5D, followed by the SF-6D (n = 12, 8%), Health Utility Index 2/3 (n = 4, 3%), and Quality of Well-Being (n = 3, 2%). One study each used the AQoL-6D, and the 15D. Ten studies (7%) included >1 measure, including the Paper Adaptive Test, standard gamble, and rating scales.

Results of meta-analysis. *Studies included in meta-analyses.* Fifty-one knee OA–related studies (Figure 2D) qualified for meta-analyses. Nine, 6, 5, and 31 related to core

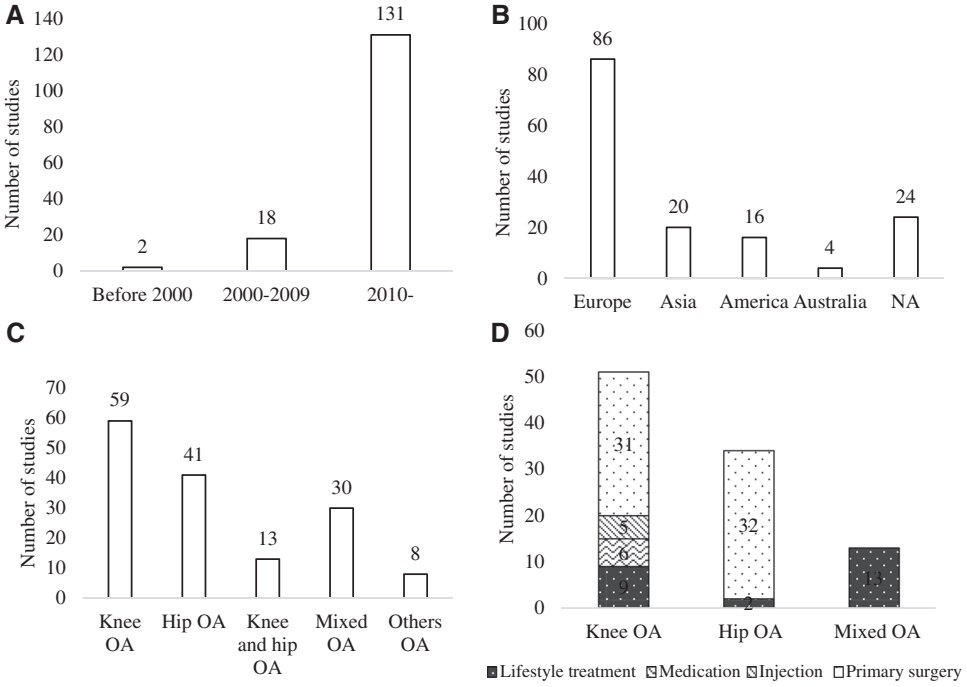

Figure 2. The distribution of included studies in systematic reviews by **A**, year of publication, **B**, study setting, **C**, osteoarthritis (OA) joint sites, and **D**, the distribution of included studies in meta-analysis by OA joint sites and treatments. **C**, This knee and hip OA group included studies reporting health state utility values for each type separately. Other OA included hand and shoulder OA studies. NA = study setting not reported.

Table 1. Studies included in systematic review and meta-analyses for each OA affected joint site and treatment*

	Systematic reviews	Meta-analyses
Knee OA		
Core intervention	10 (14)	9 (18)
Medication	6 (8)	6 (12)
Injection	5 (7)	5 (10)
Surgery	32 (44)	31 (61)
Other treatments	7 (10)	0 (0)
No treatments	12 (17)	0 (0)
Subtotal, no.	72	51
Hip OA		
Core intervention	2 (4)	2 (6)
Surgery	46 (85)	32 (94)
Other treatments	1 (2)	0 (0)
No treatments	5 (9)	0 (0)
Subtotal, no.	54	34
Mixed OA		
Core intervention	14 (47)	13 (100)
Other treatments	4 (13)	0 (0)
No treatments	12 (40)	0 (0)
Subtotal, no.	30	13
Other joint sites of OA, no.	8	0
Total, no.†	164	98

* Values are the number (%) unless indicated otherwise. OA = osteoarthritis. Mixed OA included a variety of OA patients without specifying their OA type. Other joint sites of OA included shoulder and hand OA studies.
† Thirteen studies reporting knee and hip health state utility values separately have been counted in both the hip and knee OA groups and 10 were included in the meta-analyses.

interventions, medications, injections, and primary surgery, respectively (see Supplementary Tables 2–5, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24478/abstract). Thirty-four hip OA–related studies (Figure 2D) qualified for meta-analyses. Two and 32 related to core interventions and primary surgery, respectively (see Supplementary Tables 6 and 7, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24478/abstract). Thirteen studies for mixed OA core interventions qualified for meta-analyses (Figure 2D and Supplementary Table 8, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24478/abstract). The posttreatment time points included in meta-analyses varied between different OA joint sites and treatments based on data availability (Table 2).

HSUVs of knee OA. The pooled mean baseline (pretreatment) HSUV of knee OA core interventions was 0.64 (number of HSUVs pooled, n = 19 [95% CI 0.61–0.66], I^2 = 99%). The pooled HSUVs of knee OA core interventions at 3 months postintervention were higher (0.73, n = 6 [95% CI 0.70–0.76], I^2 = 91%). The pooled 6-month and 1-year HSUVs did not differ significantly from baseline (0.65, n = 4 [95% CI 0.60–0.71], I^2 = 97% at 6 months; 0.71, n = 5 [95% CI 0.64–0.79], I^2 = 1 at 1 year). In the subgroup analyses, there were significant difference in HSUV estimates between different MAUIs at each time point (see Table 2 and Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24478/abstract).

The pooled mean HSUV for knee OA medication treatment was significantly different at baseline (0.56, n = 9 [95% CI 0.45–0.68], I^2 = 1%) than at 3-month follow-up (0.75, n = 3 [95% CI 0.70–0.80], I^2 = 87%). All knee medication related–HSUVs were based on the EQ-5D (see Table 2 and Supplementary Figure 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24478/abstract).

The pooled HSUVs for knee OA (intraarticular) injections were similar at baseline (0.58, n = 7 [95% CI 0.50–0.66], I^2 = 94%) and 1-year posttreatment (0.63, n = 1 [95% CI 0.59–0.67]). The baseline HSUV estimates significantly differed between EQ-5D and Health Utility Index 3 measures (see Table 2, and Supplementary Figure 3, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24478/abstract).

For knee OA primary surgeries, the pooled mean HSUV was 0.52 (n = 55 [95% CI 0.49–0.55], I^2 = 99.7%) at baseline. A significant difference was found between HSUVs of baseline and various postsurgery time points: 6 months (0.71, n = 21 [95% CI 0.69–0.74], I^2 = 95%), 1 year (0.77, n = 18 [95% CI 0.73–0.81], I^2 = 99%), and 2 years (0.74, n = 17 [95% CI 0.71–0.78], I^2 = 99%). Significant differences existed between different MAUIs at each time point (see Table 2 and Supplementary Figure 4, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24478/abstract).

HSUVs of hip OA. Only 2 studies focused on hip OA core interventions. HSUVs did not differ significantly between the baseline (0.71, n = 3 [95% CI 0.59–0.84], I^2 = 99%), 3-month (0.72, n = 2 [95% CI 0.59–0.84], I^2 = 98%), or 1-year (0.72, n = 2 [95% CI 0.58–0.85], I^2 = 98%) postinterventions. All HSUVs were based on the EQ-5D (see Table 2 and Supplementary Figure 5, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24478/abstract).

For hip OA primary surgery treatments, there was a significant difference between the pooled mean HSUVs of baseline (0.52, n = 46 [95% CI 0.49–0.56], I^2 = 1%) and postsurgery periods: 6 months (0.79, n = 9 [95% CI 0.76–0.82], I^2 = 94%), 1 year (0.83, n = 22 [95% CI 0.80–0.85], I^2 = 99%), and 2 years (0.84, n = 11 [95% CI 0.80–0.87], I^2 = 98%). Significant differences existed between different MAUIs at each time point (see Table 2 and Supplementary Figure 6, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24478/abstract).

HSUVs of mixed OA. For mixed OA core interventions, there was a significant difference between the pooled mean HSUVs of baseline (0.61, n = 27 [95% CI 0.59–0.64], I^2 = 99%) and 3-month postintervention (0.71, n = 10 [95% CI 0.68–0.73], I^2 = 97%), and 1-year postintervention (0.69, n = 12 [95% CI 0.66–0.71], I^2 = 98%). The same trend was found for EQ-5D HSUVs but not for SF-6D (see Table 2 and Supplementary Figure 7, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24478/abstract).

Table 2. The number of pooled HSUVs representing population in meta-analyses and the pooled mean HSUVs*

	Overall			EQ-5D			SF-6D		
	HSUVs	Population	Mean (95% CI)	HSUVs	Population	Mean (95% CI)	HSUVs	Population	Mean (95% CI)
Knee OA									
Core intervention									
Baseline	19	31,349	0.64 (0.61–0.66)	12	30,869	0.60 (0.57–0.62)	2	119	0.65 (0.55–0.75)
3 months	6	30,380	0.73 (0.70–0.76)	4	30,233	0.70 (0.69–0.72)	0	0	–
6 months	4	542	0.65 (0.60–0.71)	4	542	0.65 (0.60–0.71)	0	0	–
1 year	5	20,549	0.71 (0.64–0.79)	3	20,402	0.64 (0.56–0.72)	0	0	–
Medication									
Baseline	9	3,749	0.56 (0.45–0.68)	9	3,749	0.56 (0.45–0.68)	0	0	–
3 months	3	249	0.75 (0.70–0.80)	3	249	0.75 (0.70–0.80)	0	0	–
Injection									
Baseline	7	473	0.58 (0.50–0.66)	5	224	0.63 (0.56–0.70)	0	0	–
1 year	1	122	0.63 (0.59–0.67)	0	0	–	0	0	–
Primary surgery									
Baseline	55	53,434	0.52 (0.49–0.55)	43	45,434	0.49 (0.44–0.53)	10	7,797	0.62 (0.60–0.64)
6 months	21	4,260	0.71 (0.69–0.74)	14	3,378	0.72 (0.70–0.73)	5	719	0.71 (0.68–0.75)
1 year	18	3,790	0.77 (0.73–0.81)	12	2,179	0.78 (0.73–0.84)	4	1,456	0.75 (0.72–0.79)
2 years	17	15,160	0.74 (0.71–0.78)	11	8,872	0.77 (0.73–0.81)	5	6,270	0.71 (0.66–0.76)
Hip OA									
Core intervention									
Baseline	3	13,773	0.71 (0.59–0.84)	3	13,773	0.71 (0.59–0.84)	0	–	–
3 months	2	13,671	0.72 (0.59–0.84)	2	13,671	0.72 (0.59–0.84)	0	–	–
1 year	2	8,421	0.72 (0.58–0.85)	2	8,421	0.72 (0.58–0.85)	0	–	–
Primary surgery									
Baseline	46	59,846	0.52 (0.49–0.56)	34	52,671	0.50 (0.46–0.54)	6	6,791	0.58 (0.56–0.60)
6 months	9	3,922	0.79 (0.76–0.82)	6	3,727	0.80 (0.77–0.83)	0	0	–
1 year	22	42,788	0.83 (0.80–0.85)	20	42,468	0.83 (0.80–0.85)	1	224	0.80 (0.78–0.82)
2 years	11	16,732	0.84 (0.80–0.87)	7	10,228	0.88 (0.85–0.90)	4	6,504	0.78 (0.75–0.81)
Mixed OA									
Core intervention									
Baseline	27	9,644	0.61 (0.59–0.64)	18	5,672	0.58 (0.53–0.62)	5	3,576	0.69 (0.66–0.72)
3 months	10	6,926	0.71 (0.68–0.73)	6	3,587	0.67 (0.63–0.71)	2	3,192	0.73 (0.65–0.82)
1 year	12	7,305	0.69 (0.66–0.71)	6	3,819	0.65 (0.60–0.71)	4	3,339	0.71 (0.65–0.77)

* Values are the number, unless indicated otherwise. Pooling at all time points includes observational data on intervention, and at baseline includes both control and active treatment groups from trials, but at follow-up includes only data from active treatment groups from trials. 95% CI = 95% confidence interval; EQ-5D = EuroQol 5-domain instrument; HSUVs = health state utility values; OA = osteoarthritis; SF-6D = Short Form 6-dimension utility score.

DISCUSSION

This is the first wide-ranging systematic review of OA-related HSUVs and meta-analyses on HSUVs for people with different OA-affected joint sites before and after various treatments. Our systematic review identified important areas where the current evidence is lacking, namely underrepresented geographical locations and ethnicities, affected OA joint sites, treatment options, and HSUVs based on more sensitive MAUIs. Our meta-analyses provide an HSUV database for alternative pre- and post-OA treatments that could offer a variety of HSUV inputs for future cost-utility models of OA-related conditions. HSUVs associated with 4 key treatment categories (core interventions, medication, injection, and surgery) often differed, as expected, pre- and posttreatment. Furthermore, we found significant inter-MAUI differences in the mean HSUVs, which is as expected from alternative descriptive systems and utility algorithms. Therefore this review provides important information that could be used by health economists and policy makers to determine the cost-effectiveness of various OA treatments and long-term disease outcomes using modeling techniques.

Our systematic review identified numerous gaps in the data on OA-related HSUVs, including geographical locations and ethnicities, affected OA joint sites, treatment options, and HSUVs based on more sensitive MAUIs. We found that more than half of included studies (57%) were conducted in Europe, and none in Africa. Because HSUVs should ideally be based on local population preferences, the generalizability of our results to underrepresented populations (e.g., African and Asian) may therefore be limited. Seventy-six percent of included studies focused on knee and hip OA, while other joint sites (e.g., shoulder and hand) attracted limited attention. While these results align well with the higher clinical impact, prevalence, and societal burden of knee and hip OA (18–20), the increasing prevalence and disease burden of hand and shoulder OA as a result of population aging (21,22) mandate further primary studies investigating the HSUVs of these joint sites.

The HSUVs that we have meta-analyzed differed as expected between alternative OA joint sites, treatments, HSUVs measures, and time points. We found a mean HSUV difference of +0.09 units in patients with knee OA using core interventions between baseline and 3-month postintervention, and this difference exceeds the minimal clinically important difference for all the MAUIs reported in previous studies (from +0.04 units [EQ-5D] to 0.08 units [AQoL-8D]) (23–27). Our findings are consistent with the randomized controlled trial (RCT) evidence showing the short-lived effects of knee OA core interventions (28,29). Other possible explanations include the limited number of core intervention studies with a follow-up a period of >3 months (and hence, wider 95% CIs for our 6-month and 1-year posttreatment HSUVs), and a likely reduction in the core intervention adherence in the long-term (30,31).

Most studies of knee OA medication treatments (83.3%) had relatively shorter follow-up periods (3 months), with only 1 study with a follow-up period of >3months. Consistent with RCT evidence of effectiveness of medication treatments (32,33), the pooled HSUV of studies with follow-up at 3 months postmedication treatment was significantly higher than the pooled HSUV of studies with baseline measures. As we did not have enough data on long-term HSUVs in patients using OA medications, we leave this question on the agenda for future research when long-term data become available. We found similar HSUVs at baseline and 1-year follow-up for knee OA injection treatments. However, these results should be carefully interpreted and used in economic modeling, as they are derived from only a limited number of studies ($n = 5$ at baseline and $n = 1$ at 1-year follow-up). HSUVs of knee OA patients recorded the largest difference (+0.25 units) between baseline and 1-year postprimary surgery, and it remained relatively stable to 2-years postprimary surgery. These findings are once again consistent with the previous evidence of the effectiveness of knee surgery, suggesting that HSUVs record a significant improvement within 1 year of knee surgery, and this change in HSUVs is sustained for years (34).

Surgery was the most common treatment in hip OA HSUV studies (85%). HSUVs in patients with primary hip OA surgery were significantly higher at 6 months postsurgery than at baseline and remained improved over the long term. The difference between pooled HSUV before and after surgery over 1 year was smaller in knee OA primary surgery (+0.25 units) than hip OA (+0.31 units). These findings align well with previous research (35) advocating a relatively higher efficacy of hip OA joint surgery. Only 2 studies (both based on the EQ-5D) investigated HSUVs in patients with hip OA core intervention, which aligned well with the previous findings of the dearth of studies measuring the HSUVs in patients using hip OA core interventions (36,37). No studies on hip OA medication and injection treatments were identified in our review as expected (38,39); thus, no meta-analysis for these treatments was possible. We recommend future studies to investigate HSUVs in patients using medications and injections, subject to the availability of better long-term observed data.

The HSUVs for mixed OA core interventions showed the same trend observed for knee OA, with a significant difference (+0.10 units) between baseline and 3-month postintervention HSUVs. This finding aligns with the existing findings of short-term benefits associated with OA core interventions (29). A small number of studies of medication treatment ($n = 1$) for mixed OA did not allow us to generate HSUV estimates of before and after medication treatments for use in health-economic modeling. Future primary HSUV studies in this area should therefore be imperative in bridging this evidence gap.

The EQ-5D was the most commonly used MAUI in the included studies (79%), with little to no representation from other more detailed MAUIs (e.g., AQoL-8D) that can more fully capture and assess the complex physical and psychosocial health

aspects of OA patients (23,40). Our MAUI-specific subgroup analysis revealed significant differences between HSUVs based on alternative MAUIs (EQ-5D and SF-6D, for example), which is as expected from the MAUIs that are far from identical in terms of their descriptive systems and measurement scales (41). As the key objective of our review was not to explore the extent of agreement between alternative MAUIs, we leave the head-to-head comparison of HSUVs obtained through alternative MAUIs on the agenda for future research. Moreover, there is no consensus on the choice of MAUI to be used in measuring HSUVs of patients with OA (41,42). Future research should also endeavor to identify MAUIs that could be preferentially recommended for OA patients.

When the baseline HSUVs for various treatments were compared, the mean baseline HSUVs for patients with knee and hip OA using core interventions were significantly higher than those using surgery treatments, which is likely to be due to the specified selection criteria for RCTs. Due to the recommended stepwise approach for OA treatments (43), patients are more likely to receive core interventions at earlier stages of their OA (with better HRQoL) and surgery treatments at more severe stages of their OA (with relatively worse HRQoL), which can also explain this pattern. This result reinforces the need to use different HSUVs in modeling for treatments used at different disease stages.

The strength of this study is that this is the first comprehensive review and meta-analysis of all types of OA-related preference-based HSUVs by OA-affected joint sites, OA treatments, and utility elicitation methods. The study provided an HSUV database for alternative pre- and post-OA treatments that could offer a variety of HSUV inputs to future cost-utility models of OA-related conditions and identified important areas where evidence gaps exist in these estimates to inform future research needs. Our study has several limitations. It is important that the differences in HSUVs at different time points are not interpreted as true pre-/postchange or as direct evidence of intervention effectiveness, as the data do not examine differences in change in HSUVs between controls and intervention groups over time, and the data included in pooling at each time point come from different studies. Heterogeneity of the included studies due to the differences in terms of their study design, settings, and HSUV elicitation techniques can affect the interpretation of generated HSUVs. While we have conducted subgroup analyses where possible to highlight some possible sources of heterogeneity, we had limited capacity to explain and account for all sources of heterogeneity. The random-effects model in our meta-analyses aims to account for heterogeneity but may have consequences for the precision of model estimates (44). Therefore, in modeling, as well as the pooled mean sensitivity analyses, consideration of the potential imprecision of our estimates is important.

A further limitation is that due to the paucity of available studies, we could not conduct meta-analyses for all treatments of hip, knee, and other OA joint sites, nor could we group the treatment types in a more detailed way or perform meta-regression to account for >1 potential effect-modifying variable at a time. Quite a few potentially eligible CUA/CEA reports could not be included as they did not adequately report the required HSUVs (pre- and/or posttreatment) (45,46), despite clear reporting guidelines that recommend these be reported (47,48). We recommend that future CUA/CEA studies refer to these guidelines to help improve the availability of this important data. Also, the exploration of long-term HSUVs of patients using different OA treatments was mostly not possible. Finally, due to the paucity of data, we could not generate the estimates of HSUVs associated with alternative therapy adherence levels and medication adverse event types.

Our systematic review found that studies of OA-related HSUVs are of wide variety and differ from each other in terms of their setting, design, focused OA joint sites, utility measurement technique, generalizability, and other factors. The HSUVs that we have generated will be useful in conducting future health economic modeling for people experiencing various OA-related conditions. Our results should, however, be interpreted with caution as being derived from a relatively small number of heterogeneous studies. More research is needed to investigate changes in HSUVs of OA patients for longer follow-up periods.

ACKNOWLEDGMENT

The authors thank Mr. Petr Otahal (Statistical Officer, Menzies Institute for Medical Research, University of Tasmania) for his help in conducting the meta-analyses.

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. Palmer 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. Zhao, Winzenberg, de Graaff, Aitken, Ahmad, Palmer.

Acquisition of data. Zhao, Ahmad.

Analysis and interpretation of data. Zhao, Winzenberg, Ahmad.

REFERENCES

1. Wilson R, Blakely T, Abbott JH. Radiographic knee osteoarthritis impacts multiple dimensions of health-related quality of life: data from the Osteoarthritis Initiative. Rheumatology (Oxford) 2018;57:891–9.
2. Abbott JH, Usiskin IM, Wilson R, Hansen P, Losina E. The quality-of-life burden of knee osteoarthritis in New Zealand adults: a model-based evaluation. PLoS One 2017;12:e0185676.
3. Henderson JV, Harrison CM, Britt HC, Bayram CF, Miller GC. Prevalence, causes, severity, impact, and management of chronic pain in Australian general practice patients. Pain Med 2013;14:1346–61.
4. Brazier J, Green C. A systematic review of health state utility values for osteoporosis-related conditions. Osteoporos Int 2002;13:768–76.
5. Ahmad H, Taylor BV, van der Mei I, Colman S, O'Leary BA, Breslin M, et al. The impact of multiple sclerosis severity on health state utility values: evidence from Australia. Mult Scler 2017;23:1157–66.

6. Meregaglia M, Cairns J. A systematic literature review of health state utility values in head and neck cancer. Health Qual Life Outcomes 2017;15:174.
7. Hawthorne G, Richardson J, Day NA. A comparison of the Assessment of Quality of Life (AQoL) with four other generic utility instruments. Ann Med 2001;33:358–70.
8. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 2009;339: b2700.
9. Arber M, Glanville J, Isojarvi J, Baragula E, Edwards M, Shaw A, et al. Which databases should be used to identify studies for systematic reviews of economic evaluations? Int J Technol Assess Health Care 2018;34:547–54.
10. Papaioannou D, Brazier J, Paisley S. Systematic searching and selection of health state utility values from the literature. Value Health 2013; 16:686–95.
11. Arber M, Garcia S, Veale T, Edwards M, Shaw A, Glanville JM. Performance of Ovid Medline search filters to identify health state utility studies. Int J Technol Assess Health Care 2017;33:472–80.
12. Covidence. Better systematic review management. URL: https://www.covidence.org/home.
13. The Royal Australian College of General Practitioners. Guideline for the management of knee and hip osteoarthritis. 2nd ed. East Melbourne: Royal Australian College of General Practitioners; 2018.
14. Harris RJ, Deeks JJ, Altman DG, Bradburn MJ, Harbord RM, Sterne JA. Metan: fixed-and random-effects meta-analysis. Stata J 2008;8:3–28.
15. Wan X, Wang W, Liu J, Tong T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. BMC Med Res Methodol 2014;14:135.
16. Luo D, Wan X, Liu J, Tong T. Optimally estimating the sample mean from the sample size, median, mid-range, and/or mid-quartile range. Stat Methods Med Res 2018;27:1785–805.
17. Higgins JP, Green S. Cochrane handbook for systematic reviews of interventions. Boston: John Wiley; 2011.
18. Osteoarthritis Research Society International. Osteoarthritis: a serious disease. URL: https://www.oarsi.org/sites/default/files/docs/2016/oarsi_white_paper_oa_serious_disease_121416_1.pdf.
19. Mahir L, Belhaj K, Zahi S, Azanmasso H, Lmidmani F, El Fatimi A. Impact of knee osteoarthritis on the quality of life. Ann Phys Rehabil Med. 2016;59:e159.
20. Salaffi F, Carotti M, Stancati A, Grassi W. Health-related quality of life in older adults with symptomatic hip and knee osteoarthritis: a comparison with matched healthy controls. Aging Clin Exp Res 2005;17: 255–63.
21. Millett PJ, Gobezie R, Boykin RE. Shoulder osteoarthritis: diagnosis and management. Am Fam Physician 2008;78:605–11.
22. Haugen IK, Englund M, Aliabadi P, Niu J, Clancy M, Kvien TK, et al. Prevalence, incidence and progression of hand osteoarthritis in the general population: the Framingham Osteoarthritis Study. Ann Rheum Dis 2011;70:1581–6.
23. Campbell JA, Palmer AJ, Venn A, Sharman M, Otahal P, Neil A. A head-to-head comparison of the EQ-5D-5L and AQoL-8D multiattribute utility instruments in patients who have previously undergone bariatric surgery. Patient 2016;9:311–22.
24. Campbell JA, Hensher M, Neil A, Venn A, Wilkinson S, Palmer AJ. An exploratory study of long-term publicly waitlisted bariatric surgery patients' quality of life before and 1 year after bariatric surgery, and considerations for healthcare planners. Pharmacoecon Open 2018; 2:63–76.
25. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. J Clin Epidemiol 2003;56: 395–407.
26. Hawthorne G, Osborne R. Population norms and meaningful differences for the Assessment of Quality of Life (AQoL) measure. Aust N Z J Public Health 2005;29:136–42.
27. Walters SJ, Brazier JE. What is the relationship between the minimally important difference and health state utility values? The case of the SF-6D. Health Qual Life Outcomes 2003;1:4.
28. Fransen M, McConnell S, Harmer AR, Van der Esch M, Simic M, Bennell KL. Exercise for osteoarthritis of the knee: a Cochrane systematic review. Br J Sports Med 2015;49:1554–7.
29. Pisters MF, Veenhof C, van Meeteren NL, Ostelo RW, de Bakker DH, Schellevis FG, et al. Long-term effectiveness of exercise therapy in patients with osteoarthritis of the hip or knee: a systematic review. Arthritis Care Res (Hoboken) 2007;57:1245–53.
30. Van Gool CH, Penninx BW, Kempen GI, Rejeski WJ, Miller GD, van Eijk JT, et al. Effects of exercise adherence on physical function among overweight older adults with knee osteoarthritis. Arthritis Care Res (Hoboken) 2005;53:24–32.
31. Marks R, Allegrante JP. Chronic osteoarthritis and adherence to exercise: a review of the literature. J Aging Phys Act 2005;13:434–60.
32. Lin J, Zhang W, Jones A, Doherty M. Efficacy of topical non-steroidal anti-inflammatory drugs in the treatment of osteoarthritis: meta-analysis of randomised controlled trials. BMJ 2004;329:324.
33. Scott DL, Berry H, Capell H, Coppock J, Daymond T, Doyle DV, et al. The long-term effects of non-steroidal anti-inflammatory drugs in osteoarthritis of the knee: a randomized placebo-controlled trial. Rheumatology (Oxford) 2000;39:1095–101.
34. Nilsdotter AK, Toksvig-Larsen S, Roos E. A 5 year prospective study of patient-relevant outcomes after total knee replacement. Osteoarthritis Cartilage 2009;17:601–6.
35. Bachmeier CJ, March LM, Cross MJ, Lapsley HM, Tribe KL, Courtenay BG, et al. A comparison of outcomes in osteoarthritis patients undergoing total hip and knee replacement surgery. Osteoarthritis Cartilage 2001;9:137–46.
36. Murphy SL, Robinson-Lane SG, Niemiec SL. Knee and hip osteoarthritis management: a review of current and emerging nonpharmacological approaches. Curr Treatment Opt Rheumatol 2016; 2:296–311.
37. Lee PY, Rozewicz S, Chandrappa MH, Othman A, Jury C, Whiting B. Modern non-pharmacological and non-surgical treatments for hip pain. J Arthritis 2018;7:262.
38. Myers J, Wielage RC, Han B, Price K, Gahn J, Paget MA, et al. The efficacy of duloxetine, non-steroidal anti-inflammatory drugs, and opioids in osteoarthritis: a systematic literature review and meta-analysis. BMC Musculoskelet Disord 2014;15:76.
39. Ayral X. Injections in the treatment of osteoarthritis. Best Pract Res Clin Rheumatol 2001;15:609–26.
40. Fransen M, Edmonds J. Reliability and validity of the EuroQol in patients with osteoarthritis of the knee. Rheumatology (Oxford) 1999;38:807–13.
41. Zhang F, Yang Y, Huang T, Zhang Y, Zhao L, Li S. Is there a difference between EQ-5D and SF-6D in the clinical setting? A comparative study on the quality of life measured by AIMS2-SF, EQ-5D and SF-6D scales for osteoarthritis patients. Int J Rheum Dis 2018;21:1185–92.
42. NIfHaC E. Guide to the methods of technology appraisal 2013. London: NICE; 2013.
43. Sinusas K. Osteoarthritis: diagnosis and treatment. Am Fam Physician 2012;85:49-56.
44. Borenstein M, Hedges L, Rothstein H. Meta-analysis: Fixed effect vs. random effects. 2007. URL: http://www.meta-analysis.com/downloads/Meta-analysis_fixed_effect_vs_random_effects_sv.pdf.

45. Abbott J, Wilson R, Pinto D, Chapple C, Wright A, Team MT. Incremental clinical effectiveness and cost effectiveness of providing supervised physiotherapy in addition to usual medical care in patients with osteoarthritis of the hip or knee: 2-year results of the MOA randomised controlled trial. Osteoarthritis Cartilage 2019; 27:424–34.
46. Bulthuis Y, Mohammad S, Braakman-Jansen LM, Drossaers-Bakker KW, van de Laar MA. Cost-effectiveness of intensive exercise therapy directly following hospital discharge in patients with arthritis: results of a randomized controlled clinical trial. Arthritis Care Res (Hoboken) 2008;59:247–54.
47. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated health economic evaluation reporting standards (CHEERS) statement. BMJ 2013;346:f1049.
48. Ramsey SD, Willke RJ, Glick H, Reed SD, Augustovski F, Jonsson B, et al. Cost-effectiveness analysis alongside clinical trials II: an ISPOR Good Research Practices Task Force report. Value Health 2015;18:161–72.

Arthritis Care & Research
Vol. 74, No. 2, February 2022, pp 301–307
DOI 10.1002/acr.24452

AMERICAN COLLEGE of RHEUMATOLOGY
Empowering Rheumatology Professionals

BRIEF REPORT

Walking Disabilities in Association With Tenosynovitis at the Metatarsophalangeal Joints: A Longitudinal Magnetic Resonance Imaging Study in Early Arthritis

Yousra J. Dakkak, Fenne Wouters, Xanthe M. E. Matthijssen, Monique Reijnierse, and Annette H. M. van der Helm-van Mil

Objective. The relationship between functional disability and magnetic resonance imaging (MRI) inflammation has been studied for the hands, but has not been well established for the feet, even though walking difficulties are common. Therefore, our objective was to study whether walking difficulties were associated with MRI inflammation at metatarsophalangeal (MTP) joints in early arthritis patients, at diagnosis and during 24 months of follow-up.

Methods. A total of 532 consecutive patients presenting with early arthritis reported on the presence and severity of walking difficulties (Health Assessment Questionnaire question 4a, scale 0–3), and underwent unilateral contrast-enhanced MRI of MTP joints 1–5 at baseline. In total, 107 patients had clinical and MRI data at follow-up (4, 12, and 24 months). MRI inflammation (synovitis, tenosynovitis, and osteitis) was scored in line with the Rheumatoid Arthritis Magnetic Resonance Imaging Scoring system. At baseline, the association of walking disability with MRI inflammation was assessed using regression. Longitudinally, the association between a change in walking disability with a change in MRI inflammation was studied with linear mixed models.

Results. At baseline, 81% of patients with walking disabilities had MRI inflammation at MTP joints, versus 68% without walking disabilities ($P < 0.001$). Total MRI inflammation (i.e., the sum of tenosynovitis, synovitis, and osteitis) was associated with severity of walking disability ($\beta = 0.023$, $P < 0.001$). Studying the MRI features separately, tenosynovitis, synovitis, and osteitis were all univariably associated with severity of walking disability ($P < 0.001$, $P < 0.001$, and $P = 0.014$, respectively). In multivariable analysis, the association was strongest for tenosynovitis. During follow-up, a decrease in MTP inflammation was associated with a decrease in walking disability ($\beta = 0.029$, $P = 0.001$); in multivariable analyses only, tenosynovitis was independently associated ($\beta = 0.073$, $P = 0.049$).

Conclusion. Of the different inflamed tissues in MTP joints, predominantly MRI-detected tenosynovitis was associated with walking disabilities. Likewise a reduction in tenosynovitis related to a decrease in walking disabilities. These results increase our understanding of the involvement of tenosynovitis in walking disabilities in early arthritis.

INTRODUCTION

Rheumatoid arthritis (RA) is a disease that involves the small joints of the hands and feet. The focus in research, however, has primarily been on the hands (1), though 80% of patients report disease-related foot problems and 71% report walking difficulties (2). These difficulties have an important impact on the quality of life of patients that is often underestimated by clinicians (3) and is associated with clinical factors such as inflammation, pain, and duration of disease (2,3).

Magnetic resonance imaging (MRI) is increasingly used in RA research, as it sensitively detects inflammation, defined as tenosynovitis, synovitis, and osteitis. The association between walking difficulties and MRI inflammation has not been fully explored for the forefoot. Two previous reports have included MRI data of the metatarsophalangeal (MTP) joints, but these were not

Supported by the European Research Council under the European Union's Horizon 2020 research and innovation program (Starting grant #714312) and by the Dutch Arthritis Foundation.

[1]Yousra J. Dakkak, MD, Fenne Wouters, MsC, Xanthe M. E. Matthijssen, MD, Monique Reijnierse, MD, PhD, Annette H. M. van der Helm-van Mil, MD, PhD: Leiden University Medical Center, Leiden, The Netherlands.

No potential conflicts of interest relevant to this article were reported.

Address correspondence to Yousra J. Dakkak, MD, Department of Rheumatology, Leiden University Medical Center, P. O. Box 9600, 2300 RC Leiden, The Netherlands. Email: y.j.dakkak@lumc.nl.

Submitted for publication May 12, 2020; accepted in revised form September 10, 2020.

SIGNIFICANCE & INNOVATIONS

- Walking disability at diagnosis is associated with magnetic resonance imaging (MRI) inflammation at metatarsophalangeal joints; this association was strongest for tenosynovitis.
- A treatment-induced decrease of MRI inflammation, particularly tenosynovitis, is associated with a reduction in walking disabilities.
- This study increases our understanding of the nature of walking impairments in early arthritis.

contrast-enhanced, and tenosynovitis at the MTP joints was not included (4,5). Additionally, the reports were cross-sectional and thus did not study whether change in MRI inflammation over time related to change in disability.

Therefore, with the aim to increase our understanding of the role of inflammation at the MTP joints in functional disability, we set up a cross-sectional and longitudinal study in early arthritis patients to evaluate the association of walking disabilities with MRI inflammation, defined as tenosynovitis, synovitis, and osteitis at the MTP joints.

PATIENTS AND METHODS

Patients. Between June 2013 and July 2017, 604 consecutive patients newly presenting with clinical confirmed arthritis of ≥1 joint and a symptom duration of <2 years who were naive to disease-modifying antirheumatic drugs (DMARDs) were included in the Leiden Early Arthritis Cohort. The cohort is extensively described elsewhere (6). In short, at baseline, 4 months, 12 months, and yearly thereafter, information was obtained from physical examination, laboratory tests, questionnaires including the Health Assessment Questionnaire (HAQ), and MRI. Included patients were treated in routine care.

Of 604 patients at baseline, 65 had missing HAQ data and 7 had insufficient MRI images. The remaining 532 were studied. A flow chart is shown in Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24452/abstract. From August 2010 until February 2015, follow-up MRIs were performed in patients with the initial diagnosis of RA or undifferentiated arthritis. From the 532 with complete baseline MRI and HAQ data, follow-up MRI results were available for 107 patients (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24452/abstract). The Early Arthritis Cohort was approved by the local medical ethics committee (#P10.108). Informed consent was obtained. The data sets analyzed during the current study are available from the corresponding author on reasonable request.

Assessment of walking disability. The HAQ is a well-validated, widely used questionnaire on functional disability that consists of 20 questions covering different categories of functional activities (4), including 2 questions on walking: question 4a: "are you able to walk outdoors on flat ground?" and question 4b: "are you able to climb up 5 steps?" They are answered as 0 = no difficulty, 1 = some difficulty, 2 = much difficulty, and 3 = unable to do. Question 4a was used as the primary measure of walking disability, because walking outdoors on flat ground was assumed to importantly involve forefoot mechanics. Question 4b was used as an alternative measure in a subanalysis, as climbing stairs was assumed to assess not only forefoot mechanics, but also other joints such as ankle and knee mechanics. Data on walking disability were available at baseline and at 12 and 24 months. The HAQ also evaluates equipment dependency, like walking sticks. Equipment dependency can be the result of disability in different domains, such as the knee or hip. To avoid the introduction of noise, equipment dependency was therefore not incorporated in the analyses, although excluding this dependency from the evaluation can potentially lead to an underestimation of the severity of walking disability.

MRI. Unilateral contrast-enhanced MRI of MTP joints 1–5 of the more painful side, or the dominant side in case of symmetric symptoms, was performed with a 1.5T extremity MRI (General Electric). Baseline MRI was obtained ≤2 weeks after the first presentation and before DMARD initiation, with follow-up MRIs at 4, 12, and 24 months. MRIs were scored for tenosynovitis, synovitis, osteitis, and erosions at MTP joints 1–5, in line with the Rheumatoid Arthritis Magnetic Resonance Imaging Scoring system, with researchers blinded from any clinical data. A detailed description is given in Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24452/abstract (7–9). Total MRI inflammation was defined as the total sum of the semiquantitative scores of tenosynovitis (range 0–30), synovitis (range 0–15), and osteitis (range 0–30) at MTP joints. Follow-up MRI was scored in known time order. Reliability of scoring was excellent (intraclass correlation coefficient ≥0.92). Additional information is given in Supplementary Appendix A and Supplementary Figure 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24452/abstract).

Statistical analysis. At baseline the association between walking difficulties and total MRI inflammation was assessed using linear regression, with severity of walking disability as the outcome. Although erosions were expected to be infrequent at the time of diagnosis, they were also studied in relation to walking disability. Next, the association of tenosynovitis, synovitis, and osteitis was assessed separately. Univariable and multivariable analyses were performed: multivariable analyses adjusted for the simultaneous presence of different types of MRI inflammation,

because these 3 features often co-occur, and in a separate analysis for the following clinical features: age, 66 swollen joint count (SJC), and C-reactive protein (CRP) level. We adjusted for these factors because they may associate with walking disability and MRI inflammation, and to elucidate whether MRI inflammation is associated with walking disability regardless of the level of systemic and local inflammation (CRP level and SJC, respectively) (6,10). The analyses were repeated for the presence of walking disability as a dichotomous outcome using logistic regression.

To assess whether a change in the severity of walking disability was associated with a change in MRI inflammation, linear mixed models were used. First, the association was studied for total MRI inflammation with walking disability as the outcome. Subsequently, tenosynovitis, synovitis, and osteitis were assessed separately. Also here analyses were performed univariably and multivariably, adjusting for the simultaneous presence of different types of MRI inflammation and for clinical features. Linear mixed models have the advantage that all patient information is used, including for those who had missing data, as this method assumes that missing outcomes can be estimated using available measurements.

RA patients may have more severe inflammation, potentially influencing the relationship between walking difficulties and MRI inflammation. Therefore, as a subanalysis, the analyses between walking difficulties and MRI inflammation at baseline and during follow-up were repeated in the subgroup of RA patients (clinical diagnosis and fulfilment of 1987 or 2010 criteria at 2 weeks).

We prioritized walking difficulties in the main analyses. As a subanalysis we analyzed whether MRI inflammation at MTP joints could possibly also be related to difficulty climbing stairs. Therefore, analyses were repeated with difficulty climbing stairs as the outcome.

RESULTS

Patient characteristics. Of 532 patients with baseline data, the mean age was 58 years, 60% were female, and the mean symptom duration was 10 weeks. Walking disability was present in 202 patients (38%). This finding was comparable in the subgroup of patients who were studied longitudinally. Patient characteristics are shown in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24452/abstract.

Walking disability and MRI-detected inflammation at baseline. Mean MRI scores and the results from regression analyses are shown in Tables 1 and 2. At baseline, more severe walking disabilities were associated with more severe total MRI inflammation ($\beta = 0.023$, $P < 0.001$). The severity of walking disability was associated with tenosynovitis, synovitis, and osteitis scores in univariable analyses ($P < 0.001$, $P < 0.001$, and $P = 0.014$, respectively). In a multivariable analysis that included all 3 features, the effect size was largest for tenosynovitis ($\beta = 0.042$, $P = 0.060$). In a separate multivariable analysis, the results for total MRI inflammation and tenosynovitis were adjusted for clinical features (age, SJC, and CRP level); MRI inflammation and tenosynovitis remained associated with walking disability ($P = 0.014$ and $P = 0.042$, respectively). Walking disability was not associated with erosion scores ($P = 0.18$). In additional multivariable analyses, the results for total MRI inflammation and tenosynovitis were adjusted for clinical features and MRI-detected erosions that revealed similar results ($\beta = 0.014$, $P = 0.026$ for MRI inflammation and $\beta = 0.035$, $P = 0.047$ for tenosynovitis, results not shown in tables). Next the association of the

Table 1. Severity of walking disability and the association between MRI-detected inflammation at the MTP joints and walking-disability at disease presentation in 532 early arthritis patients*

	Total inflammation score†	Tenosynovitis score	Synovitis score	Osteitis score	Erosion score
MRI score, mean ± SD‡					
Disability positive	4.6 ± 6.3	1.3 ± 2.2	1.6 ± 2.1	1.7 ± 3.2	0.7 ± 1.1
Disability negative	2.7 ± 4.1	0.7 ± 1.4	1.0 ± 1.5	1.1 ± 2.1	0.6 ± 0.9
Univariable analysis					
β (95% CI)	0.023 (0.01, 0.03)	0.064 (0.03, 0.1)	0.063 (0.03, 0.1)	0.029 (0.02, 0.05)	0.043 (–0.02, 0.1)
P	<0.001	<0.001	<0.001	0.014	0.18
Multivariable analysis					
MRI features§					
β (95% CI)	–	0.042 (–0.02, 0.09)	0.026 (–0.02, 0.08)	0.007 (–0.02, 0.04)	–
P	–	0.06	0.27	0.66	–
Clinical features¶					
β (95% CI)	0.015 (0.00, 0.03)	0.036 (0.00, 0.07)	–	–	–
P	0.014	0.042	–	–	–

* Assessed using linear regression. 95% CI = 95% confidence interval; MRI = magnetic resonance imaging; MTP = metatarsophalangeal.
† Defined as the summed scores of tenosynovitis, synovitis, and osteitis.
‡ Mean score of MRI features in patients with walking disability (defined as Health Assessment Questionnaire [HAQ] question 4a = ≥1) and patients without walking disability (HAQ question 4a = 0).
§ Multivariable analyses including MRI-detected tenosynovitis, synovitis, and osteitis at the MTP joints.
¶ Multivariable analyses including swollen joint count, age at inclusion, and C-reactive protein level, performed separately for the total inflammation score and for tenosynovitis. Due to the risk of overfitting, this multivariable analysis only included variables that were most importantly associated with walking disability.

Table 2. Presence of walking disability: association between MRI-detected inflammation at the MTP joints and walking disability at disease presentation in 532 early arthritis patients*

	Presence of any inflammation†	Tenosynovitis score	Synovitis score	Osteitis score	Erosion score
MRI feature present, no. (%)‡					
Disability positive	163 (81)	91 (45)	128 (64)	117 (60)	93 (46)
Disability negative	223 (68)	105 (32)	161 (49)	165 (50)	147 (45)
Univariable analysis					
OR (95% CI)	1.08 (1.0, 1.1)	1.22 (1.1, 1.4)	1.21 (1.1, 1.3)	1.10 (1.0, 1.2)	1.12 (0.9, 1.3)
P	<0.001	<0.001	<0.001	0.013	0.23
Multivariable analysis					
MRI features§					
OR (95% CI)	–	1.15 (1.003, 1.31)	1.09 (0.93, 1.27)	1.03 (0.94, 1.12)	–
P	–	0.045	0.29	0.57	–
Clinical features¶					
OR (95% CI)	1.06 (1.02, 1.10)	1.15 (1.03, 1.28)	–	–	–
P	0.005	0.017	–	–	–

* Dichotomous outcome, association assessed using logistic regression. 95% CI = 95% confidence interval; MRI = magnetic resonance imaging; MTP = metatarsophalangeal; OR = odds ratio.
† Defined as the presence of tenosynovitis, synovitis, and/or osteitis.
‡ Presence of an MRI feature in patients with walking disability (defined as Health Assessment Questionnaire [HAQ] question 4a = ≥1) and patients without walking disability (HAQ question 4a = 0).
§ Multivariable analyses including MRI-detected tenosynovitis, synovitis, and osteitis at the MTP joints.
¶ Multivariable analyses including swollen joint count, age at inclusion, and C-reactive protein level, performed separately for the total inflammation score and for tenosynovitis. Due to the risk of overfitting, this multivariable analysis only included variables that were most importantly associated with walking disability.

3 inflammatory features was studied with the presence of walking disability as a dichotomous outcome (Table 2). This analysis revealed similar results (Tables 1 and 2).

Course of walking disability and MRI inflammation during 2 years of follow-up. Then we assessed whether a change in the severity of walking disability was associated with a change in MRI inflammation during 2 years of follow-up (Figure 1 and Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24452/abstract). A decrease in total MRI inflammation was associated with a decrease in walking disability (β = 0.022, *P* = 0.019). For the separate features, a decrease in tenosynovitis and synovitis was associated with a decrease

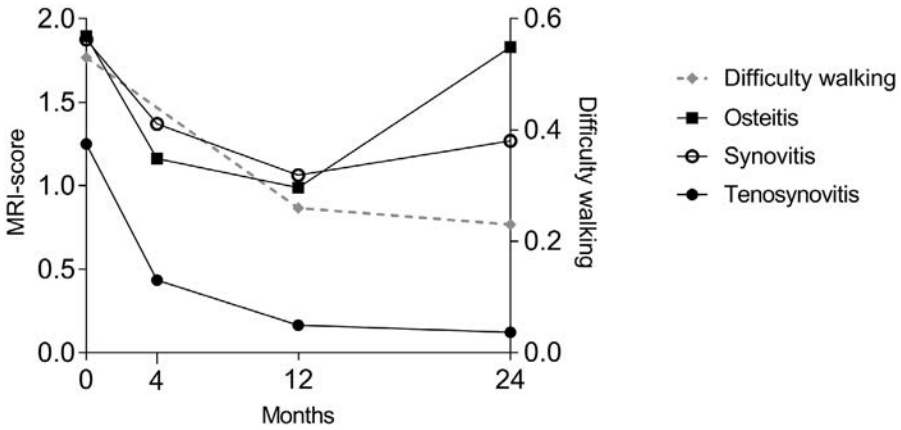

Figure 1. Difficulty walking and magnetic resonance imaging (MRI) mean scores for osteitis, synovitis, and tenosynovitis during 24 months of follow-up. Difficulty walking was assessed by the Health Assessment Questionnaire question 4a: "are you able to walk outdoors on flat ground?" Patients answered on a scale from 0 to 3 (0 = without difficulty, 1 = with some difficulty, 2 = with much difficulty, 3 = unable to do). In univariable analyses, a decrease in tenosynovitis and synovitis was associated with a decrease in walking disability (*P* = 0.001 and *P* = 0.002, respectively; see Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24452/abstract). In multivariable analyses that included osteitis, synovitis, and tenosynovitis, only the association for tenosynovitis remained (β = 0.073, *P* = 0.049). Follow-up MRI data were available as follows: 107, 100, 80, and 41 patients at baseline and at 4, 12, and 24 months, respectively. Data on walking difficulty were available for 107, 78, and 70 patients at baseline and at 12 and 24 months, respectively. The increase in osteitis score at 24 months was caused by missing data for patients with resolution of symptoms who were lost to follow-up at 24 months, while patients with more severe disease kept coming for follow-up (see Supplementary Figure 4, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24452/abstract).

in walking disability ($P = 0.001$ and $P = 0.002$, respectively, in univariable analyses), while osteitis was not statistically significant ($P = 0.058$). The association for tenosynovitis remained in multivariable analysis adjusted for synovitis and osteitis ($\beta = 0.073$, $P = 0.049$) and when adjusted for clinical features ($\beta = 0.091$, $P = 0.002$). The analyses were repeated and adjusted for baseline values, which revealed similar results for MRI inflammation ($\beta = 0.024$, $P = 0.014$) and for tenosynovitis ($\beta = 0.069$, $P = 0.034$) (Figure 2).

Subanalysis in RA patients. As a subanalysis, the association between walking difficulties and MRI inflammation at baseline was assessed in the subgroup of RA patients (n = 192), as RA patients may have more severe inflammation that may influence the relationship between walking difficulties and MRI inflammation (see Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24452/abstract). Indeed, walking difficulties were more frequently present (46% versus 38% of patients; see

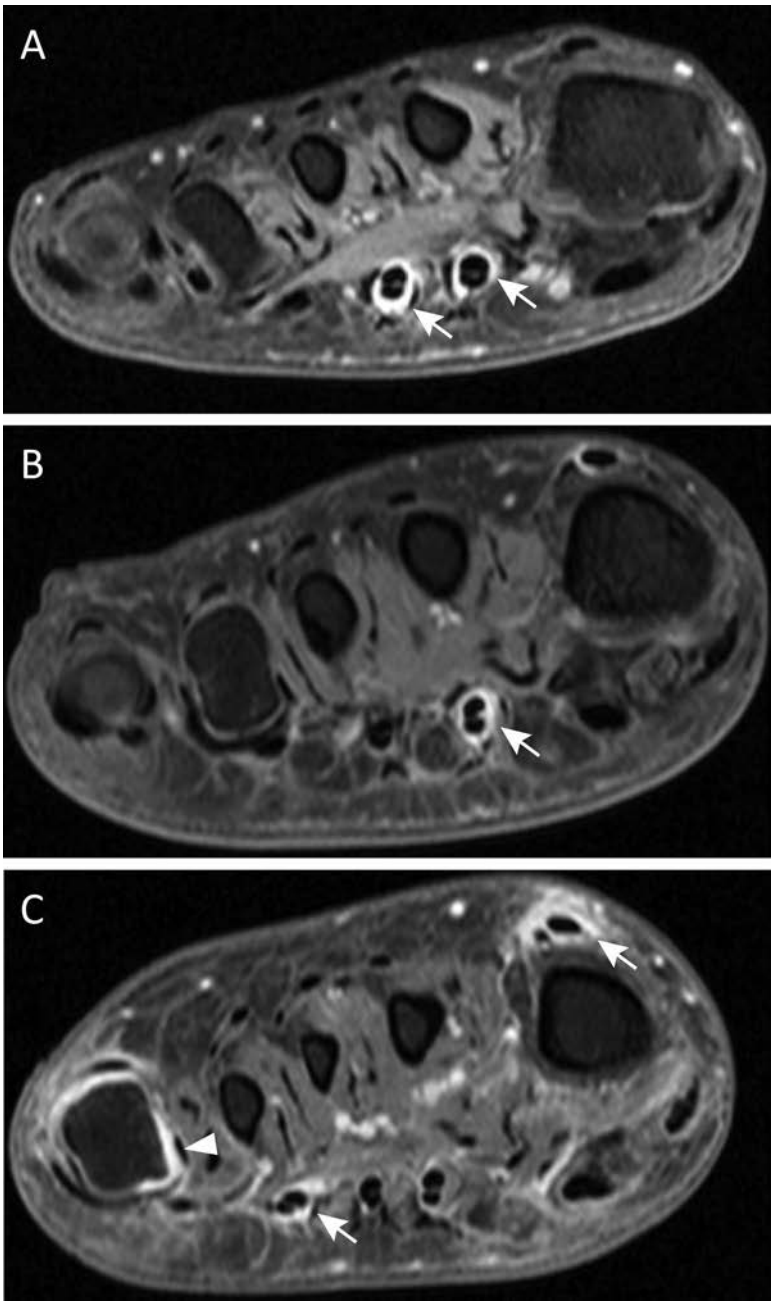

Figure 2. Examples of magnetic resonance imaging–detected tenosynovitis (arrows) in the coronal plane. **A**, Tenosynovitis of the common flexor digitorum at the 2nd and 3rd metatarsophalangeal (MTP) joint; **B**, Tenosynovitis of the common flexor digitorum at the 2nd MTP joint; and **C**, Tenosynovitis of the extensor hallucis longus and common flexor digitorum of the 4th MTP joint, with synovitis of MTP-5 (arrowhead).

Supplementary Table 1, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24452/abstract), and the total MRI inflammation score was higher in RA patients than in early arthritis patients with walking disabilities (6.4 versus 4.6; see Supplementary Table 3, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24452/abstract). Similar results were observed as in the main analyses; in multivariable analysis, the association only remained for tenosynovitis ($\beta = 0.078$, $P = 0.004$). The effect of tenosynovitis remained when adjusted for clinical features ($\beta = 0.054$, $P = 0.015$).

Next we assessed whether a decrease in walking difficulties was associated with a decrease in MRI inflammation in the subgroup of RA patients (n = 72) (see Supplementary Figure 3 and Supplementary Table 4, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24452/abstract). This analysis revealed similar results, and in multivariable analysis the association only remained for tenosynovitis ($\beta = 0.097$, $P = 0.021$). The effect of tenosynovitis remained when adjusted for clinical features ($\beta = 0.12$, $P < 0.001$).

Subanalysis for difficulty climbing stairs and MRI inflammation. Finally, the analyses at baseline were repeated with difficulty climbing stairs (HAQ question 4b) as the outcome (see Supplementary Table 5, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24452/abstract). Also here, the severity of walking disability was associated with the total MRI inflammation score ($\beta = 0.023$, $P < 0.001$). In multivariable analysis with all 3 inflammatory features, the effect was strongest for tenosynovitis ($\beta = 0.045$, $P = 0.045$).

DISCUSSION

Walking disabilities were frequent, and in our study 38% of early arthritis patients and 46% of RA patients reported having difficulties with walking. We aimed to increase our understanding of the role of inflammation at the MTP joints in this disability and observed that the severity of inflammation at MTP joints as detected with MRI was associated with walking disabilities at diagnosis. Interestingly, although univariable analyses revealed that synovitis, tenosynovitis, and osteitis were associated with walking difficulties, and that synovitis and tenosynovitis often co-occur (risk of collinearity in multivariable analyses), in multivariable analyses including the different inflammatory lesions, tenosynovitis had the strongest association. This finding suggests that tenosynovitis not only causes walking difficulties because it co-occurs with synovitis, but that by itself it can also lead to walking difficulties in patients with early arthritis. The association of tenosynovitis with walking disability was also independent of regular measures of local and systemic inflammation (SJC and CRP level). Finally, serial MRIs revealed that a decrease of MRI inflammation was associated with a reduction in walking disabilities and that here as well the association was strongest for tenosynovitis. These results suggest that tenosynovitis at the level of MTP joints importantly contributes to physical impairments.

These results add to the increasing evidence on the importance of tenosynovitis in early RA (11). Most of this research, however, has focused on the hands (1,4), also regarding disability. In the hands, tenosynovitis also had the strongest association. A recent study showed that of the 3 inflammatory MRI features at the MTP joints, tenosynovitis had the strongest association with early RA (12). In that report, tenosynovitis at both flexor and extensor tendons was associated with RA and occurred in 31% and 28% of RA patients, respectively, of which the most common site was extensor tenosynovitis of MTP-1, which occurred in 20% of RA patients. Our current study is the first to report on MTP tenosynovitis with respect to functional disability.

To the best of our knowledge, no longitudinal studies on walking disability in relation to imaging-detected inflammation exist in early disease. Previous studies have reported on the occurrence of walking disabilities in established RA, where walking disability remained moderate to severe during follow-up (13). We have found a decrease in walking disability from the moment of diagnosis until 2 years of follow-up. This decrease after diagnosis is most likely the result of treatment initiation. Nevertheless, after having observed that the severity of walking disabilities is associated with the severity of tenosynovitis, we see confirmation in the fact that improvement in walking is associated with a reduction in tenosynovitis.

A limitation of this study is that the HAQ is validated for integral use and not for the individual questions. Validated questionnaires that specifically study foot-related disability exist, like the Leeds Foot Impairment Score that more specifically studies impairment and activity limitation (14). The relationship of imaging-detected MTP inflammation in these different aspects of foot disability would certainly be interesting and is a subject for further research.

MTP-1 is a predilection site for degenerative disease, and part of the inflammation (synovitis, osteitis) at advanced age in MTP-1 is possibly related to osteoarthritis. However, research has also reported that tenosynovitis at MTP-1 is RA specific, and that involvement of MTP-1 is especially specific for RA in younger patients (12,15). We therefore did not exclude MTP-1 from our analyses. Although radiographic information on osteoarthritis of MTP-1 was not systematically available, we did adjust for age in our analyses, as osteoarthritis is mostly age related.

Walking disabilities in established RA can be due to inflammation but also due to damage and joint deformity. We studied patients at first presentation to the outpatient clinic and found no relation between erosions and walking disabilities. This finding is not surprising, as the prevalence of erosions, and thus the contribution to functional impairment, is low at disease onset.

Interestingly, in addition to the studied MRI features, in the forefoot, synovium-lined intermetatarsal bursae are present that can clearly be differentiated from tenosynovitis, synovitis, and osteitis, as they have no anatomical connection with MTP joints and are surrounded bilaterally by the interosseous tendons (16). Inflammation of these bursae, referred to as intermetatarsal bursitis, relates to clinical joint swelling (17) and may predict the development of foot impairment in RA patients (13). How inflammation relates to tenosynovitis, synovitis, and osteitis in respect to disability is unknown.

In conclusion, in patients with arthritis, we traditionally consider synovitis to be the cause of disability. We have shown that, additionally, tenosynovitis at the MTP joints is an important feature that needs to be considered. Appreciating this role of tenosynovitis increases our understanding of walking disabilities in patients with early arthritis and of functional disability in RA patients.

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. Dakkak 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. Dakkak, Reijnierse, van der Helm-van Mil.

Acquisition of data. Dakkak, Wouters.

Analysis and interpretation of data. Dakkak, Matthijssen, van der Helm-van Mil.

REFERENCES

1. Dakkak YJ, van der Heijde DM, Reijnierse M, van der Helm-van Mil AH. Validity of the rheumatoid arthritis MRI score applied to the forefeet using the OMERACT filter: a systematic literature review. RMD Open 2018;4:e000796.
2. Grondal L, Tengstrand B, Nordmark B, Wretenberg P, Stark A. The foot: still the most important reason for walking incapacity in rheumatoid arthritis. Distribution of symptomatic joints in 1,000 RA patients. Acta Orthop 2008;79:257–61.
3. Otter SJ, Lucas K, Springett K, Moore A, Davies K, Cheek L, et al. Foot pain in rheumatoid arthritis prevalence, risk factors and management: an epidemiological study. Clin Rheumatol 2010;29:255–71.
4. Burgers LE, Nieuwenhuis WP, van Steenbergen HW, Newsum EC, Huizinga TW, Reijnierse M, et al. Magnetic resonance imaging-detected inflammation is associated with functional disability in early arthritis: results of a cross-sectional study. Rheumatology (Oxford) 2016;55:2167–75.
5. Rondina RG, de Mello RA, Valim V, Lourenco RB, Batista EF, de Oliveira Junior R. Discordance between clinical and imaging criteria: assessment by magnetic resonance imaging of the foot of patients with rheumatoid arthritis. Rheumatol Int 2017;37:1357–64.
6. Nieuwenhuis WP, Mangnus L, van Steenbergen HW, Newsum EC, Huizinga TW, Reijnierse M, et al. Older age is associated with more MRI-detected inflammation in hand and foot joints. Rheumatology (Oxford) 2016;55:2212–9.
7. Haavardsholm EA, Ostergaard M, Ejbjerg BJ, Kvan NP, Kvien TK. Introduction of a novel magnetic resonance imaging tenosynovitis score for rheumatoid arthritis: reliability in a multireader longitudinal study. Ann Rheum Dis 2007;66:1216–20.
8. Ostergaard M, Peterfy C, Conaghan P, McQueen F, Bird P, Ejbjerg B, et al. OMERACT rheumatoid arthritis magnetic resonance imaging studies core set of MRI acquisitions, joint pathology definitions, and the OMERACT RA-MRI scoring system. J Rheumatol 2003;30:1385–6.
9. Ostergaard M, Peterfy CG, Bird P, Gandjbakhch F, Glinatsi D, Eshed I, et al. The OMERACT rheumatoid arthritis magnetic resonance imaging (MRI) scoring system: updated recommendations by the OMERACT MRI in Arthritis working group. J Rheumatol 2017;44:1706–12.
10. Turner DE, Helliwell PS, Siegel KL, Woodburn J. Biomechanics of the foot in rheumatoid arthritis: identifying abnormal function and the factors associated with localised disease 'impact'. Clin Biomech (Bristol, Avon) 2008;23:93–100.
11. Rogier C, Hayer S, van der Helm-van Mil A. Not only synovitis but also tenosynovitis needs to be considered: why it is time to update textbook images of rheumatoid arthritis. Ann Rheum Dis 2020;79:546–7.
12. Dakkak YJ, Jansen FP, DeRuiter MC, Reijnierse M, van der Helm-van Mil AH. Rheumatoid arthritis and tenosynovitis at the metatarsophalangeal joints: an anatomic and MRI study of the forefoot tendon sheaths. Radiology 2020;295:146–54.
13. Hooper L, Bowen CJ, Gates L, Culliford DJ, Ball C, Edwards CJ, et al. Prognostic indicators of foot-related disability in patients with rheumatoid arthritis: results of a prospective three-year study. Arthritis Care Res (Hoboken) 2012;64:1116–24.
14. Helliwell P, Reay N, Gilworth G, Redmond A, Slade A, Tennant A, et al. Development of a foot impact scale for rheumatoid arthritis. Arthritis Rheum 2005;53:418–22.
15. Boeters DM, Nieuwenhuis WP, van Steenbergen HW, Reijnierse M, Landewe RB, van der Helm-van Mil AH. Are MRI-detected erosions specific for RA? A large explorative cross-sectional study. Ann Rheum Dis 2018;77:861–8.
16. Theumann NH, Pfirrmann CW, Chung CB, Mohana-Borges AV, Haghighi P, Trudell DJ, et al. Intermetatarsal spaces: analysis with MR bursography, anatomic correlation, and histopathology in cadavers. Radiology 2001;221:478–84.
17. Dakkak YJ, Boer AC, Boeters DM, Niemantsverdriet E, Reijnierse M, van der Helm-van Mil AH. The relation between physical joint examination and MRI-depicted inflammation of metatarsophalangeal joints in early arthritis. Arthritis Res Ther 2020;22:67.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 2, February 2022, pp 308–316
DOI 10.1002/acr.24577

Discrete Choice Experiment on a Magnetic Resonance Imaging Scoring System for Temporomandibular Joints in Juvenile Idiopathic Arthritis

Mirkamal Tolend,[1] Thitiporn Junhasavasdikul,[2] Randy Q. Cron,[3] Emilio J. Inarejos Clemente,[4] Thekla von Kalle,[5] Christian J. Kellenberger,[6] Bernd Koos,[7] Elka Miller,[8] Marion A. van Rossum,[9] Rotraud K. Saurenmann,[6] Lynn Spiegel,[10] Jennifer Stimec,[10] Marinka Twilt,[11] Nikolay Tzaribachev,[12] Shelly Abramowicz,[13] Simone Appenzeller,[14] Linda Z. Arvidsson,[15] Saurabh Guleria,[16] Jacob L. Jaremko,[17] Eva Kirkhus,[18] Tore A. Larheim,[15] Arthur B. Meyers,[19] Jyoti Panwar,[20] Cory M. Resnick,[21] Susan C. Shelmerdine,[22] Brian M. Feldman,[10] and Andrea S. Doria,[10] for the OMERACT MRI in JIA Working Group

Objective. To determine the relative importance weights of items and grades of a newly developed additive outcome measure called the juvenile idiopathic arthritis (JIA) magnetic resonance imaging (MRI) scoring system for the temporomandibular joint (TMJ) (JAMRIS-TMJ).

Methods. An adaptive partial-profile, discrete choice experiment (DCE) survey using the 1000Minds platform was independently completed by members of an expert group consisting of radiologists and non-radiologist clinicians to determine the group-averaged relative weights for the JAMRIS-TMJ. Subsequently, an image-based vignette ranking exercise was done, during which experts individually rank ordered 14 patient vignettes for disease severity while blinded to the weights and unrestricted to JAMRIS-TMJ assessment criteria. Validity of the weighted JAMRIS-TMJ was tested by comparing the consensus-graded, DCE-weighted JAMRIS-TMJ score of the vignettes with their unrestricted image-based ranks provided by the experts.

Results. Nineteen experts completed the DCE survey, and 21 completed the vignette ranking exercise. Synovial thickening and joint enhancement showed higher weights per raw score compared to bone marrow items and effusion in the inflammatory domain, while erosions and condylar flattening showed nonlinear and higher weights compared to disk abnormalities in the damage domain. The weighted JAMRIS-TMJ score of the vignettes correlated highly with the ranks from the unrestricted comparison method, with median Spearman's ρ of 0.92 (interquartile range [IQR] 0.87–0.95) for the inflammation and 0.93 (IQR 0.90–0.94) for the damage domain.

Conclusion. A DCE survey was used to quantify the importance weights of the items and grades of the JAMRIS-TMJ. The weighted score showed high convergent validity with an unrestricted, holistic vignette ranking method.

INTRODUCTION

Juvenile idiopathic arthritis (JIA) is the most common form of chronic arthritis in children and youth, with a prevalence of 1 in 1,000 children worldwide (1). In large consecutive series of JIA patients, ~40% have been found to develop some degree of inflammation and structural changes in the temporomandibular joint (TMJ) (2–4). While arthritis of the TMJ can be asymptomatic (5), it was recently reported that orofacial pain and functional disability are common and seem to persist over time in most

[1]Mirkamal Tolend, BSc: The Hospital for Sick Children and University of Toronto, Toronto, Ontario, Canada; [2]Thitiporn Junhasavasdikul, MD: Ramathibodi Hospital, Bangkok, Thailand; [3]Randy Q. Cron, MD, PhD: Children's of Alabama, Birmingham; [4]Emilio J. Inarejos Clemente, MD: Hospital Sant Joan de Deu, Barcelona, Spain; [5]Thekla von Kalle, MD: Olgahospital Klinikum Stuttgart, Stuttgart, Germany; [6]Christian J. Kellenberger, MD, Rotraud K. Saurenmann, MD: University Children's Hospital Zürich, Zürich, Switzerland; [7]Bernd Koos, DMD: University Hospital Tübingen, Tübingen, Germany; [8]Elka Miller, MD, CHEO: Children's Hospital of Eastern Ontario, Ottawa, Ontario, Canada; [9]Marion A. van Rossum, MD, PhD: Emma Children's Hospital, Academic Medical Centre and Amsterdam Rheumatology and Immunology Center, Reade, Amsterdam, The Netherlands; [10]Lynn Spiegel, MD, Jennifer Stimec, MD, Brian M. Feldman, MD, MSc, Andrea S. Doria, MD, PhD, MSc, MBA: The Hospital for Sick Children, Toronto, Ontario, Canada; [11]Marinka Twilt, MD, MScE, PhD: Alberta Children's Hospital and University of Calgary, Calgary, Alberta, Canada; [12]Nikolay Tzaribachev, MD, PhD: Pediatric Rheumatology Research Institute, Bad Bramstedt, Germany; [13]Shelly Abramowicz, DMD, MPH: Emory University School of Medicine and Children's Healthcare of Atlanta, Atlanta, Georgia; [14]Simone Appenzeller, MD, PhD: University of Campinas, Campinas, Brazil; [15]Linda Z. Arvidsson, DDS, PhD, Tore A. Larheim, DDS, PhD: University of Oslo, Oslo, Norway; [16]Saurabh Guleria, MD: Austin Radiological Association, Austin, Texas; [17]Jacob L. Jaremko, MD, PhD: University of Alberta, Edmonton, Alberta, Canada; [18]Eva Kirkhus, MD, PhD: Oslo University Hospital, Oslo, Norway; [19]Arthur B. Meyers, MD: Cincinnati Children's Hospital

SIGNIFICANCE & INNOVATIONS

- A discrete choice experiment was used to develop a weighting scheme for the items and grades of a newly developed magnetic resonance imaging scoring system for assessing the inflammation and damage in the temporomandibular joints (TMJ) of children with juvenile idiopathic arthritis (JAMRIS-TMJ).
- In the inflammatory domain of the scoring system, the importance weights for joint enhancement (34% of domain score) and synovial thickening (31%) were higher than the bone marrow items (9% and 10%) and effusion (16%).
- In the damage domain, erosions and condylar flattening were both weighted higher compared to disk abnormalities (38% and 49% versus 13%).
- The weighted JAMRIS-TMJ score showed high convergent validity when compared to an unrestricted image-based method of ranking vignettes (median Spearman's ρ of 0.92 and 0.93 for the 2 domains).

patients, negatively impacting oral health-related quality of life (6). Early detection of arthritis of the TMJ may facilitate intervention to prevent joint damage and dysfunction.

Arthritis of the TMJ cannot be assessed comprehensively by physical examination, ultrasound, conventional radiographs, or computed tomography imaging (7–13). Contrast-enhanced magnetic resonance imaging (MRI) remains the best available diagnostic tool as it allows for visualization of both soft tissue and osteochondral changes in the TMJ. Since many early changes are subtle, the evaluation of TMJ MRI remains subjective and necessitates a standardized and feasible outcome measure. To this end, the JIA MRI working group (JAMRI) within the Outcome Measures in Rheumatology (OMERACT) research network has recently developed the JIA MRI scoring system for the TMJ (JAMRIS-TMJ) (14).

The JAMRIS-TMJ is constructed as a multiitem, additive outcome measure with each joint graded by inflammatory and damage domains. Once the scoring items and feasible grading criteria are defined, the relative importance weights of the items and their grades must be determined and validated for deriving composite domain scores. For example, studies have identified that mild levels of effusion and synovial enhancement are not specific to TMJ arthritis (15–17), emphasizing that MRI-observable features and their levels have different and context-specific importance when interpreting MRI of TMJs. A discrete choice experiment (DCE) is helpful in this regard, offering a formalized and quantitative approach for eliciting the opinions of an expert panel in defining the relative importance weights of items in this type of measure (18–21). For a brief background on DCE, see Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24577/abstract.

In this study, we determined the relative importance weights of items and grades of the JAMRIS-TMJ using a DCE survey (22). The resulting weighting scheme enables the calculation of percentage-wise inflammation and damage domain scores using the JAMRIS-TMJ method of MRI evaluation. To test the validity of the elicited weights, we conducted a vignette ranking exercise. The weighted JAMRIS-TMJ score ranking approach was tested against a holistic, image-to-image comparison approach as the reference standard, because the latter method allows greater differentiation and does not entail the reductionistic assumptions of the DCE process or the restrictions inherent in the JAMRIS-TMJ grading criteria. The specific aims of the study were as follows: 1) to determine the relative importance weights of the items and grades in the JAMRIS-TMJ using an adaptive DCE method within a multicenter, multispecialty group of experts; and 2) to assess the validity of the DCE-derived importance weights using an image vignette-based exercise by testing the correlation of the JAMRIS-TMJ weighted vignette score with the vignette rank given through a scoring system–independent method.

MATERIALS AND METHODS

This study was approved by the Research Ethics Board (REB) of The Hospital for Sick Children (Toronto, Canada; study reference 1000042164). Information letters were provided to the participants before each activity to explain the study and that their voluntary completion and submission of the study surveys constituted their implied consent to participate in the study. Considering the practical limitations, and that the imaging exams used for creating the vignettes were anonymized and retrospective in nature, written consent requirement was waived by the REB. The study was conducted in 2 phases, the first being the DCE survey to develop the relative importance weights, and the second being the vignette ranking exercise that tested the face and convergent validity (23) of the DCE–weighted scoring system. Figure 1 summarizes the methods in a flow chart.

DCE survey. An adaptive, partial-profile DCE survey administered through the 1000Minds software (22) was completed independently and anonymously by a multidisciplinary group of experts. Radiologists and other clinicians were invited if they

Medical Center, Cincinnati, Ohio; [20]Jyoti Panwar, MD: Christian Medical College and Hospital, Vellore, Tamil Nadu, India; [21]Cory M. Resnick, DMD, MD: Boston Children's Hospital, Boston, Massachusetts; [22]Susan C. Shelmerdine, MBBS: Great Ormond Street Hospital, London, UK.

No potential conflicts of interest relevant to this article were reported.

Address correspondence to Andrea S. Doria, MD, PhD, MSc, MBA, The Hospital for Sick Children, 555 University Avenue, 2nd floor, Toronto, Ontario M5G1X8, Canada. Email: andrea.doria@sickkids.ca.

Submitted for publication April 17, 2020; accepted in revised form February 4, 2021.

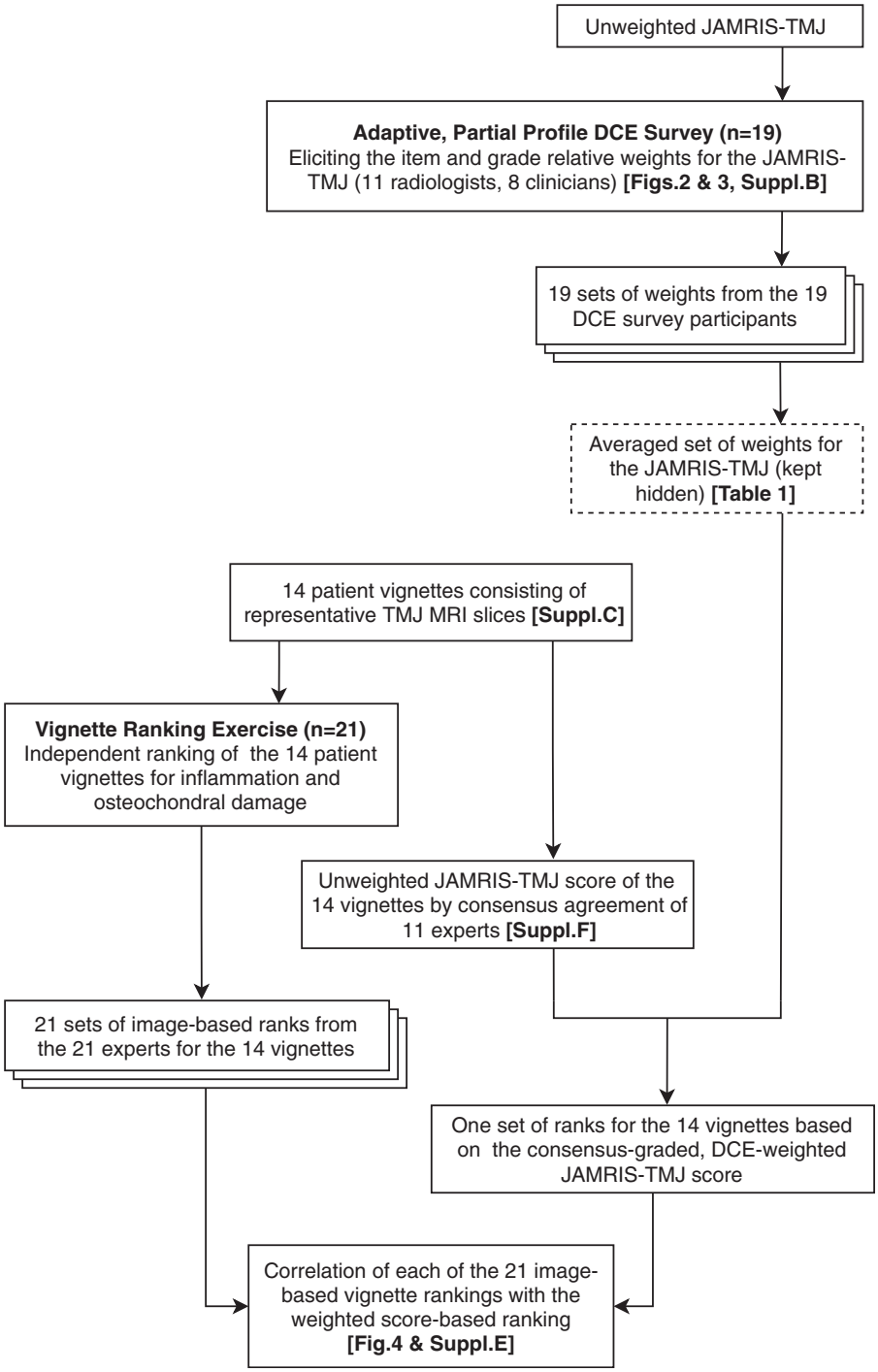

Figure 1. Flow chart summary outlining the progression of the study tasks in chronological order from top to bottom. First, an adaptive, partial-profile, discrete choice experiment (DCE) survey was completed individually by a group of experts (n = 19) to determine the importance weights of the items and grades of the juvenile idiopathic arthritis magnetic resonance imaging (MRI) scoring system for the temporomandibular joint (TMJ) (JAMRIS-TMJ). Second, blinded to the DCE-derived weights, an image-based vignette ranking exercise was completed individually by experts (n = 21), producing 21 sets of both the inflammatory disease and osteochondral damage severity rankings for a set of 14 patient vignettes based on a full-profile, scoring system–independent method of comparison. Then, the item-wise JAMRIS-TMJ grades for the vignettes were agreed upon by consensus of experts (n = 11), and the DCE-derived weights were applied to obtain the consensus-weighted score for the vignettes for the 2 domains. Finally, the resulting vignette rankings from the 2 methods were correlated to test for convergent validity of the weighted JAMRIS-TMJ score. Fig. = figure; Suppl. = Supplementary Appendix, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24577/abstract.

routinely assessed MRIs of the TMJ in patients with JIA. Each expert participant completed separate DCE surveys for the inflammatory and damage domains. All discrete choice questions asked the expert to compare 2 hypothetical sets of findings with different, nondominating grades in the same 2 JAMRIS-TMJ domain items and to choose which scenario represented "more severe disease, assuming all else being equal," or to rate them as equal (see Supplementary Appendix B, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24577/abstract). The relative weights were derived by the 1000Minds software utilizing the PAPRIKA (Potentially All Pairwise Rankings of All Possible Alternatives) method (22). A complete set of item- and grade-importance weights was obtained for each DCE survey participant. The individual sets of weights were averaged over the entire group of experts to serve as the relative weights for the scoring system for testing. The weights were kept hidden until after the ranking exercise.

Vignette ranking exercise. Convergent validity of the weighted JAMRIS-TMJ was tested through a vignette ranking exercise conducted by a multidisciplinary group of radiologists and other non-radiologist clinicians within the JAMRI working group. Fourteen vignettes representing single TMJs from JIA patients were constructed from representative slices from each of the 6 imaging sequences from a TMJ MRI protocol for JIA utilizing dedicated surface coils (see Supplementary Appendix C, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24577/abstract). The 6 images consisted of 3 precontrast sequences (fat-suppressed sagittal oblique T2, sagittal oblique proton density–weighted, and coronal T1-weighted) and 3 gadolinium-enhanced T1-weighted fat-suppressed sequences in 3 planes (axial, sagittal oblique, and coronal). Participants independently ranked these vignettes in increasing order of severity of inflammation and osteochondral damage, allowing for tied ranks. Item-wise grades of the 14 vignettes achieved by consensus of 2 radiologists (TJ and ASD) were provided for a subgroup of clinician participants who do not regularly interpret TMJ MRI exams themselves; hence, they ranked graded images. To simulate a pragmatic and holistic method of vignette-to-vignette comparison that is independent of any scoring method, all participants were instructed not to base their ranking on any summation of scores, allowing for the possibility that more important items or certain combinations of item grades can disproportionally influence the disease severity ranking.

The item-wise JAMRIS-TMJ raw scores for each of the 14 vignettes were decided by consensus during a face-to-face and video conference meeting among a subgroup of participants ($n = 11$) who regularly interpret TMJ MRI examinations. Weighted JAMRIS-TMJ scores for the 14 vignettes were produced using these consensus grades and the importance weights derived from the DCE. The weighted JAMRIS-TMJ score was then correlated with the ranking provided by each of the participating experts. This correlation tested the combined impact of several factors related to the face and content validity of the weighted JAMRIS-TMJ: the items, grades, and their relative weights; the joint factor independence (21), transitivity, and other assumptions of the adaptive partial-profile DCE method used to derive the weights (22); as well as the discriminative capacity and feasibility of the grading criteria.

Sample size considerations. The adaptive DCE method from 1000Minds that we used for generating the weights provides a complete set of weights for each item and grade level of the scoring system for every participant (see Supplementary Appendix A and D, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24577/abstract) (22). Therefore, the sample size requirement for the number of participants was not based on quantitative simulations for model convergence, but instead, on achieving a comprehensive and saturated opinion base that is representative of the level of heterogeneity among clinicians from multiple centers and specialties. Convenience sampling from an international research interest group was used to enroll experts from multiple specialties for the 2 study exercises. The number of vignettes used for the ranking was also subjectively determined to provide a balance between representing the common item combinations across the spectrum of 2 disease domains and reducing respondent error.

Statistical analysis. In the DCE survey, homogeneity of the relative weights within the expert group was assessed in 2 ways. First, the representativeness of the group-averaged set of relative weights was tested by calculating the Spearman's rank correlation coefficients (ρ) between rankings of all potential item combinations produced by group-averaged weights and each of the participants' weights (22). Second, the agreement of the relative weights among the participants was assessed by calculating the intraclass correlation coefficients (ICC; 2-way random, single measure, absolute agreement type). In the vignette ranking exercise, agreement in the vignette rankings among the participants was assessed visually per vignette by scatterplots and quantitatively by calculating the ICC of ranks given to each of the 14 vignettes. Spearman's ρ was used for correlating the image-based ranking with the weighted JAMRIS-TMJ score. For both correlation coefficients, values ≤ 0.4 were defined as poor correlation, 0.41–0.60 as moderate, 0.61–0.80 as substantial, and ≥ 0.81 as high correlation. Statistical analyses were performed using SPSS, version 23. For further details regarding the DCE survey and the statistical tests used, see Supplementary Appendix D, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24577/abstract.

RESULTS

Nineteen experts completed the DCE survey in total, including 11 pediatric or maxillofacial radiologists, 7 pediatric rheumatologists (6 of whom self-identified as not regularly interpreting TMJ

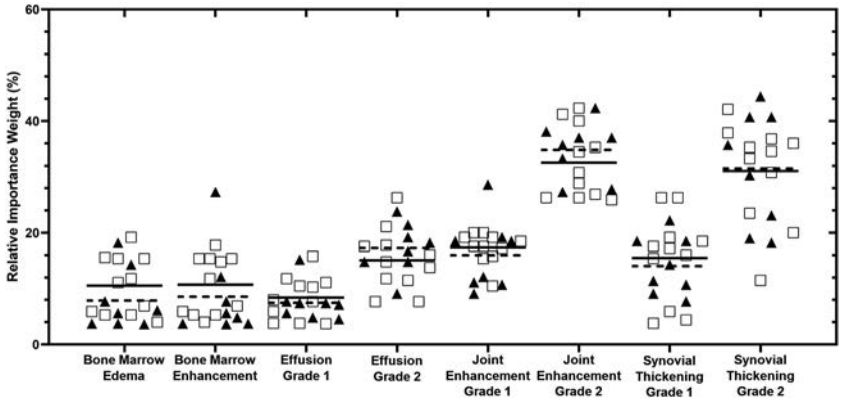

Figure 2. Scatterplot of the item- and grade-relative weights obtained from the discrete choice experiment survey for the juvenile idiopathic arthritis magnetic resonance imaging scoring system for the temporomandibular joint inflammatory domain. Relative importance weights from each of the participants are plotted. Lines indicate the average weight for radiologists (solid lines and squares; n = 11) and non-radiologist clinicians (broken lines and triangles; n = 8) for each of the item grades.

MRIs themselves), and 1 orthodontist, yielding 19 sets of item and grade weights. Approximately 20–25 discrete choice questions (see Supplementary Appendix B, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24577/abstract) were required to obtain a full set of relative importance weights for the 2 domains of the scoring system for each participant (Figures 2 and 3): the 5-item inflammatory domain required between 14 and 18 questions, and the 3-item damage question required between 5 and 7. The number of questions varied between the participants due to the differences in opinion and the order in which the questions were presented (see Supplementary Appendix A and D, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24577/abstract). Quantitative indices of the group's homogeneity on these weights were sufficiently high: the ranking of all possible combinations of items that is produced from each expert's weights correlated highly with the rank produced by the group-averaged weights, with a median Spearman's ρ of 0.96 for the inflammation domain (interquartile range [IQR] 0.93–0.96) and 0.97 for the damage domain (IQR 0.95–0.99); group-wide, 19-rater agreement on these 8 and 5 non-zero weights for the 2 domains was substantial, at 0.71 for the inflammatory domain weights and 0.77 for the damage domain. Therefore, the average of the 19 sets of weights from the experts was deemed representative to be used as the JAMRIS-TMJ weights (Table 1), which were kept hidden prior to the vignette ranking exercise.

In total, 21 experts, consisting of 11 pediatric or maxillofacial radiologists, 7 pediatric rheumatologists, 2 oral and maxillofacial surgeons, and 1 orthodontist (13 overlapping with the experts who participated in the DCE) completed the vignette ranking exercise. Overall, the ranks given to the 14 vignettes correlated substantially among the 21 participants, with the ICC of the inflammatory domain vignette ranking at 0.85, and the damage domain ranking at 0.91.

The group-averaged relative weights from the DCE survey revealed several differences between the items of the JAMRIS-TMJ

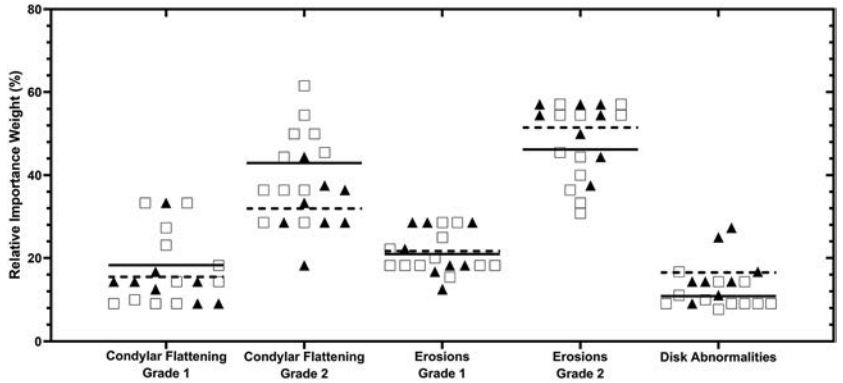

Figure 3. Scatterplot of the item- and grade-relative weights obtained from the discrete choice experiment survey for the juvenile idiopathic arthritis magnetic resonance imaging scoring system for the temporomandibular joint damage domain. Relative importance weights from each of the participants are plotted. Lines indicate the average weight for radiologists (solid lines and squares; n = 11) and non-radiologist clinicians (broken lines and triangles; n = 8) for each of the item grades.

Table 1. Relative importance weights derived from the discrete choice experiment for the items and levels of the juvenile idiopathic arthritis magnetic resonance imaging scoring system for the temporomandibular joint*

Item	Grading level and weight					
	0	%	1	%	2	%
Inflammatory domain						
Bone marrow edema	Absent	0	Present	9	–	–
Bone marrow enhancement	Absent	0	Present	10	–	–
Effusion	Normal: ≤1 mm in the largest joint recess	0	Mild: >1 and ≤2 mm in the largest joint recess	8	Moderate/severe: >2 mm focally and/or extension to entire joint	16
Joint enhancement	Normal: no exceeding joint enhancement	0	Mild: localized exceeding joint enhancement	17	Moderate/severe: exceeding joint enhancement diffusely involving the joint	34
Synovial thickening	Normal: no synovium visible	0	Mild: ≤2 mm thickness at the point of maximum synovial thickening	15	Moderate/severe: >2 mm	31
Damage domain						
Condylar flattening	Normal round/ovoid shape	0	Mild: extent of flattening involves part of the surface of the condyle	17	Moderate/severe: extent of flattening involves the entire surface of the condyle, or loss of height in the condyle	38
Erosions	No irregularities or deep breaks	0	Mild: presence of irregularities involving only part of the articular surface of the condyle	21	Moderate/severe: presence of deep breaks in the subchondral bone seen in 2 planes, or irregularities involving the entire articular surface of the condyle	49
Disk abnormalities	Absent	0	Present	13	–	–

* After an image has been graded, the total score for each domain is calculated by adding the percentage weight of each given grade for all items to yield a scaled percentage disease severity score ranging from 0–100% for each domain separately. Weights presented in this table are the group-averaged weights from Figures 2 and 3.

and their grade levels (Table 1). Highest grade joint enhancement showed a 34% relative weight for assessment of inflammation compared to the highest-grade joint effusion (16%) and bone marrow enhancement (10%). Condylar flattening and erosions showed nonlinear changes between grade levels, with the second grade-level being weighted higher per score than the first. Differences between the radiologists and non-radiologist clinicians on the relative weights were not statistically different when adjusted for multiple testing. The participants agreed that the group-averaged set of importance weights seem to be an appropriate representation of the group's opinion for use in subsequent construct validity studies and justifiable considering the current understanding of TMJ arthritis and the clinimetric properties of observable items in contrast-enhanced MRI. Nevertheless, in examining the range of potential item combinations for the 2 domains (up to 108 and 18 for the inflammatory and damage domain, respectively), it was identified that 3 of 4 potential item-grade combinations in the damage domain between weighted scores of 52% to 78% may be quite rare or impossible to obtain: grade 2 flattening and grade 1 erosions (59%), grade 1 flattening and grade 2 erosions (66%), both with no disk abnormalities, and grade 2 erosions with disk abnormalities but no flattening (62%).

The consensus DCE-weighted JAMRIS-TMJ score for the 14 vignettes correlated very highly with the 21 sets of vignette ranks generated from the image-based ranking exercise, with median Spearman's ρ of 0.92 (IQR 0.87–0.95) for the inflammatory domain and 0.93 (IQR 0.90–0.94) for the damage domain (Figure 4). Vignettes that received weighted scores placing midway in the disease severity spectrum showed more variability in the image-based ranking than those with weighted scores near the 2 extremes (see Supplementary Appendix E, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24577/abstract). No significant subgroup differences were observed between the participants who performed the image-only ranking (those who self-identified as reading TMJ MRIs regularly; $n = 15$) versus those who performed the graded image ranking (those who do not usually interpret TMJ MRIs themselves; $n = 6$).

The full-profile comparison of the patient vignettes was not restricted in terms of the items and grading cutoffs of the scoring system, allowing higher levels of differentiation between disease stages and therefore a greater potential for disagreement in vignette ranks between the 2 methods. During the postexercise discussions, it was identified that there were subtle but appreciable differences in the image-based ranking of the vignettes that were not differentiated by change in the JAMRIS-TMJ score. These scenarios could be described as "high" grade 1 versus "low" grade 1 within the confines of the grading threshold. Vignettes with unreliable, borderline grading (e.g., considering a

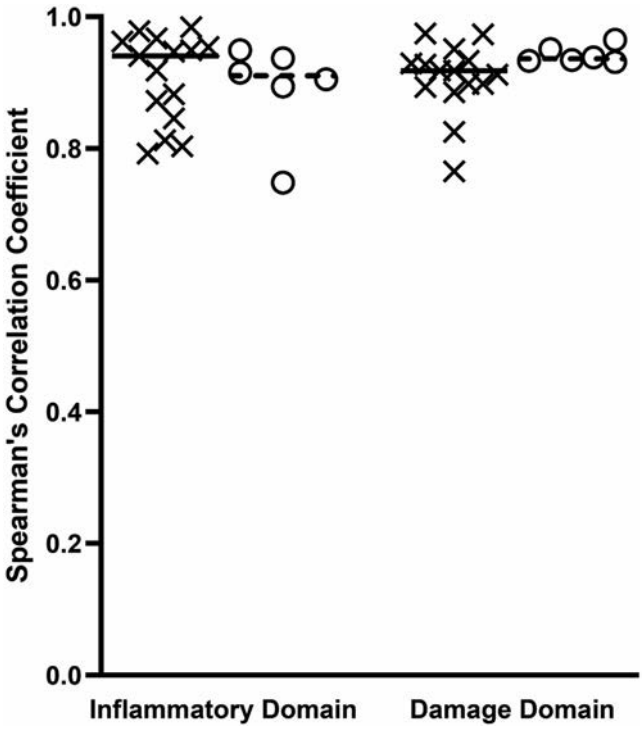

Figure 4. Correlation of the vignette ranks produced by the unrestricted, image-based ranking method and weighted juvenile idiopathic arthritis magnetic resonance imaging (MRI) scoring system for the temporomandibular joint (TMJ) (JAMRIS-TMJ) score. Separate Spearman's rank correlation coefficients (ρ) are plotted for each of the participants (n = 21) comparing their image-based ranking of the 14 patient vignettes with 1 set of consensus-graded, discrete choice experiment–weighted JAMRIS-TMJ score ranks for the vignettes. Horizontal lines indicate the median Spearman's ρ for each subgroup of participants: one group ranked the vignettes by the images only (cross markers and solid lines; n = 15); the other group, consisting of pediatric rheumatologists who do not regularly interpret and grade TMJ MRIs themselves, ranked the vignettes by the unweighted, pre-graded images (circles and broken lines; n = 6). The data from which these coefficients derive are visualized as scatterplots in Supplementary Appendix E, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24577/abstract.

feature as "high grade 1" or "low grade 2") or those that require knowing the patient's age or comparison with the contralateral TMJ challenged the ranking task, especially with regard to interpreting condylar flattening and bone marrow changes.

DISCUSSION

In this study, we used an adaptive partial-profile DCE method to formalize the assignment of quantitative importance weights to the items and grades of the JAMRIS-TMJ. Synovial thickening and joint enhancement items were considered by the expert panel on average twice as important per raw score compared to the other 3 inflammatory domain items (Figure 2 and Table 1). This finding underlines the diagnostic importance for contrast administration for assessment of TMJs, although this has become more restricted in clinical practice due to potential concerns with cumulative deposition of gadolinium in the body (24,25). In the damage domain, erosions were weighted the most important, followed by condylar flattening, both with nonlinear per score weights, then disk abnormalities (Figure 3 and Table 1). The nonlinear increase in weights of grades for these damage domain items better represents the ordinal scaling of the grading definitions for these items compared to unweighted scoring. In general, the weighting scheme represents the features of both the progressive and additive TMJ MRI scoring systems that the JAMRIS-TMJ was derived from (26–28), emphasizing the diagnostic features with higher specificity for active inflammation while still allowing for further differentiation by ancillary items.

The JAMRIS-TMJ grading method focuses on measuring the items as independently as possible. Synovial thickening is measured only on fluid-sensitive sequences as presenting with intermediate signal intensity on MRI; pockets of fluid need to be considered in grading joint enhancement to distinguish them from enhanced synovium; and the bone marrow edema signal is considered only on precontrast images (14). To the extent that the items can be measured independently and that the various combinations of these items are realistic and informative, it should be useful to add these items to produce composite domain scores. For example, a region of synovium that does not enhance after contrast may suggest residual pannus from prior disease that is not currently inflamed, differentiating it from active disease. However, practical issues still can cause correlation or restriction of grades between items. When there is severe structural damage in the joint, some soft tissue components, such as inactive pannus, become difficult to identify and grade. Disease may also be overestimated when a given finding cannot be reliably attributed to a specific item: differentiation of soft tissue components is difficult if not impossible to assess using only postcontrast images, and comparing postcontrast with corresponding precontrast images may not always be helpful. Grading of these changes may be improved with the utilization of measurement aids for different stages of joint inflammation and degeneration such as by using an imaging atlas (29).

Nonlinearity in the change in weighted JAMRIS-TMJ score between adjacently ranked vignettes (see Supplementary Appendix F, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24577/abstract) likely resulted from the limited vignette selection. However, it may also suggest that some theoretical combinations of item scores in these intervals too rare or transient to be captured. A cross-sectional study using the scoring system on a large consecutive series of patients would be helpful to study the true prevalence in these intervals of the scoring spectrum.

The chief limitation of this study was that the number of vignettes that could be rank ordered was relatively low, precluding

a more complex study design that could directly quantify the advantage of weighted scores over unweighted scores when correlating to the holistic, image-based rank. To achieve this, it would be necessary to select the vignettes in a manner that maximizes the difference between the raw score and the weighted scores, allowing for a more efficient differentiation between the 2 correlations. Increasing the number of vignettes to serve this purpose would also be challenging because it would increase the cognitive burden of ranking, potentially leading the participants to use simplifying heuristics in their comparison of vignettes, and hence, skewing the way they applied the relative weights. Instead, the vignettes were selected to better represent the various common presentations across the entirety of the scoring spectrum in both domains, thus capturing more of the nuances in item combinations.

In conclusion, the DCE survey facilitated the development of relative importance weights of items and grades in the JAMRIS-TMJ, which showed high convergent validity with a holistic, scoring system–independent method of image assessment when applied to rank a series of TMJ MRI vignettes. The relative weights derived from the DCE revealed differences between the items as well as between the different grades of items, which would not be captured by the number of grades allotted to the items. The weighting scheme is therefore crucial for scaling the JAMRIS-TMJ inflammatory and damage domain scores in accordance with the perceived differences in the items and their grade levels, enabling their application as standardized outcome measures in clinical practice and research, including clinical trials in JIA. Our methodology combining adaptive DCE with validation by subsequent holistic vignette ranking exercise could be applied to relative weighting of components of other imaging-based grading systems.

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. Doria 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. Tolend, Junhasavasdikul, Miller, Spiegel, Tzaribachev, Abramowicz, Appenzeller, Jaremko, Feldman, Doria.

Acquisition of data. Tolend, Junhasavasdikul, Clemente, von Kalle, Kellenberger, Koos, Miller, van Rossum, Saurenmann, Spiegel, Stimec, Twilt, Tzaribachev, Arvidsson, Guleria, Kirkhus, Larheim, Meyers, Panwar, Resnick, Shelmerdine, Doria.

Analysis and interpretation of data. Tolend, Junhasavasdikul, Cron, Miller, Appenzeller, Jaremko, Meyers, Resnick, Shelmerdine, Feldman, Doria.

REFERENCES

1. Manners PJ, Bower C. Worldwide prevalence of juvenile arthritis why does it vary so much? J Rheumatol 2002;29:1520–30.
2. Larheim TA, Doria AS, Kirkhus E, Parra DA, Kellenberger CJ, Arvidsson LZ. TMJ imaging in JIA patients—an overview. Semin Orthod 2015;21:102–10.
3. Cannizzaro E, Schroeder S, Müller LM, Kellenberger CJ, Saurenmann RK. Temporomandibular joint involvement in children with juvenile idiopathic arthritis. J Rheumatol 2011;38:510–5.
4. Stoll ML, Sharpe T, Beukelman T, Good J, Young D, Cron RQ. Risk factors for temporomandibular joint arthritis in children with juvenile idiopathic arthritis. J Rheumatol 2012;39:1880–7.
5. Twilt M, Mobers SM, Arends LR, ten Cate R, van Suijlekom-Smit L. Temporomandibular involvement in juvenile idiopathic arthritis. J Rheumatol 2004;31:1418–22.
6. Rahimi H, Twilt M, Herlin T, Spiegel L, Pedersen TK, Küseler A, et al. Orofacial symptoms and oral health-related quality of life in juvenile idiopathic arthritis: a two-year prospective observational study. Pediatr Rheumatol Online J 2018;16:47.
7. Weiss PF, Arabshahi B, Johnson A, Bilaniuk LT, Zarnow D, Cahill AM, et al. High prevalence of temporomandibular joint arthritis at disease onset in children with juvenile idiopathic arthritis, as detected by magnetic resonance imaging but not by ultrasound. Arthritis Rheum 2008; 58:1189–96.
8. Munir S, Patil K, Miller E, Uleryk E, Twilt M, Spiegel L, et al. Juvenile idiopathic arthritis of the axial joints: a systematic review of the diagnostic accuracy and predictive value of conventional MRI. AJR Am J Roentgenol 2014;202:199–210.
9. Koos B, Twilt M, Kyank U, Fischer-Brandies H, Gassling V, Tzaribachev N. Reliability of clinical symptoms in diagnosing temporomandibular joint arthritis in juvenile idiopathic arthritis. J Rheumatol 2014;41:1871–7.
10. Muller L, Kellenberger CJ, Cannizzaro E, Ettlin D, Schraner T, Bolt IB, et al. Early diagnosis of temporomandibular joint involvement in juvenile idiopathic arthritis: a pilot study comparing clinical examination and ultrasound to magnetic resonance imaging. Rheumatol Oxf Engl 2009;48:680–5.
11. Rongo R, Alstergren P, Ammendola L, Bucci R, Alessio M, D'Antò V, et al. Temporomandibular joint damage in juvenile idiopathic arthritis: Diagnostic validity of diagnostic criteria for temporomandibular disorders. J Oral Rehabil 2019;46:450–9.
12. Bernini JM, Kellenberger CJ, Eichenberger M, Eliades T, Papageorgiou SN, Patcas R. Quantitative analysis of facial asymmetry based on three-dimensional photography: a valuable indicator for asymmetrical temporomandibular joint affection in juvenile idiopathic arthritis patients? Pediatr Rheumatol Online J 2020;18:10.
13. Zwir LF, Terreri MT, Castro AD, Rodrigues WD, Fernandes AR. Is power Doppler ultrasound useful to evaluate temporomandibular joint inflammatory activity in juvenile idiopathic arthritis? Clin Rheumatol 2020;39:1237–40.
14. Tolend MA, Twilt M, Cron RQ, Tzaribachev N, Guleria S, von Kalle T, et al. Toward establishing a standardized magnetic resonance imaging scoring system for temporomandibular joints in juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2018;70: 758–67.
15. Stoll ML, Guleria S, Mannion ML, Young DW, Royal SA, Cron RQ, et al. Defining the normal appearance of the temporomandibular joints by magnetic resonance imaging with contrast: a comparative study of children with and without juvenile idiopathic arthritis. Pediatr Rheumatol Online J 2018;16:8.
16. Ma GM, Calabrese CE, Donohue T, Peacock ZS, Caruso P, Kaban LB, et al. Imaging of the temporomandibular joint in juvenile idiopathic arthritis: how does quantitative compare to semiquantitative MRI scoring? J Oral Maxillofac Surg 2019;77:951–8.
17. Tzaribachev N, Fritz J, Horger M. Spectrum of magnetic resonance imaging appearances of juvenile temporomandibular joints (TMJ) in non-rheumatic children. Acta Radiol Stockh Swed 1987 2009;50: 1182–6.
18. Ryan M, Farrar S. Using conjoint analysis to elicit preferences for health care. BMJ 2000;320:1530–3.

19. Burnett HF, Regier DA, Feldman BM, Miller FA, Ungar WJ. Parents' preferences for drug treatments in juvenile idiopathic arthritis: a discrete choice experiment. Arthritis Care Res (Hoboken) 2012;64: 1382–91.
20. Neogi T, Aletaha D, Silman AJ, Naden RL, Felson DT, Aggarwal R, et al. The 2010 American College of Rheumatology/European League Against Rheumatism classification criteria for rheumatoid arthritis: phase 2 methodological report. Arthritis Rheum 2010;62:2582–91.
21. Krantz DH. Measurement structures and psychological laws. Science 1972;175:1427–35.
22. Hansen P, Ombler F. A new method for scoring additive multi-attribute value models using pairwise rankings of alternatives. J Multi Criteria Decis Anal 2008;15:87–107.
23. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol 2010;63:737–45.
24. Kuno H, Jara H, Buch K, Qureshi MM, Chapman MN, Sakai O. Global and regional brain assessment with quantitative MR imaging in patients with prior exposure to linear gadolinium-based contrast agents. Radiology 2016;283:195–204.
25. Murata N, Murata K, Gonzalez-Cuyar LF, Maravilla KR. Gadolinium tissue deposition in brain and bone. Magn Reson Imaging 2016;34: 1359–65.
26. Vaid YN, Dunnavant FD, Royal SA, Beukelman T, Stoll ML, Cron RQ. Imaging of the temporomandibular joint in juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2014;66:47–54.
27. Koos B, Tzaribachev N, Bott S, Ciesielski R, Godt A. Classification of temporomandibular joint erosion, arthritis, and inflammation in patients with juvenile idiopathic arthritis. J Orofac Orthop 2013;74: 506–19.
28. Kellenberger CJ, Arvidsson LZ, Larheim TA. Magnetic resonance imaging of temporomandibular joints in juvenile idiopathic arthritis. Semin Orthod 2015;21:111–20.
29. Kellenberger CJ, Junhasavasdikul T, Tolend M, Doria AS. Temporomandibular joint atlas for detection and grading of juvenile idiopathic arthritis involvement by magnetic resonance imaging. Pediatr Radiol 2018;48:411–26.

Arthritis Care & Research
Vol. 74, No. 2, February 2022, pp 317–323
DOI 10.1002/acr.24475

American College of Rheumatology
Empowering Rheumatology Professionals

Gout Flare Severity From the Patient Perspective: A Qualitative Interview Study

Andrea Garcia-Guillen,[1] Sarah Stewart,[2] Isabel Su,[2] William J. Taylor,[3] Angelo L. Gaffo,[4] Merryn Gott,[2] Julia Slark,[2] Anne Horne,[2] and Nicola Dalbeth[2]

Objective. The patient experience of a gout flare is multidimensional. To establish the most appropriate methods of flare measurement, there is a need to understand the complete experience of a flare. This qualitative study aimed to examine what factors contribute to the severity of a flare from the patient perspective.

Methods. Face-to-face interviews were conducted with patients with gout. Participants were asked to share their experience with their worst gout flare and contrast it to their experience of a less severe or mild flare. Interviews were audio recorded and transcribed verbatim. Data were analyzed using a reflexive thematic approach.

Results. In total, 22 participants with gout (17 male participants, mean age 66.5 years) were interviewed at an academic center in Auckland, New Zealand. Four key themes were identified as contributing to the severity of a flare: 1) flare characteristics (pain intensity, joint swelling, redness and warmth, duration, and location); 2) impact on function (including walking, activities of daily living, wearing footwear, and sleep); 3) impact on family and social life (dependency on others, social connection, and work); and 4) psychological impact (depression, anxiety, irritability, and sense of control).

Conclusion. A wide range of interconnecting factors contribute to the severity of a gout flare from the patient perspective. Capturing these domains in long-term gout studies would provide a more meaningful and accurate representation of cumulative flare burden.

INTRODUCTION

Gout flares (sometimes referred to as "gout attacks" or "acute gout") are a characteristic feature of gout and a central concern to patients (1). An important goal in the management of gout is complete suppression and prevention of gout flares. However, there is currently no standardized method for the assessment of gout flares in clinical trials. Content analyses have shown a wide variation in methods used to measure and report flares in clinical trials of flare prevention (2,3). The majority of studies capture data related to flare frequency, with few studies also reporting data related to flare duration and pain severity (2).

The patient experience of a gout flare is multidimensional and goes far beyond the data routinely captured in clinical trials. A recent meta-synthesis of qualitative studies illustrated the impact of gout flares on many aspects of patients' lives, including physical, social, and family life and psychological well-being (4). The interconnecting nature of these domains highlighted the complexity of the flare experience, but it remains unclear which aspects of a flare are most important to patients in influencing the overall flare burden.

Establishing meaningful flare reporting in clinical trials would involve defining the most appropriate methods of gout flare measurement. An important step in achieving this is to better understand, from the patient perspective, which factors make the experience of a flare more or less severe. This qualitative study aimed to examine what factors contribute to the severity of a gout flare from the patient perspective.

Dr. Garcia-Guillen's work was supported by the Spanish Foundation for Rheumatology and the Catalan Society of Rheumatology. Dr. Stewart's work was supported by a fellowship from the Auckland Medical Research Foundation.

[1]Andrea Garcia-Guillen, MD: University of Auckland, Auckland, New Zealand, and Hospital de la Santa Creu i Sant Pau, Barcelona, Spain; [2]Sarah Stewart, PhD, Isabel Su, BSc, Merryn Gott, MA (Oxon), PhD, Julia Slark, PhD, Anne Horne, MBChB, Nicola Dalbeth, MBChB, MD, FRACP: University of Auckland, Auckland, New Zealand; [3]William J. Taylor, MBChB, PhD, FRACP, FAFRM: University of Otago, Wellington, New Zealand; [4]Angelo L. Gaffo, MD, MsPH, FACP: UAB Health System and Birmingham VA Medical Center, Birmingham, Alabama.

Dr. Gaffo has received research support from Amgen. Dr. Dalbeth has received consulting fees, speaking fees, and/or honoraria from AstraZeneca, Horizon, Amgen, AbbVie, Pfizer, and Janssen (less than $10,000 each). No other disclosures relevant to this article were reported.

Address correspondence to Sarah Stewart, PhD, University of Auckland Faculty of Medical and Health Sciences, Medicine, 85 Park Road Grafton, Auckland 1023, New Zealand. Email: sarah.stewart@auckland.ac.nz.

Submitted for publication June 30, 2020; accepted in revised form September 29, 2020.

SIGNIFICANCE & INNOVATIONS

- This is the first study to examine factors contributing to the overall severity of a flare from the patient perspective.
- Multiple interconnected domains contribute to the overall severity of a flare; flare characteristics, impact on function and activities of daily living, psychological impact, and impact on family and social life were the key themes.
- Measuring these domains in studies assessing flare management or prevention may provide a more meaningful and accurate representation of cumulative flare burden from the patient perspective.

PATIENTS AND METHODS

Participants. Participants were recruited through existing databases of patients with gout who had participated in research at the Clinical Research Centre, University of Auckland, New Zealand and consented to be contacted for future studies. Purposive sampling was used to ensure a broad and diverse representation of demographic variables (age, ethnicity, sex) and gout disease characteristics (disease duration, tophaceous gout, flare frequency). Participants were included if they had gout according to the 2015 American College of Rheumatology/European Alliance of Associations for Rheumatology gout classification criteria (5), if they had at least 1 gout flare in the last 12 months, if they were age ≥18 years, and if they were English speaking. Participants were excluded if they had a cognitive impairment that would preclude completion of the interview or had other forms of inflammatory arthritis. Ethics approval was obtained from the University of Auckland Human Participants Ethics Committee (UAHPEC 023965), and all participants provided written informed consent.

Data collection. In-depth, semistructured face-to-face interviews were conducted by a rheumatologist who was not involved in the medical care of the participant (AG-G). Participants were asked to share their experience of flares during the course of their disease by recalling their worst gout flare as well as a less severe/mild gout flare in order to capture factors contributing to overall flare severity. An interview schedule containing key focused, open-ended questions and probes was used to encourage conversation. These questions included: "Can you tell me about the worst gout flare you have had?" "What was it about this flare that made it so severe?" "Can you tell me about a mild gout flare?" and "What was it about this flare that made it less severe?" The questions and probes were elicited from a meta-synthesis of qualitative studies reporting the patient experience of gout flares (4).

The interviews took place in a private room at the Clinical Research Centre (University of Auckland, New Zealand) and lasted between 20 and 45 minutes. Each interview was digitally audio recorded, transcribed verbatim, and anonymized to ensure confidentiality. Participants had the opportunity to review the transcripts to check for completeness and representativeness. Demographic and clinical data were also obtained during the participants' study visit, including age at onset of gout, ethnicity, and presence and history of clinical features of gout and treatment.

Data analysis. Data collection and analysis occurred simultaneously, and initial results informed successive sampling and data collection as themes emerged. Interviews continued until no new themes were identified from the data and the purposive sampling framework was completed. Data was analyzed using a reflexive thematic approach (6). Transcripts from the interviews were read and re-read to immerse the researcher in the data. Emergent themes identified from the transcripts were initially coded and categorized by a single researcher (AG-G) using NVivo software, version 12 (QSR International Property). Initial codes and concepts were reviewed by 2 further researchers (SS and IS), and final codes were then grouped into potential themes and subthemes. The researchers met regularly to discuss the data throughout the analysis stage, and the final themes were defined, named, and agreed upon by all authors. Illustrative quotes from transcripts were selected to provide evidence for each theme and subtheme.

RESULTS

Study participants. A total of 22 participants with gout were interviewed. There was diversity across age, sex, ethnicity, and clinical features (Table 1).

Themes. Four key themes were identified from the data. Participants described the characteristics of the flares, impact on physical function and activities of daily living, impact on social and family life, and psychological impact as contributing to the overall severity of a flare. Illustrative quotes are shown in Tables 2–5. A thematic map showing the 4 themes and subthemes is provided in Figure 1.

Gout flare characteristics. Pain intensity was reported by all participants as contributing to gout flare severity. Participants described the pain of a severe flare as "intense," "extreme," "excruciating," "horrible," and "horrendous." Several participants also described it as the worst pain they had ever experienced, which for some was worse than a broken bone, abdominal surgery, or giving birth. The pain of a severe gout flare was described as constant and unchanging with no ability to alleviate it. Many participants described the intensity of pain using a numeric rating scale that varied from 6 to 11 of 10. In contrast, the pain of a mild flare was described as "uncomfortable" and "awkward." The pain of a mild flare was compared to having tight muscles, exercise-related soreness, feeling very stiff, or stubbing a toe. Mild flares were described with numeric pain ratings ranging from 2 to 4 of 10.

Table 1. Participant demographic and clinical characteristics (n = 22)*

Characteristic	Value
Sex	
Male	17 (77)
Female	5 (23)
Age, median (range) years	67 (27–84)
Ethnicity	
New Zealand European	12 (55)
Māori	5 (23)
Asian	3 (14)
Pacific peoples	2 (9)
Disease duration, median (range)	10.5 (6 months–35 years)
Age at onset of gout, median (range) years	48.5 (20–81)
No. of flares in the last 12 months	
1–4	16 (73)
5–9	2 (9)
≥10	4 (18)
Tophaceous gout	5 (23)
Last serum urate level, median (range) mmoles/liter	0.31 (0.18–0.64)
Urate-lowering therapy	
Allopurinol	18 (82)
Febuxostat	1 (5)
Allopurinol plus probenecid	1 (5)
None	2 (9)

* Values are the number (%) unless indicated otherwise.

A severe gout flare was also accompanied by intense joint swelling, warmth, and redness. Participants described worsening of these symptoms as the flare progressed. In contrast, participants noted a low level or complete absence of these characteristics during mild flares.

Flare duration was an important factor in contributing to flare severity. Participants described a severe flare as lasting multiple days and even for weeks, while mild flares resolved much more quickly, sometimes within hours.

The location of the gout flare also influenced overall flare severity. Gout flares affecting larger joints, such as the knees, created greater functional difficulty than flares in smaller joints, such as the feet. Some participants also felt that flares involving joints that were easier to rest and had less impact on functional activities, such as wrists, were less severe compared with flares involving joints required for mobility, including feet and ankles.

Impact on function. The level of disability also influenced overall flare severity. During severe flares, walking was described as extremely difficult or completely impossible. Performing any function involving weightbearing, including standing, exacerbated the pain of the flare. Participants could be completely immobile during a severe flare and had to stay seated in a chair or lying in bed, while others described using a wheelchair or crawling or hopping to move around. In contrast, the ability to walk was not affected to the same extent during a mild flare. Although walking was still difficult and uncomfortable for some, the milder symptoms meant that it was easier to move around compared to a severe flare.

Participants described greater difficulty with activities of daily living during severe flares. Tasks such as holding a mug, getting

Table 2. Quotes illustrating flare characteristics*

Subtheme	Severe flares	Mild flares
Pain intensity	"[It's] just so intense…it's one of the worst pains I've ever had. I've had abdominal surgery, and it's not as bad as the gout" (Participant 3, M, 59 years, NZ European). "I've had broken ankles and broken knees and joints and stuff like that and I would rate those pain as, probably, out of 10, probably up in about 6, 7. But the gout would sit pretty close to a 9" (Participant 5, M, 57 years, Māori). "It's like someone's stabbing me with a bottle" (Participant 8, M, 44 years, Māori).	"You can feel it, it's there, but it's not a real, serious one… you can just sort of sense it, you know it's there" (Participant 6, M, 72 years, Asian). "It's somewhere more in background than in the foreground. Probably maybe more like a 4 out of 10, or something like that, and so you can sort of cope with it a bit more" (Participant 11, M, 57 years, NZ European). "It's a bit like if you're walking around the house in bare feet and you stub your toe—you accidentally kick a piece of furniture with your toe—and it hurts. It feels like that" (Patient 9, M, 69 years, NZ European).
Joint swelling, redness, warmth	"When it flares up, it's just redder and tighter than the not-so-bad ones" (Participant 16, F, 68 years, NZ European).	"There was no redness, no nothing" (Participant 21, F, 73 years, NZ European).
Duration	"Seven, 8 days, I was in real pain, and then it subsided, bit by bit" (Participant 10, M, 73 years, NZ European).	"It's just probably less than 24 hours or 24 hours" (Patient 14, M, 60 years, Asian).
Location	"It was 2 joints at the same time. And it was the same leg, so moving was especially hard" (Participant 12, M, 48 years, Asian). "Depending on where it is, like, even a mild one in my knees is still… incapacitating" (Participant 8, M, 44 years, Māori).	"But like if it's in the side of my foot, or my toe, or you know, you can sort of manage it…I can go to work. It's not comfortable, but it's tolerable, if you know what I mean?" (Participant 8, M, 44 years, Māori). "I would rather have it in the wrist than in the foot. [If] it was on the wrist I could manage it more. I mean, with your foot, I'm not, like, flexible, so it's hard to do anything with it. With the wrist, it's just much, much easier, 'cause I can rest it anywhere" (Participant 12, M, 48 years, Asian).

* NZ = New Zealand.

Table 3. Quotes illustrating the impact on physical function and activities of daily living*

Subtheme	Severe flares	Mild flares
Walking ability	"I thought that using a [walking] stick might help to walk around; it didn't help anything" (Participant 18, F, 61 years, Māori). "I had to ask the person that takes the wheelchair people if I could get a ride – I couldn't walk" (Participant 4, M, 59 years, Māori).	"It is a bit sore, but I can still walk and move around. It's a bit easier to manage" (Participant 17, M, 27 years, Pacific Island). "Walking becomes a little bit more difficult, but not impossible" (Patient 7, M, 82 years, NZ European).
Impact on activities of daily living	"I had trouble driving. [It] was difficult changing gear in a right-hand drive car. So you're changing gear with your left hand" (Participant 1, M, 74 years, NZ European). "When it's at its worst, I [wasn't] able to move my arm…it was difficult to even just get out of bed…shower, toilet. All those daily things, it was just, pretty much, very difficult" (Participant 19, M, 30 years, Pacific Island).	"You can't go right back to normal lifestyle, but you can do things, more things, than [if] it was severe" (Participant 19, M, 30 years, Pacific Island).
Ability to wear footwear	"It makes it very difficult to put proper shoes on, you've got to then go to a sandal-type thing until it reduces" (Participant 21, F, 77 years, NZ European).	"Sneakers – they're alright when it's mild – you sort of feel like maybe it's just keeping [the foot] still. Whereas you can't stand it being firm when it's really bad" (Participant 16, F, 68 years, NZ European).
Impact on sleep	"When it hits hard, you can't even put a sheet over because it hurts at night. And if you move at night it just hurts, so you keep waking up" (Participant 12, M, 48 years, Asian). "In bed, at night-time…it'll be throbbing and aching and hot. It makes it harder to get off to sleep" (Participant 20, F, 84 years, NZ European).	"When it's [not] really bad, usually you can sleep, get in a position where it's comfortable and you're not feeling anything" (Participant 8, M, 44 years, Māori).

* NZ = New Zealand.

out of bed, showering, going to the toilet, and walking up and down stairs were difficult during severe flares. In contrast, during mild flares, participants were still able to participate in most daily activities, even working out or playing golf.

Gout flare severity was also influenced by the ability to wear footwear. During severe flares, many were unable to wear shoes at all due to the accompanying swelling and pain. Some participants opted for more open-style or looser fitting shoes, such as sandals or jandals (flip-flops) or wore different shoes on each foot. In contrast, during mild flares, participants were often able to wear firmer, fitting sneakers, which could not be worn during a severe flare.

Most participants had difficulty sleeping during severe flares. The affected area was described as hypersensitive. Participants described difficulty finding a comfortable position to sleep in, with even the slightest movement causing pain. Pillows were used in an attempt to relieve pressure in the area and find a position that was comfortable enough to sleep in. The pain would sometimes wake participants up and prevent them from getting back to sleep. In contrast, during mild flares, the pain did not always wake participants from sleep, and participants found it easier to find comfortable positions.

Impact on family and social life. During severe flares, participants reported being dependent on others, including relying on family members to bring them things around the house, asking others to drive the car, and push them in a wheelchair.

A severe gout flare disrupted social connections, leading to physical and psychological withdrawal from family and friends. During a severe flare, participants wanted to be left alone and did not want to interact with their spouses or children.

During severe flares, many participants had to take days off work. For others who attended work, they described staggering around or remaining seated. Not being able to walk or drive meant that some participants had no way of getting to work during severe flares. In contrast, during a mild flare, participants did not

Table 4. Quotes illustrating the impact on family and social life*

Subtheme	Severe flares	Mild flares
Dependency on others	"I was dependent on mum and dad, and just my little sisters to get me things around the house" (Participant 17, M, 27 years, Pacific Island). "I remember my wife driving the car for me" (Participant 15, M, 78 years, NZ European).	No relevant quotes
Social connection	"I'll withdraw from my engagement with family and friends and what I might be doing, and so, they'll notice that you're off the grid" (Participant 11, M, 57 years, NZ European).	No relevant quotes
Impact on work	"I had to take a week off. Oh, 4 and a half days off work. Lucky I have an understanding boss" (Participant 12, M, 48 years, Asian).	"I can still go to work; it's still not ideal or comfortable, but I don't have to waste a sick day on not going" (Participant 8, M, 44 years, Māori). "Never had to take days off with a mild one" (Participant 17, M, 27 years, Pacific Island).

* NZ = New Zealand.

Table 5. Quotes illustrating psychological impact*

Severe flares	Mild flares
"I felt horrible. I've been suicidal…and when I got the gout it played around with my mind" (Participant 18, F, 61 years, Māori).	"No [it didn't impact my mood]. I had come to accept it for what it was" (Participant 2, M, 65 years, NZ European).
"You're all the time worrying about not hitting it against something or somebody just bumping you" (Participant 16, F, 68 years, NZ European).	No relevant quotes
"[With a severe flare] you don't have a lot of patience, even the cat kept away from me" (Participant 9, M, 69 years, NZ European).	No relevant quotes
"Just knowing that it's out of my control – like, no matter how much meds I take, doesn't really mean that it will stop the flare" (Participant 17, M, 27 years, Pacific Island).	"All the others I knew instantly what it was…I could feel it coming on… and so I got on the drugs as quickly as I could" (Participant 2, M, 65 years, NZ European).

* NZ = New Zealand.

have to take days off work and were able to comfortably perform sedentary or computer-based tasks.

Psychological impact. A severe flare led to feelings of helplessness and not being able to escape the pain. Participants experienced depression and low mood at the time of a severe flare. Participants also reported feeling anxiety and constantly worried about someone bumping them and exacerbating the pain. During milder flares, participants felt anxious about whether the flare would get worse, while other participants experienced less worry, as they knew it would get better again.

In a severe flare, many participants felt irritable. Not being able to do anything and having to take time off work resulted in frustration. During severe flares, participants also reported losing patience with others and became grumpy and cranky around family members, including their children. In contrast, participants described feeling only slight frustration during mild flares.

Another important contributor to the overall severity of flares was the sense of control. For many participants, their worst gout flare was their first one because they had never experienced anything like it and did not understand what was happening. Feelings

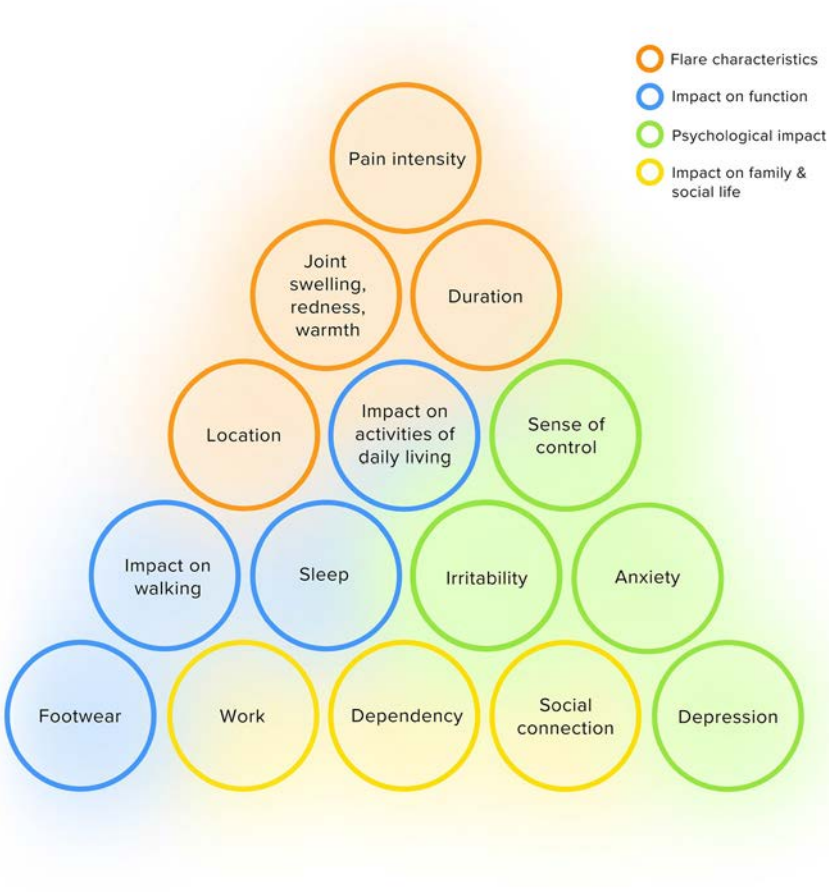

Figure 1. Mind map representing the 4 key themes and subthemes contributing to the overall severity of a gout flare.

of shock and an inability to control the symptoms contributed to the severity of the flare. During mild flares, participants described being able to sense a flare coming on. Knowing what to expect made the situation less shocking, and participants were able to initiate treatment quickly, which also prevented the pain from reaching the same peak as a severe flare.

DISCUSSION

This qualitative study provides in-depth insights into factors that contribute to the overall severity of a gout flare from the patient perspective. Although numerous studies have reported on the patients' experience of flares (summarized in ref. 4), this is the first study to specifically examine what factors contribute to the severity of a flare from the patient perspective. The impact of flare characteristics on function and activities of daily living, psychological impact, and impact on family and social life were the key themes.

Pain intensity was the dominant reported flare characteristic distinguishing a severe flare from a mild flare. The importance of pain is also reflected in its inclusion as a mandatory outcome measure proposed by Outcome Measures in Rheumatology (OMERACT) for acute and chronic gout studies (7). The experience of pain varied greatly between severe and milder flares. In the current study, mild flares, which were given ratings of between 2 to 4 on a 0–10-point pain scale, may not all have met the recently validated Gaffo definition of a flare, which requires a pain rating of at least 3 (8). This highlights the variability in pain intensity of a flare, which may not be comprehensively captured with a binary (present/absent) definition of a flare.

Physical disability, including difficulty walking and performing other activities of daily living, coupled with reliance on family members for assistance, were also commonly identified factors influencing the perceived severity of a flare. Feelings of depression, anxiety, and irritability also contributed to the overall severity of a flare. Previous research has shown associations between decreased physical and mental well-being and flare frequency (9). Flare frequency is also important to the patient perception of being in a state of low disease activity or remission (10). Given the intermittent nature of the flare experience and the complete resolution of symptoms between flares, the extent to which these factors contribute to the cumulative burden of flares over time would be of interest.

Sense of control was an important psychological factor contributing to the patient perception of flare severity. Knowledge and experience of previous flares, not present during a patient's first gout flare, meant that patients were able to initiate treatment to control the symptoms and prevent it from escalating into a more severe flare. These findings align with previous work, which has shown that patients who have a greater perceived understanding of the illness report more personal- and treatment-related control of the disease (11).

In this study, multiple domains contributed to the overall severity of a flare. These findings are consistent with previous work, which has shown that the experience of a gout flare is multidimensional with several interconnecting factors (4,12). However, it is unclear how much the overall severity of a flare is driven by pain alone. For example, severe flares were associated with greater pain intensity, which then impacted on patients' ability to function and therefore to attend work and undertake usual activities, which in turn impacted psychological health. Further work is warranted to determine the relative importance of factors that influence the overall severity of individual flares as well as the cumulative burden of flares over time.

This study has a number of strengths and limitations. First, the purposeful sampling method ensured that participants represented a wide range of demographic and clinical features of gout, which provides a diverse view of different patient experiences. However, participants were predominantly male, and although this reflects the sex differences in gout prevalence, this may reduce generalizability of the findings to female patients with gout. A further strength was the continuation of recruitment and analysis until theoretical saturation was reached, which provides confidence that a comprehensive understanding of the patient perspective was covered.

In conclusion, this qualitative study identified 4 key domains that together contribute to the overall severity of a gout flare from the patient perspective. In addition to flare characteristics, impact on function, psychological health, and family and social life all contribute to the severity of a gout flare. Measuring these domains in long-term studies assessing flare management or prevention, in addition to simply measuring reductions in flare frequency, would provide a more meaningful and accurate representation of cumulative flare burden from the patient perspective.

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. Stewart 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. Stewart, Taylor, Gaffo, Gott, Slark, Dalbeth.
Acquisition of data. Garcia-Guillen, Horne.
Analysis and interpretation of data. Garcia-Guillen, Stewart, Su, Taylor, Gaffo, Gott, Slark, Dalbeth.

REFERENCES

1. Tatlock S, Rudell K, Panter C, Arbuckle R, Harrold L, Taylor W, et al. What outcomes are important for gout patients? In-depth qualitative research into the gout patient experience to determine optimal endpoints for evaluating therapeutic interventions. Patient 2017;10:65–79.
2. Stewart S, Tallon A, Taylor W, Gaffo A, Dalbeth N. How flare prevention outcomes are reported in gout studies: a systematic review and content analysis of randomized controlled trials. Semin Arthritis Rheum 2020;50:303–13.
3. Stamp L, Morrillon M, Taylor W, Dalbeth N, Singh J, Lassere M, et al. Variability in the reporting of serum urate and flares in gout clinical trials: need for minimum reporting requirements. J Rheumatol 2018; 45:419–24.

4. Stewart S, Taylor W, Gaffo A, Slark A, Gott M, Dalbeth N. The experience of a gout flare: a meta-synthesis of qualitative studies. Semin Arthritis Rheum 2020;50:805–11.
5. Neogi T, Jansen TL, Dalbeth N, Fransen J, Schumacher HR, Berendsen D, et al. 2015 gout classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheumatol 2015;67:2557–68.
6. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3:77–101.
7. Schumacher H, Taylor W, Edwards L, Grainger R, Schlesinger N, Dalbeth N, et al. Outcome domains for studies of acute and chronic gout. J Rheumatol 2009;36:2342–5.
8. Gaffo AL, Dalbeth N, Saag KG, Singh JA, Rahn EJ, Mudano AS, et al. Validation of a definition of flare in patients with established gout. Arthritis Rheumatol 2018;70:462–7.
9. Khanna P, Nuki G, Bardin T, Tausche AK, Forsythe A, Goren A, et al. Tophi and frequent gout flares are associated with impairments to quality of life, productivity, and increased healthcare resource use: results from a cross-sectional survey. Health Qual Life Outcomes 2012:10:117–27.
10. Taylor W, Dalbeth N, Singh JA, Rahn EJ, Mudano AS, Chen Y, et al. Flare rate thresholds for patient assessment of gout disease activity states. J Rheumatol 2021;48:293–8.
11. Dalbeth N, Petrie KJ, House M, Chong J, Leung W, Chegudi R, et al. Illness perceptions in patients with gout and the relationship with progression of musculoskeletal disability. Arthritis Care Res (Hoboken) 2011;63:1605–12.
12. Chandratre P, Mallen C, Roddy E, Liddle J, Richardson J. "You want to get on with the rest of your life": a qualitative study of health-related quality of life in gout. Clin Rheumatol 2016;35:1197–205.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 2, February 2022, pp 324–335
DOI 10.1002/acr.24468

Characteristics of Patients With Antiphospholipid Antibody Positivity in the APS ACTION International Clinical Database and Repository

Ecem Sevim,[1] Diane Zisa,[2] Danieli Andrade,[3] Savino Sciascia,[4] Vittorio Pengo,[5] Maria G. Tektonidou,[6] Amaia Ugarte,[7] Maria Gerosa,[8] H. Michael Belmont,[9] Maria Angeles Aguirre Zamorano,[10] Paul R. Fortin,[11] Lanlan Ji,[12] Maria Efthymiou,[13] Hannah Cohen,[13] D. Ware Branch,[14] Guilherme Ramires de Jesus,[15] Laura Andreoli,[16] Michelle Petri,[17] Esther Rodriguez,[18] Ricard Cervera,[19] Jason S. Knight,[20] Tatsuya Atsumi,[21] Rohan Willis,[22] Robert Roubey,[23] Maria Laura Bertolaccini,[24] Doruk Erkan,[2] and Medha Barbhaiya,[2] on behalf of the APS ACTION Investigators

Objective. To describe the baseline characteristics of patients with positivity for antiphospholipid antibodies (aPLs) who were enrolled in an international registry, the Antiphospholipid Syndrome (APS) Alliance for Clinical Trials and International Networking (APS ACTION) clinical database and repository, overall and by clinical and laboratory subtypes.

Methods. The APS ACTION registry includes adults who persistently had positivity for aPLs. We evaluated baseline sociodemographic and aPL-related (APS classification criteria and "non-criteria") characteristics of patients overall and in subgroups (aPL-positive without APS, APS overall, thrombotic APS only, obstetric APS only, and both thrombotic APS/obstetric APS). We assessed baseline characteristics of patients tested for the presence of three aPLs (lupus anticoagulant [LAC] test, anticardiolipin antibody [aCL], and anti–β_2-glycoprotein I [anti-β_2GPI]) antibodies by aPL profiles (LAC only, single, double, and triple aPL positivity).

Results. The 804 aPL-positive patients assessed in the present study had a mean age of 45 $\pm$ 13 years, were 74% female, and 68% White; additionally, 36% had other systemic autoimmune diseases. Of these 804 aPL-positive patients, 80% were classified as having APS (with 55% having thrombotic APS, 9% obstetric APS, and 15% thrombotic APS/obstetric APS). In the overall cohort, 71% had vascular thrombosis, 50% with a history of pregnancy had obstetric morbidity, and 56% had experienced at least one non-criteria manifestation. Among those with three aPLs tested (n = 660), 42% were triple aPL–positive. While single–, double–, and triple aPL–positive subgroups had similar frequencies of vascular, obstetric, and non-criteria events, these events were lowest in the single aPL subgroup, which consisted of aCLs or anti-β_2GPI only.

Conclusion. Our study demonstrates the heterogeneity of aPL-related clinical manifestations and laboratory profiles in a multicenter international cohort. Within single aPL positivity, LAC may be a major contributor to clinical events. Future prospective analyses, using standardized core laboratory aPL tests, will help clarify aPL risk profiles and improve risk stratification.

INTRODUCTION

Antiphospholipid syndrome (APS) is characterized as an autoimmune disease marked by thromboses and/or pregnancy morbidity with persistent positivity for antiphospholipid antibodies (aPLs), lupus anticoagulant (LAC) test, anticardiolipin antibodies (aCLs), and/or anti–β_2-glycoprotein I (anti-β_2GPI) antibodies, as defined by the revised Sapporo criteria for APS (1,2). Other well-recognized "non-criteria" clinical manifestations may occur in aPL-positive patients, including thrombocytopenia, autoimmune

Dr. Fortin is recipient of a Tier 1 Canada Research Chair on systemic autoimmune rheumatic diseases. The Hopkins Lupus Cohort is supported by NIH grant RO1-AR-069572 to Dr. Petri. Dr. Barbhaiya's work is supported by a Rheumatology Research Foundation Investigator Award. The APS ACTION Registry was created using REDCAP, which was provided by the Clinical and Translational Science Center at Weill Cornell Medical College (CTSC grant UL1-TR000457).

[1]Ecem Sevim, MD: Hospital for Special Surgery, New York, New York, and Montefiore Medical Center, Bronx, New York; [2]Diane Zisa, MD, Doruk Erkan, MD, MPH, Medha Barbhaiya, MD, MPH: Hospital for Special Surgery and Weill Cornell Medicine, New York, New York; [3]Danieli Andrade, MD, PhD: University of Sao Paulo, Sao Paulo, Brazil; [4]Savino Sciascia, MD, PhD: University of Turin, Turin, Italy; [5]Vittorio Pengo, MD: University Hospital Padova, Padova, Italy; [6]Maria G. Tektonidou, MD: National and Kapodistrian University of Athens,

SIGNIFICANCE & INNOVATIONS

- Using the multicenter, international Antiphospholipid Syndrome (APS) Alliance for Clinical Trials and International Networking (ACTION) registry, we described baseline clinical and laboratory characteristics of patients with persistent positivity for antiphospholipid antibodies (aPLs), including 36% of patients with other systemic autoimmune disease included in this registry.
- One-fifth of the registry patients did not fulfill clinical classification criteria for clinical APS. Among the registry patients, 71% experienced vascular events, 25% had aPL-related obstetric morbidity, and 56% had at least one other non-criteria clinical aPL manifestation, most commonly thrombocytopenia and white matter lesions of the central nervous system.
- Although single–, double–, and triple–aPL–subgroups had similar frequencies of vascular, pregnancy, and non-criteria events, these events were less common in the single aPL subgroup after excluding LAC-positive patients, suggesting the importance of LAC positivity in APS.
- Future prospective analyses, using standardized core laboratory testing for aPLs, will help clarify risk profiles for individuals with aPL positivity.

hemolytic anemia, livedo reticularis, aPL-associated nephropathy, cardiac valve disease, cognitive dysfunction, and skin ulcers (1,3). APS can occur in isolation (primary APS) or in association with other autoimmune diseases, most notably systemic lupus erythematosus (SLE) (4).

Patients with positivity for aPLs can have heterogeneous clinical manifestations, including asymptomatic aPL positivity (no thrombosis or pregnancy morbidity), thrombotic APS (which is characterized by venous, arterial, or microvascular involvement), and obstetric APS (which is characterized by pregnancy complications such as fetal loss, recurrent early miscarriages, placental insufficiency, or preeclampsia). Furthermore, not every positive result for aPL testing is clinically significant, and transient low titer aPL positivity may occur in settings of infection or malignancy (5,6). Despite accumulating data showing an important role for aPL laboratory profiles in APS assessment (7–9), the risk of aPL-related clinical events by aPL laboratory profile remains under investigation. Few large cohorts have estimated the prevalence of aPL-related clinical manifestations (10–12). Furthermore, the distribution of demographic and clinical factors by aPL-related clinical subtypes or laboratory profiles is not well-established.

The Antiphospholipid Syndrome Alliance for Clinical Trials and International Networking (APS ACTION) is an international network established in 2010 to conduct large-scale multicenter studies and clinical trials in persistently aPL-positive patients (2). The APS ACTION clinical database and repository ("registry") was created to study the natural course of persistently aPL-positive patients with or without autoimmune disorders. See Appendix A for APS ACTION registry investigators and their locations. In the present study, our primary objective was to retrospectively evaluate the baseline demographic and clinical characteristics of aPL-positive patients enrolled in the APS ACTION registry since 2010, overall and by clinical subtype (aPL positive without APS classification, thrombotic APS, and obstetric APS). Secondly, we also assessed the clinical characteristics of aPL-positive patients who were tested at baseline for all three "criteria" aPLs (LAC, aCLs, and anti-β_2GPI antibodies), categorized by aPL profile (LAC positivity only and single, double, and triple aPL positivity).

PATIENTS AND METHODS

APS ACTION registry and data collection. The inclusion criteria for the APS ACTION Registry were the following: individuals ages 18 to 60 years with persistent positivity for aPLs (at least 12 weeks apart) according to the revised Sapporo criteria (1), within 12 months prior to screening. Patients referred to APS ACTION sites were referred from hospital or outpatient settings and had received aPL testing for a variety of reasons such as thrombosis, pregnancy morbidity, false-positive serologic test for syphilis, prolonged activated partial thromboplastin time, thrombocytopenia, or concomitant systemic autoimmune diseases. As part of the registry entry criteria, patients must have had persistent aPL positivity prior to registry entry. Positivity for aCLs and/or anti-β_2GPI antibodies was defined as an individual having IgG, IgM, or IgA titers of ≥40 units/ml (medium-to-high titers). LAC activity was detected by coagulation assays according to the

Athens, Greece; [7]Amaia Ugarte, MD: Hospital Universitario Cruces, Barakaldo, País Vasco, Spain; [8]Maria Gerosa, MD, PhD: University of Milan, Milan, Italy; [9]H. Michael Belmont, MD: New York University Langone Medical Center, New York, New York; [10]Maria Angeles Aguirre Zamorano, MD, PhD: Maimonides Institute for Biomedical Research of Cordoba, Cordoba, Spain; [11]Paul R. Fortin, MD, PhD: CHU de Québec and Université Laval, Quebec City, Canada; [12]Lanlan Ji, MD: Peking University First Hospital, Beijing, China; [13]Maria Efthymiou, PhD, Hannah Cohen, MD: University College London, London, UK; [14]D. Ware Branch, MD: University of Utah and Intermountain Healthcare, Salt Lake City, Utah; [15]Guilherme Ramires de Jesus, MD, PhD: Rio de Janeiro State University, Rio de Janeiro, Brazil; [16]Laura Andreoli, MD, PhD: University of Brescia, Brescia, Italy; [17]Michelle Petri, MD, MPH: Johns Hopkins University School of Medicine, Baltimore, Maryland; [18]Esther Rodriguez, MD: Hospital Universitario 12 de Octubre, Madrid, Spain; [19]Ricard Cervera, MD, PhD, FRCP: Hospital Clínic Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Catalonia, Spain; [20]Jason S. Knight, MD, PhD: University of Michigan, Ann Arbor; [21]Tatsuya Atsumi, MD, PhD: Hokkaido University Hospital, Sapporo, Japan; [22]Rohan Willis, MD: University of Texas Medical Branch, Galveston; [23]Robert Roubey, MD: University of North Carolina, Chapel Hill; [24]Maria Laura Bertolaccini, MD, PhD: St Thomas' Hospital, London, UK.

No potential conflicts of interest relevant to this article were reported.

Address correspondence to Medha Barbhaiya, MD, MPH, Hospital For Special Surgery, 535 East 70th Street, New York, NY 10021. Email: barbhaiyam@hss.edu.

Submitted for publication January 24, 2020; accepted in revised form September 22, 2020.

International Society on Thrombosis and Hemostasis guidelines for lupus anticoagulant detection (13).

An international web-based application, REDCap, was used to store and manage data on baseline sociodemographic information, aPL-related clinical events, pregnancy history, medications, and laboratory profile (14). Blood samples were also collected at registry entry for confirmation of aPL positivity. Patients were followed every 12 ± 3 months with clinical data and blood collection, or at the time of a new aPL-related thrombosis and/or pregnancy morbidity.

Study cohort. All participants who were persistently positive for aPLs and who were enrolled in the APS ACTION registry between May 2010 and March 2019 were included in the study cohort. We categorized patients into two groups by clinical subtype at baseline—1) patients who had positivity for aPLs without APS and who met laboratory criteria for APS classification, but not the clinical revised Sapporo criteria (1) and 2) patients with overall APS who met both laboratory and clinical criteria for definite APS. Patients with overall APS were further categorized into three mutually exclusive groups as follows: 1) patients with thrombotic APS, which is defined by a history of any vascular event (including any arterial thrombosis, venous thrombosis, or microvascular involvement, but excluding only superficial vascular thrombosis); 2) patients with obstetric APS, which is defined by a history of any pregnancy morbidity event (defined by the revised Sapporo classification criteria [1]); and 3) patients with thrombotic APS/obstetric APS, who were defined as individuals who have experienced any vascular thrombosis event and any pregnancy morbidity event (Table 1).

After categorizing patients into subgroups, we evaluated the baseline laboratory profiles of aPL-positive patients in the registry. We assessed the baseline clinical characteristics of aPL-positive patients with different laboratory profiles (single, double, and triple aPL positivity) among patients tested for all three aPLs (LAC, aCLs, and anti-β_2GPI antibodies). We also subcategorized the subgroup with single aPL positivity by separately evaluating those with LAC only and those with single aPL positivity excluding LAC (Table 2). For the purposes of this study, positivity for aCL IgG, IgM, and IgA and anti-β_2GPI IgG, IgM, and IgA was defined as a patient having a titer of ≥40 units, with the highest titer among all test results taken into consideration during analysis.

Data collection for baseline characteristics. Demographic characteristics collected included mean age, race (White, Latin American Mestizo, Asian, Black, or "Other"), ethnicity (Non–Latin American or Latin American [for the US, Canada, and Europe], Afro-descendent, Mestizo, or Caucasian [for South America], Afro-descendent [for South Africa], or "Other"), and region of residence (Europe, North America [for the US and Canada], Latin America, and Asia-Pacific). Clinical manifestations were subgrouped into vascular events (arterial thrombosis, venous thrombosis, microvascular involvement), catastrophic APS (CAPS), pregnancy morbidity, and "other." Other clinical manifestations included livedo reticularis/racemosa, persistent thrombocytopenia defined as a platelet count of <100,000 per microliter tested twice at least 12 weeks apart, autoimmune hemolytic anemia, echocardiography-proven cardiac valve disease, aPL-related nephropathy, skin ulcers, chorea, seizure disorder, radiographic white matter lesions (only identified in those patients who had magnetic resonance imaging performed), and neuropsychiatric test–proven cognitive dysfunction (Supplementary Table 1, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24468/abstract). CAPS was defined as "definite" or "probable" based on the international consensus statement on classification criteria and treatment guidelines for CAPS (15). Past and current medications, including aspirin, warfarin, low molecular weight heparin, direct oral anticoagulants, glucocorticoids, hydroxychloroquine, intravenous immunoglobulin, rituximab, azathioprine, cyclophosphamide, cyclosporine, methotrexate, and mycophenolate mofetil, were collected at the time of registry entry.

Study design and statistical analysis. Data from the APS ACTION registry were locked in March 2019. First, we evaluated the baseline demographic and clinical characteristics of aPL-positive patients overall, and by clinical subtype: aPLs without APS, APS (overall), obstetric APS, thrombotic APS, and thrombotic APS/obstetric APS. We also classified aPL-positive patients (overall and by aPL-related clinical subtypes) as having primary aPL/APS or aPL/APS with other systemic autoimmune disease, including SLE, rheumatoid arthritis, mixed connective tissue disease, Sjögren's syndrome, systemic sclerosis, inflammatory muscle disease, and vasculitis.

Second, we assessed the clinical characteristics of aPL-positive patients with different baseline laboratory profiles (LAC positivity only, single aPL positivity, single aPL positivity after excluding LAC positivity, double aPL positivity, and triple aPL positivity), among patients tested for all three aPLs. Descriptive statistics were used to describe continuous variables (mean ± SD, minimum, median, and maximum).

RESULTS

Baseline characteristics in overall cohort. As of March 2019, 804 patients who were persistently positive for aPLs were enrolled from 26 centers worldwide (mean age of 45 ± 13 years at study entry, with 594 [74%] of patients being female, 546 [68%] being White, 87 [11%] being Latin American Mestizos, 387 [48%] from Europe, and 232 [29%] from North America). Table 1 shows the baseline demographic and clinical characteristics of aPL-positive patients at registry entry, both overall and by clinical subtype. In the study cohort, 642 (80%) of patients met the clinical criteria for definite APS; among these

Table 1. Baseline demographic and clinical characteristics of patients with positivity for aPLs between different groups of aPL-positive patients included in the APS ACTION registry according to aPL-related clinical phenotype (2010–2019)*

	All patients, (N = 804)	Patients with aPL positivity without APS, 162 (20)	Patients with APS (overall), 642 (80)	Patients with OAPS only, 74 (9)	Patients with TAPS only, 446 (55)	Patients with TAPS + OAPS, 122 (15)
Primary aPL/APS	516 (64)	89 (55)	427 (67)	55 (74)	295 (66)	77 (63)
Concomitant systemic autoimmune disease†	288 (36)	73 (45)	215 (33)	19 (26)	151 (34)	45 (37)
SLE	242 (30)	60 (37)	182 (28)	15 (20)	129 (29)	38 (31)
Sociodemographic characteristics						
Age at registry entry, mean ± SD years	45.12 ± 13	43.80 ± 13	45.45 ± 13	41.47 ± 11	46.69 ± 14	43.34 ± 12
Female sex	594 (74)	127 (78)	467 (73)	74 (100)	271 (61)	122 (100)
Race‡						
White	546 (68)	118 (73)	428 (67)	52 (70)	305 (68)	71 (58)
Latin American Mestizos	87 (11)	6 (4)	81 (13)	6 (8)	47 (11)	28 (23)
Asian	56 (7)	17 (10)	39 (6)	8 (11)	24 (5)	7 (8)
Black	26 (3)	7 (4)	19 (3)	2 (3)	12 (3)	5 (4)
American Indian or Alaskan	2	1 (1)	1	0	1	0
Native American	0	0	0	0	0	0
Reported as "other"§	14 (2)	2 (1)	12 (2)	1 (1)	9 (2)	2 (2)
Ethnicity¶						
US, Canada, and Europe	377 (47)	92 (57)	285 (44)	43 (58)	201 (45)	41 (34)
Non-Hispanic	356 (44)	88 (54)	268 (42)	38 (51)	194 (43)	36 (30)
Hispanic	21 (3)	4 (2)	17 (3)	5 (7)	7 (2)	5 (4)
South America	137 (17)	8 (5)	129 (20)	8 (11)	82 (18)	39 (32)
Mestizos	72 (9)	2 (1)	70 (11)	4 (5)	42 (9)	24 (20)
Caucasian	47 (6)	4 (2)	43 (7)	2 (3)	31 (7)	10 (8)
African descendent	18 (2)	2 (1)	16 (2)	2 (3)	9 (2)	5 (4)
Other#	135 (17)	35 (22)	100 (16)	16 (22)	65 (15)	19 (16)
Australia	4 (1)	0	4 (1)	0	2	2 (2)
Not Aboriginal	4 (1)	0	4 (1)	0	2	2 (2)
Aboriginal	0	0	0	0	0	0
Region of residence						
Europe	387 (48)	84 (52)	303 (47)	37 (50)	221 (50)	45 (37)
North America	232 (29)	60 (37)	172 (27)	23 (31)	117 (26)	32 (26)
US	201 (25)	56 (35)	145 (23)	21 (28)	95 (21)	29 (24)
Canada	31 (4)	4 (2)	27 (4)	2 (3)	22 (5)	3 (4)
Latin America	131 (16)	6 (4)	125 (19)	7 (9)	83 (19)	35 (29)
Asia Pacific	54 (7)	12 (7)	42 (7)	7 (9)	25 (6)	10 (8)
Clinical manifestations						
Any vascular event	568 (71)	0	568 (71)	0	446 (100)	122 (100)
Any arterial thrombosis	300 (37)	0	300 (37)	0	239 (54)	61 (50)
Stroke	165 (21)	0	165 (26)	0	127 (28)	38 (31)
Transient ischemic attacks	69 (9)	0	69 (11)	0	50 (11)	19 (16)
Myocardial infarction	31 (4)	0	31 (5)	0	29 (7)	2 (2)
Intracardiac thrombus	3	0	3	0	2	1 (1)
Peripheral artery**	30 (4)	0	30 (5)	0	27 (6)	3 (3)
Visceral	10 (1)	0	10 (2)	0	9 (2)	1 (1)
Retinal	5 (1)	0	5 (1)	0	3 (1)	2 (2)
Any venous thrombosis	347 (43)	0	347 (54)	0	269 (60)	78 (64)
Central venous sinus	13 (2)	0	13 (2)	0	12 (3)	1 (1)
Pulmonary embolism	76 (9)	0	76 (12)	0	64 (14)	12 (10)
Upper extremity	7 (1)	0	7 (1)	0	7 (2)	0
Lower extremity	217 (27)	0	217 (34)	0	177 (40)	40 (33)
Visceral	8 (1)	0	8 (1)	0	4 (1)	4 (3)
Retinal	6 (1)	0	6 (1)	0	5 (1)	1
Any microvascular involvement	93 (12)	3 (2)	90 (14)	2 (3)	67 (15)	21 (17)
Biopsy-proven	32 (4)	0	32 (5)	0	26 (6)	6 (5)
Kidney	15 (2)	0	15 (2)	0	11 (2)	4 (3)
Skin	9 (1)	0	9 (1)	0	9 (2)	0
Pulmonary	3	0	3	0	3 (1)	0
Other	5 (1)	0	5 (1)	0	3 (1)	2 (2)

(Continued)

Table 1. (*Cont'd*)

	All patients, (N = 804)	Patients with aPL positivity without APS, 162 (20)	Patients with APS (overall), 642 (80)	Patients with OAPS only, 74 (9)	Patients with TAPS only, 446 (55)	Patients with TAPS + OAPS, 122 (15)
Clinical suspicion, no biopsy	61 (8)	3 (2)	58 (9)	2 (3)	41 (9)	15 (12)
Kidney	14 (2)	0	14 (2)	2 (3)	10 (2)	2 (2)
Skin	37 (5)	3 (2)	34 (5)	0	24 (5)	10 (8)
Pulmonary	2	0	2	0	2	0
Other	8 (1)	0	8 (1)	0	5 (1)	3 (2)
Both arterial and venous thrombosis	92 (11)	0	92 (14)	0	72 (16)	20 (16)
Recurrent vascular events (arterial and/or venous)	225 (28)	0	225 (35)	0	173 (39)	52 (43)
Catastrophic APS††	9 (1)	0	9 (1)	0	7 (2)	2 (2)
History of pregnancy	393/594 (66)	70/127 (55)	323 (50)	74 (100)	127/271 (47)	122 (100)
Pregnancy morbidity	196/393 (50)	0	196/323 (61)	74 (100)	0	122 (100)
Unexplained fetal death at 10 weeks of gestation or later	136/196 (69)	0	136/196 (69)	51/74 (69)	0	85/122 (70)
Premature birth prior to 34 weeks of gestation due to eclampsia, preeclampsia, or placental insufficiency	68/196 (35)	0	68/196 (35)	25/74 (34)	0	43/122 (35)
≥3 unexplained spontaneous abortions prior to 10 weeks of gestation	34/196 (17)	0	34/196 (17)	11/74 (15)	0	23/122 (19)
3 consecutive unexplained spontaneous abortions prior to 10 weeks of gestation	29/196 (15)	0	29/196 (15)	9/74 (12)	0	20/122 (16)
Other clinical manifestations‡‡	451 (56)	76 (47)	375 (58)	30 (41)	264 (59)	81 (66)
Livedo reticularis/racemosa	100 (12)	10 (6)	90 (14)	8 (11)	56 (13)	26 (21)
Persistent thrombocytopenia (platelet count <100,000/μl)	151 (19)	32 (20)	119 (19)	14 (19)	75 (17)	30 (25)
Autoimmune hemolytic anemia	40 (5)	9 (6)	31 (5)	4 (5)	22 (5)	5 (4)
Cardiac valve disease	65/688 (9)	10/142 (7)	56/546 (10)	2/52 (4)	34/391 (9)	20/103 (19)
Skin ulcer	50 (5)	3 (2)	47 (6)	0	36 (6)	11 (7)
aPL-associated nephropathy	29/755 (4)	0/156 (0)	29/599 (5)	2/69 (3)	21/414 (5)	6/116 (5)
Neurologic presentations						
Cognitive dysfunction	85 (11)	11 (7)	74 (12)	3 (4)	53 (12)	18 (15)
MS-like disease	6 (1)	1 (1)	5 (1)	0	5 (1)	0
Chorea	13 (2)	2 (1)	11 (2)	0	7 (2)	4 (3)
Seizure disorder	67 (8)	8 (5)	59 (9)	3 (4)	42 (9)	14 (11)
White matter lesions	136/549 (25)	17/103 (17)	119/446 (27)	6/35 (17)	90/326 (28)	23/85 (27)
Medications (registry entry)						
Any anticoagulation therapy	497 (62)	18 (11)	479 (75)	9 (12)	372 (83)	98 (80)
Warfarin	434 (54)	13 (8)	421 (66)	4 (5)	328 (74)	89 (73)
LMWH	48 (6)	4 (2)	44 (7)	5 (7)	30 (7)	9 (7)
Factor Xa inhibitor	28 (3)	1 (1)	27 (4)	0	26 (6)	1 (1)
Thrombin inhibitor	0	0	0	0	0	0
Acetylsalicylic acid (aspirin)	366 (46)	108 (67)	258 (40)	52 (70)	168 (38)	38 (31)
Clopidogrel	29 (4)	2 (1)	27 (4)	1 (1)	21 (5)	5 (4)
Hydroxychloroquine	364 (45)	90 (56)	274 (43)	31 (44)	189 (42)	54 (44)
Statins	191 (24)	16 (10)	175 (27)	9 (12)	143 (32)	23 (19)
ACE inhibitor/ARB	163 (20)	21 (13)	142 (22)	9 (12)	106 (24)	27 (22)
Intravenous immunoglobulin	5 (1)	0	5 (1)	0	5 (1)	0
Plasma exchange	1	0	1	0	0	1 (1)
Rituximab	16 (2)	3 (2)	13 (2)	0	12 (3)	1 (1)
Other immunosuppression§§	202 (25)	42 (26)	160 (25)	10 (14)	116 (26)	34 (28)
No medications	28 (3)	16 (10)	12 (2)	9 (12)	2	1 (1)
Medications (ever)						
Any anticoagulation therapy	763 (95)	37 (23)	566 (88)	43 (58)	407 (91)	116 (95)
Warfarin	526 (65)	20 (12)	506 (79)	10 (14)	388 (87)	108 (89)
LMWH	340 (42)	21 (13)	319 (50)	41 (55)	200 (45)	78 (64)
Factor Xa inhibitor	43 (5)	2 (1)	41 (6)	1 (1)	37 (8)	3 (3)
Thrombin inhibitor	4 (1)	0	0	0	4 (1)	0
Acetylsalicylic acid (aspirin)	516 (64)	121 (75)	395 (62)	63 (85)	250 (56)	82 (67)
Clopidogrel	50 (6)	3 (2)	47 (7)	1 (1)	38 (9)	8 (7)
Hydroxychloroquine	428 (53)	101 (62)	327 (51)	34 (46)	223 (50)	70 (57)

(Continued)

Table 1. (*Cont'd*)

	All patients, (N = 804)	Patients with aPL positivity without APS, 162 (20)	Patients with APS (overall), 642 (80)	Patients with OAPS only, 74 (9)	Patients with TAPS only, 446 (55)	Patients with TAPS + OAPS, 122 (15)
Statins	210 (26)	20 (12)	190 (30)	9 (12)	153 (34)	28 (23)
ACE inhibitor/ARB	192 (24)	23 (14)	169 (26)	11 (15)	121 (27)	37 (30)
Intravenous immunoglobulin	57 (7)	11 (7)	46 (7)	4 (5)	35 (8)	7 (6)
Plasma exchange	16 (2)	1 (1)	15 (2)	2 (3)	8 (2)	5 (4)
Rituximab	48 (6)	9 (6)	39 (6)	1(1)	34 (8)	4 (3)
Other immunosuppression§§	297 (37)	57 (35)	240 (37)	18 (24)	174 (39)	48 (39)
No medications	11 (1)	9 (6)	2	2	0	0

* Except where indicated otherwise, values are the number (%) of patients. Missing data and other categories are not included. ACE = angiotensin-converting enzyme; aPLs = antiphospholipid antibodies; APS = antiphospholipid syndrome; APS ACTION = APS Alliance for Clinical Trials and International Networking; ARB = angiotensin receptor blocker; LMWH = low molecular weight heparin; MS = multiple sclerosis; OAPS = obstetric APS; SLE = systemic lupus erythematosus; TAPS = thrombotic APS.
† Systemic autoimmune diseases included SLE, rheumatoid arthritis, mixed connective tissue disease, Sjögren's syndrome, systemic sclerosis, inflammatory muscle disease, and vasculitis.
‡ Races were collected in a total of 731 patients (162 patients with aPL positivity only, 69 patients with OAPS, 403 patients with TAPS, and 97 patients with both TAPS and OAPS). Latin American Mestizo refers to a person of combined European and Indigenous American descent.
§ Includes American Indian or Alaskan, Native Hawaiian or Pacific Islander, and other unspecified races as indicated by the patient.
¶ Ethnicities were collected in a total of 653 patients (146 patients with aPL positivity only, 65 patients with OAPS, 354 patients with TAPS, and 88 patients with both TAPS and OAPS).
# Other unspecified ethnicities as indicated by the patient.
** Consists of the arteries not in the chest or abdomen (i.e., in the arms, hands, legs, and feet).
†† Catastrophic APS was diagnosed as "probable" or "definite" based on the international consensus statement on classification criteria and treatment guidelines for catastrophic APS (15).
‡‡ Livedo reticularis/racemosa, persistent thrombocytopenia, and autoimmune hemolytic anemia, which patients were considered as "ever" or "never" having had these conditions at the time of registry entry.
§§ Other immunosuppression treatment included azathioprine, glucocorticoids, cyclophosphamide, cyclosporine, methotrexate, mycophenolate mofetil, and other therapies.

patients, 74 (12%) had obstetric APS only, 446 (69%) had thrombotic APS only, and 122 (19%) had both obstetric APS and thrombotic APS. One-hundred sixty-two patients (20%) did not meet the clinical criteria for definite APS; among these patients who have aPL positivity without APS, 76 (47%) had one or more other (non-criteria) clinical manifestations associated with aPLs, and 86 (53%) were asymptomatic. Thirty-six percent of the overall cohort had at least one concomitant systemic autoimmune disease, with 30% having SLE, 2% having Sjögren's syndrome, 2% having mixed connective tissue disease, 1% having rheumatoid arthritis, 1% having vasculitis, 1% having systemic sclerosis, and 4% having other systemic autoimmune diseases. The frequency of systemic autoimmune diseases was slightly higher in the group that had aPL positivity without APS as compared to APS patients (45% and 33%, respectively).

Among the 804 registry participants, 568 (71%) experienced at least one vascular event (arterial thrombosis, venous thrombosis, or microvascular involvement), and 28% experienced recurrent vascular events. Venous thrombosis occurred more frequently than arterial thrombosis (43% versus 37%) in the overall cohort, with both types of thrombosis appearing in 11% of the cohort; 12% of the cohort had microvascular involvement, and 1% had CAPS. Among those with arterial thrombosis, strokes (21%) occurred much more frequently than cardiac events (4%); events in the lower extremities were the most common type of venous thrombosis (27%) recorded. Of the 393 women in the registry who had a history of pregnancy, 50% had experienced a pregnancy morbidity event, most commonly due to unexplained fetal death at ≥10 weeks of gestation (69%). Over half (56%) of the overall cohort had at least one non-criteria manifestation; among these, the most common were white matter lesions of the central nervous system and persistent thrombocytopenia.

In terms of medications that were being used at the time of registry entry, 62% of aPL-positive patients were receiving anticoagulation, with 54% receiving warfarin, 6% receiving low molecular weight heparin, and 3% receiving factor Xa inhibitor. Other commonly used medications were aspirin (46%), hydroxychloroquine (45%), and statins (24%).

Baseline characteristics by clinical subtype. When comparing characteristics by aPL-related clinical subtypes (Table 1), the incidence of concomitant systemic autoimmune disease was highest in patients with aPL positivity without APS (45%) and lowest in patients with obstetric APS (26%). A similar pattern was reflected in aPL-positive patients with concomitant SLE specifically (37% in patients who had aPL positivity without APS versus 20% in patients with obstetric APS). The mean age was lowest among patients with obstetric APS (41.47 ± 11 years) and highest among patients with thrombotic APS (46.69 ± 14 years). The majority of patients in each clinical subtype were White, whereas there were very few Black patients in each clinical subtype; the highest percentage of Latin American

Table 2. Clinical characteristics of patients with aPL positivity in the APS ACTION registry (2010-2019) who were tested for 3 aPL, categorized according to aPL profile (N = 660)*

	LAC only positivity, 168 (25)	Positivity for any single aPL (including LAC only)†, 215 (32)	Single aPL positivity (excluding LAC only), 47 (7)	Double aPL positivity†, 167 (25)	Triple aPL positivity†, 278 (42)
Any vascular events	127 (73)	148 (67)	21 (45)	118 (68)	195 (70)
Arterial thrombosis	61 (36)	73 (34)	12 (26)	68 (41)	96 (35)
Venous thrombosis	81 (48)	92 (43)	11 (23)	66 (40)	132 (47)
Microvascular thrombosis	10 (6)	12 (6)	2 (4)	9 (5)	16 (6)
Transient ischemic attacks	11 (7)	13 (6)	2 (4)	19 (11)	20 (7)
Any pregnancy morbidity	43/84 (51)	53/108 (49)	10/24 (42)	41/86 (48)	62/117 (53)
>1 fetal death at 10 weeks of gestation or later	31 (51)	39 (52)	8 (57)	29 (53)	43 (56)
>1 preterm delivery prior to 34 weeks of gestation	12 (20)	13 (17)	1 (7)	14 (25)	29 (38)
≥3 pre-embryonic/ embryotic losses prior to 10 weeks of gestation	9 (15)	14 (19)	5 (36)	6 (11)	6 (8)
Any other clinical manifestation	93 (55)	107 (50)	14 (30)	104 (62)	158 (57)
Livedo reticularis/racemosa	27 (16)	28 (13)	1 (2)	24 (14)	32 (12)
Persistent thrombocytopenia†	26 (15)	30 (14)	4 (9)	29 (17)	70 (25)
Hemolytic anemia‡	9 (5)	10 (5)	1 (2)	7 (4)	16 (6)
Cardiac valve disease	11/146 (8)	12/188 (6)	1/42 (2)	12/142 (8)	32/234 (14)
Skin ulcers	6 (4)	7 (3)	1 (2)	7 (4)	10 (4)
aPL-associated nephropathy	4/157 (3)	4/201 (2)	0	5/160 (3)	11/256 (4)
Cognitive dysfunction	14 (8)	17 (8)	3 (6)	20 (12)	33 (12)
Multipe sclerosis–like disease	2 (1)	3 (1)	1 (2)	3 (2)	0
Chorea	2 (1)	2 (1)	0	4 (2)	6 (2)
Seizure	17 (10)	21 (10)	4 (9)	13 (8)	23 (8)
White matter lesions	33/120 (28)	40/155 (26)	7/35 (20)	33/116 (28)	45/190 (24)

* Except where indicated otherwise, values are the number (%) of patients. Patients in this analysis were tested for three aPLs (LAC, aCL, and anti-β_2GPI antibodies). An additional 8 patients were tested for these three aPLs but had low titers (20–39 units) on enzyme-linked immunosorbent assay with negative LAC test, and were thus excluded from the analysis. Single aPL positivity was defined as positivity on one of the three aPL tests for aPLs, double aPL positivity was defined as positivity on two of three tests, and triple aPL positivity was defined as positivity on all three tests. All groups except LAC only and single aPL positivity were mutually exclusive. Only 5 (0.8%) of 660 patients had isolated aCL/anti-β_2GPI IgA positivity with negative results for LAC and aCL/anti-β_2GPI IgG and IgM. In total, 6 patients had catastrophic APS, distributed between double aPL– and triple aPL–positive groups. aCL = anticardiolipin antibody; anti-β2GPI = anti–β_2-glycoprotein I; aPL = antiphospholipid antibody; APS = antiphospholipid syndrome; APS ACTION = APS Alliance for Clinical Trials and International Networking; LAC = lupus anticoagulant.
† Cutoff value for aCL positivity and anti-β_2GPI IgG, IgM, and IgA positivity was defined as a patient having a titer of at least 40 units upon antibody testing.
‡ Defined as a platelet count of <100,000 per microliter tested twice at least 12 weeks apart.
§ Defined as anemia in the presence of antibodies directed against red blood cells, evidenced by either direct or indirect Coombs' tests.

Mestizo patients were in the thrombotic APS/obstetric APS group (Table 1). Approximately 50% of patients in each clinical subtype were recruited from Europe, except in individuals with thrombotic APS/obstetric APS, which occurred less frequently in European patients (37%). Approximately 30% of thrombotic APS/obstetric APS patients were recruited from Latin America, which was the most common clinical subtype observed in recruited patients from this region.

Within the thrombotic APS group compared to the thrombotic APS/obstetric APS groups, while we observed similar frequencies of arterial thrombotic events (54% and 50%, respectively) and venous thrombotic events (60% and 64%, respectively) within the these subgroups, lower extremity venous thrombosis events were slightly higher among patients with thrombotic APS only compared to patients with both thrombotic and obstetric APS (40% and 33%, respectively) (Table 1). Between the thrombotic APS group and the thrombotic APS/obstetric APS group, we also observed a similar rate of microvascular involvement (15% and 17%, respectively) and catastrophic APS (2% each). When comparing pregnancy morbidity in patients with obstetric APS to that in patients with thrombotic APS/obstetric APS, we found similar frequencies of unexplained death of the fetus at 10 weeks of gestation or later (69% versus 70%, respectively), premature birth occuring earlier than 34 weeks of gestation due to eclampsia, pre-eclampsia or placental insufficiency (34% versus 35%, respectively), and at least 3 unexplained spontaneous abortions occuring prior to 10 weeks of gestation (15% versus 19%, respectively).

Compared to the overall cohort, patients who had aPL positivity without APS had a slightly lower rate of other clinical manifestations (47% versus 58%). Other clinical manifestations were highest in the thrombotic APS/obstetric APS group (66%) and lowest in the obstetric APS group (41%). In particular, patients with thrombotic APS/obstetric APS had substantially higher rates of livedo reticularis/racemosa, persistent thrombocytopenia, cardiac valve disease, skin ulcer, and cognitive dysfunction compared to patients with other subtypes of APS and the overall study cohort (Table 1). Thrombotic APS patients also had a higher rate of other clinical manifestations (59%) compared to obstetric APS patients (41%).

Compared to APS patients at the time of registry entry, patients who had aPL positivity without APS had higher rates of current use of aspirin (67% versus 40%) and hydroxychloroquine (56% versus 43%) and lower rates of anticoagulation, statin, and antihypertensive use. Aspirin use was highest in patients with a history of obstetric APS (70%) compared to patients with thrombotic APS (38%) or those with both thrombotic and obstetric APS (31%). A majority of APS patients had been receiving anticoagulation therapy with warfarin (66%) at the time of registry entry. Current use of any anticoagulation therapy (warfarin, low molecular weight heparin, factor Xa inhibitor, thrombin inhibitor) at registry entry was highest among patients with thrombotic APS (83%) compared to patients with aPL positivity without APS (11%). A similar pattern of medication use was observed for "ever" use at the time of registry entry among the aPL-related clinical subgroups (Table 1).

Baseline characteristics by aPL profile. Of the 804 aPL-positive patients, 660 (83%) were tested for all three aPLs (LAC, aCLs, and anti-β_2GPI antibodies), and 42% had triple positivity for aPLs. We excluded eight patients who were tested for three aPLs but had low titers (20–39 units) of aPLs measured by enzyme-linked immunosorbent assay with negative LAC test. Approximately one-fifth (17%) of patients were missing at least one aPL test; in this group, the proportion with single positivity was similar to that with double positivity (50% and 46%, respectively). Among those without testing for the three aPLs and with single positivity only (50%), LAC positivity was most common (37%); the combination of LAC plus aCL positivity was more common than aCL plus anti-β_2GPI antibody positivity in those with double aPL positivity (Figure 1).

While similar frequencies of vascular thrombosis, pregnancy morbidity, and other clinical manifestations were observed across subgroups with single, double, and triple aPL positivity, the subgroup with single aPL positivity (excluding the patients with only LAC positivity) had substantially lower frequencies of all three event types (Table 2). Compared to the other aPL profile subgroups, triple aPL positivity had the highest proportion of patients with at least one preterm delivery before 34 weeks of gestation,

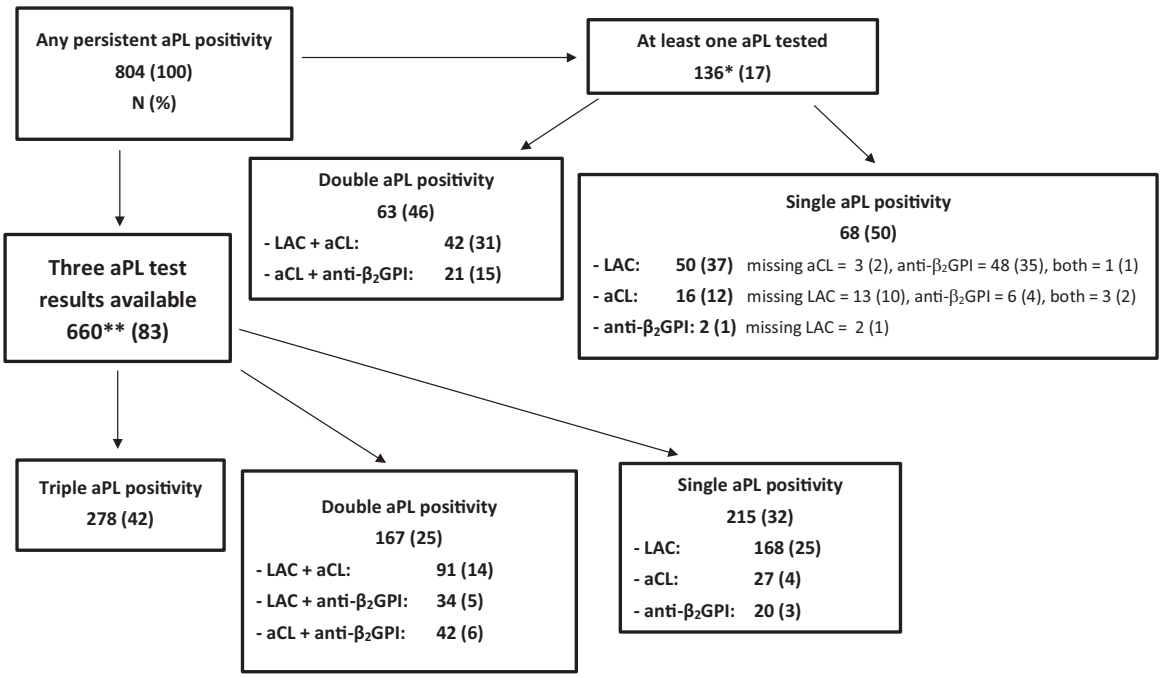

Figure 1. Antiphospholipid antibody (aPL) profile at baseline in patients (n = 804) with persistent aPL positivity who were included in the Antiphospholipid Syndrome Alliance for Clinical Trials and International Networking Registry. * = Of 804 patients, 136 (17%) had missing data for their aPL profiles, with 38 (5%) patients lacking test results for LAC, 3 (0.3%) lacking test results for aCLs, 100 (12%) lacking test results for anti-β_2GPI antibodies. ** = An additional 8 patients (1%) were tested for 3 aPLs but were excluded from the study due to low titers (20–39 units) on an aPL enzyme-linked immunosorbent assay with negative LAC test. Anti-β_2GPI = anti–β_2-glycoprotein I; aCL = anticardiolipin antibody; LAC = lupus anticoagulant.

persistent thrombocytopenia, aPL-related nephropathy, and cardiac valve disease. Within the group with single aPL positivity, LAC only positivity had the highest proportion of patients who experienced any vascular events, pregnancy morbidity, and other clinical manifestations (Table 2).

DISCUSSION

Based on our multi-center international cohort of aPL-positive patients, 20% of patients who met the entry criteria did not fulfill classification criteria for clinical APS, 71% experienced vascular events, 50% of individuals with a history of pregnancy had aPL-related obstetric morbidity, and 56% had at least one non-criteria clinical aPL manifestation—most commonly thrombocytopenia and white matter lesions. Non-criteria clinical manifestations were highest in the thrombotic APS/obstetric APS group versus thrombotic APS or obstetric APS only. APS patients overall had higher rates of current anticoagulation therapy and statin use, but lower aspirin and hydroxychloroquine use than patients with aPL positivity without APS at registry entry. Compared to single–, double–, and triple–aPL-positive subgroups, the subgroup consisting of patients with single aPL positivity excluding LAC only had substantially lower frequencies of vascular events, pregnancy morbidity, and other clinical events; this suggests that LAC positivity appears to be a major contributor to aPL-related clinical features.

Our study adds to prior work demonstrating the clinical heterogeneity of aPLs, which can result in a broad spectrum of clinical manifestations. Although the current revised classification criteria for APS incorporates vascular events and pregnancy morbidity, various "non-criteria" manifestations, known to occur frequently in aPL-positive patients, were not included (16–18). Since then, various systematic reviews and meta-analyses in SLE patients have aimed to better characterize the role of aPL-related "non-criteria" manifestations, demonstrating an increased likelihood of cardiac valve disease, pulmonary hypertension, livedo reticularis, thrombocytopenia, hemolytic anemia, and renal impairment in aPL-positive SLE patients compared to aPL-negative SLE patients (11,19). Other investigators have assessed these manifestations in APS patients in the absence or presence of concomitant systemic autoimmune disease and have demonstrated increased rates of cognitive dysfunction, white matter lesions, aPL-related nephropathy, thrombocytopenia, and livedo reticularis among these individuals (10,20). The present study adds to this literature by demonstrating that non-criteria manifestations, most commonly white matter lesions and thrombocytopenia, occurred in the majority (56%) of international aPL-positive patients and were more likely to occur in thrombotic APS/obstetric APS patients (66%), suggesting that non-criteria manifestations are prevalent in aPL-positive patients and potentially associated with more severe disease (20,21). In fact, efforts are underway using cluster analysis methodology, a data-driven method that groups patients by combinations of aPL profiles and clinical features, to further identify clinical phenotypes and distinct "clusters" of patients enrolled in the APS ACTION registry (22,23).

Assessment of clinical phenotypes, along with a better understanding of the role of aPL laboratory profiles, may play a critical role in risk stratification for aPL-positive patients (24). Although the definition of a "clinically significant" and "high-risk" aPL profile has not been clearly defined, different aPL profiles appear to confer different thrombosis risks (7,25–27). Positive LAC test (compared to enzyme-linked immunosorbent testing for aCLs or anti-β_2GPI antibodies), moderate-to-high (≥40 units) titers of aCLs or anti-β_2GPI antibodies (compared to lower titers), IgG isotype (compared to IgM and IgA isotype), and triple positivity for aPLs (compared to single or double positivity for aPLs) have a stronger correlation with aPL-related clinical events (28–30). However, there is ongoing debate about the clinical significance of isolated LAC positivity and whether it is as important as triple aPL positivity. Additionally, one recent study demonstrated that aCL IgG, but not IgM, and LAC test positivity are associated with higher rates of thromboses in SLE patients (31). Our cross-sectional analysis, demonstrating a relatively similar frequency of aPL-related clinical events in single, double, and triple aPL positivity, and a substantially lower frequency in single aPL positivity patients without positive findings on LAC test, supports the association of clinical events with LAC positivity. Furthermore, while accumulating data show that LAC positivity may be a stronger risk factor for thrombosis and pregnancy morbidity compared to positivity for either aCLs or anti-β_2GPI antibodies (1,32), standardization of laboratory testing and cutoff thresholds are still needed (1). Prospective studies will determine the association between laboratory study levels and clinically relevant disease.

While anticoagulation is the mainstay in treatment of aPL-related clinical events in thrombotic APS (33), alternative treatments are needed in patients with refractory disease or microvascular APS (24,34–39). Although the majority of APS patients overall received anticoagulation therapy, less than half received aspirin or immunosuppression treatment, and few received other treatments, such as intravenous immunoglobulin, plasma exchange, or rituximab. This finding may reflect an inherently low rate of refractory/microvascular APS or selection bias in our cohort. While data regarding treatment of obstetric APS are controversial in regard to the need for prophylactic low-dose aspirin versus the addition of unfractionated heparin to low-dose aspirin (40–44), the majority of patients in our cohort with obstetric APS only received aspirin (ever and at registry entry) and low molecular weight heparin (ever).

Furthermore, no clear consensus exists on primary prevention management of the symptoms of patients with persistent positivity for aPLs (45), including the use of aspirin, hydroxychloroquine, or anticoagulation therapy, although recent European Alliance of Associations for Rheumatology guidelines suggest that low-dose aspirin may be beneficial for various patients who were

aPL-positive (46). Our registry data show that the majority of patients who have aPL positivity without APS were treated with aspirin (67%) and hydroxychloroquine (56%), which may be driven by use of these medications for the prevention of thrombosis, underlying concomitant systemic autoimmune disease (45% of patients who have aPL positivity without APS), or other comorbid medical diseases, including cardiovascular risk factors. Patients who had positivity for aPLs without a diagnosis of APS had the highest percentage of concomitant SLE, which may have prompted aPL testing in this group.

Although we previously reported that LAC positivity, livedo reticularis, and cognitive dysfunction are more common in patients recruited from Brazil compared to those recruited from other parts of the world (47), the current study did not investigate specific clinical and laboratory differences by geographic region as a comprehensive regional analysis of the registry is already ongoing. Additionally, the low rate of Black patients (3–4%) in the registry may reflect selection bias (e.g., half of the patients were recruited from Europe) or disparities in access to care and would be worth investigation in future studies.

While the present study was limited in its retrospective, cross-sectional study design, we used data from a large, multicenter international patient cohort enriched with granular sociodemographic, clinical, laboratory, and medication information. Epidemiologic studies focusing on APS are limited; few large APS cohorts that are inclusive of different genders, races, and geographic regions are available to estimate the distribution of APS across clinical and laboratory subtypes. As data collection is ongoing in our registry, our data represents an interim assessment of baseline characteristics. Future analyses will use statistical testing and APS ACTION core laboratory aPL test results to evaluate significant differences between subgroups. Selection bias could be a factor in the low percentage of "other" clinical manifestations and systemic autoimmune disease in the obstetric APS group, as some patients in this group are recruited from obstetrics clinics. Our future prospective study will assess the risk of incident systemic autoimmune disease development after the diagnosis of primary obstetric APS.

Although selection and referral bias to APS "experts" should be considered in the interpretation of our registry data, our study demonstrated a low rate of CAPS or use of medications suggestive of refractory disease. Additionally, given that aPL profiles were not necessarily collected at the time of clinical events, our results should be confirmed in prospective studies. Moreover, while other "non-criteria" aPL tests such as those for anti–phosphatidylserine/prothrombin and anti–domain 1 antibodies, have increasingly shown to contribute to a diagnosis of APS and risk assessment for thrombosis (48,49), our study did not evaluate these laboratory tests as they are not currently standardized or widely commercially available. Finally, although we did not stratify our cohort by those with or without an systemic autoimmune disease, in a previous analysis of APS ACTION registry patients, the frequencies of thrombosis and pregnancy morbidity were similar in aPL-positive patients in the absence or presence of concomitant SLE; however, SLE in patients with persistent aPL positivity was associated with increased frequency of thrombocytopenia, hemolytic anemia, low complement, and positive findings for IgA anti–β_2GPI antibodies (50).

In conclusion, our study demonstrates the heterogeneity of aPL-related clinical manifestations and laboratory profiles in a multicenter, international cohort of aPL-positive patients. Identification of APS patients by different clinical phenotypes and aPL profiles may improve risk stratification and help physicians and researchers better characterize the disease and understand clinical outcomes. Future prospective analyses, using standardized core laboratory aPL tests, will help clarify the role of aPL risk profiles.

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. Barbhaiya 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. Sevim, Zisa, Bertolaccini, Erkan, Barbhaiya.

Acquisition of data. Sevim, Zisa, Andrade, Sciascia, Pengo, Tektonidou, Ugarte, Gerosa, Belmont, Aguirre Zamorano, Fortin, Ji, Efthymiou, Cohen, Branch, Ramires de Jesus, Andreoli, Petri, Rodriguez, Cervera, Knight, Atsumi, Willis, Roubey, Bertolaccini, Erkan, Barbhaiya.

Analysis and interpretation of data. Sevim, Zisa, Andrade, Sciascia, Pengo, Tektonidou, Ugarte, Gerosa, Belmont, Aguirre Zamorano, Fortin, Ji, Efthymiou, Cohen, Branch, Ramires de Jesus, Andreoli, Petri, Rodriguez, Cervera, Knight, Atsumi, Willis, Roubey, Bertolaccini, Erkan, Barbhaiya.

REFERENCES

1. Miyakis S, Lockshin M, Atsumi T, Branch D, Brey R, Cervera R, et al. International consensus statement on an update of the classification criteria for definite antiphospholipid syndrome (APS). J Thromb Haemost 2006;4:295–306.
2. Barbhaiya M, Andrade D, Bertolaccini M, Erkan D. Antiphospholipid Syndrome Alliance for Clinical Trials and International Networking (APS ACTION). In: Erkan D, Lockshin MD, editors. Antiphospholipid Syndrome: Current Research Highlights and Clinical Insights. Cham: Springer International Publishing; 2017. p. 267–76.
3. Levine J, Branch, DW, Rauch J. The antiphospholipid syndrome. N Engl J Med 2002;346:752–63.
4. Cervera R, Serrano R, Pons-Estel G, Ceberio-Hualde L, Shoenfeld Y, Ramón E, et al. Morbidity and mortality in the antiphospholipid syndrome during a 10-year period: a multicentre prospective study of 1000 patients. Ann Rheum Dis 2014;74:1011–8.
5. Abdel-Wahab N, Lopez-Olivio M, Pinto-Patarroyo G, Suarez-Almazor M. Systematic review of case reports of antiphospholipid syndrome following infection. Lupus 2016;25:1520–31.
6. Reinstein E, Shoenfeld Y. Antiphospholipid syndrome and cancer. Clin Rev Allergy Immunol 2007;32:184–7.
7. Pengo V, Biasiolo A, Pegoraro C, Cucchini U, Noventa F, Iliceto S. Antibody profiles for the diagnosis of antiphospholipid syndrome. Thromb Haemost 2005;93:1147–52.

8. Pengo V, Ruffatti A, Legnani C, Gresele P, Barcellona D, Erba N, et al. Clinical course of high-risk patients diagnosed with antiphospholipid syndrome. J Thromb Haemost 2010;8:237–42.
9. Yelnik C, Urbanski G, Drumez E, Sobanski V, Maillard H, Lanteri A, et al. Persistent triple antiphospholipid antibody positivity as a strong risk factor of first thrombosis, in a long-term follow-up study of patients without history of thrombosis or obstetrical morbidity. Lupus 2017;26:163–9.
10. Cervera R, Piette J, Font J, Khamashta M, Shoenfeld Y, Camps M, et al. Antiphospholipid syndrome: clinical and immunologic manifestations and patterns of disease expression in a cohort of 1,000 patients. Arthritis Rheum 2002;46:1019–27.
11. Ünlü O, Zuily S, Erkan D. The clinical significance of antiphospholipid antibodies in systemic lupus erythematosus. Eur J Rheumatol 2016;3:75–84.
12. Duarte-García A, Pham MM, Crowson CS, Amin S, Moder KG, Pruthi RK, et al. The epidemiology of antiphospholipid syndrome: a population-based study. Arthritis Rheumatol 2019;71:1545–52.
13. Pengo V, Tripodi A, Reber G, Rand J, Ortel T, Galli M, et al. Update of the guidelines for lupus anticoagulant detection. J Thromb Haemost 2009;7:1737–40.
14. Harris P, Taylor R, Thielke R, Payne J, Gonzalez N, Conde J. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Infor 2009;42.
15. Asherson R, Cervera R, de Groot P, Erkan D, Boffa M-C, Piette JC. Catastrophic antiphospholipid syndrome: international consensus statement on classification criteria and treatment guidelines. Lupus 12:530–4.
16. Unlu O, Domingues V, de Jesús GR, Zuily S, Espinosa G, Cervera R, et al. Definition and epidemiology of antiphospholipid syndrome. In: Erkan D, Lockshin MD, editors. Antiphospholipid Syndrome: Current Research Highlights and Clinical Insights. Cham: Springer International Publishing; 2017. p. 147–69.
17. Blank M, Cohen J, Toder V, Shoenfeld Y. Induction of antiphospholipid syndrome in naive mice with mouse lupus monoclonal and human polyclonal anti-cardiolipin antibodies. Proc Natl Acad Sci U S A 1991;88:3069–73.
18. Pierangeli SS, Liu XW, Barker JH, Anderson G, Harris EN. Induction of thrombosis in a mouse model by IgG, IgM and IgA immunoglobulins from patients with the antiphospholipid syndrome. Thromb Haemost 1995;74:1361–7.
19. Zuily S, Wahl D. Pulmonary hypertension in antiphospholipid syndrome. Curr Rheumatol Rep 2015;17:478.
20. Tektonidou MG, Varsou N, Kotoulas G, Antoniou A, Moutsopoulos HM. Cognitive deficits in patients with antiphospholipid syndrome: association with clinical, laboratory, and brain magnetic resonance imaging findings. Arch Intern Med 2006;166:2278–84.
21. Abreu MM, Danowski A, Wahl DG, Amigo MC, Tektonidou M, Pacheco MS, et al. The relevance of "non-criteria" clinical manifestations of antiphospholipid syndrome: 14th International Congress on Antiphospholipid Antibodies Technical Task Force Report on Antiphospholipid Syndrome Clinical Features. Autoimmun Rev 2015;14:401–14.
22. Zuily S, Clerc-Urmes I, Wahl D, Erkan D. Antiphospholipid Syndrome Alliance for Clinical Trials and International Networking (APS ACTION) Clinical Database and Repository Cluster Analysis for Identification of Different Clinical Phenotypes Among Antiphospholipid Antibody-Positive Patients. Lupus 2016;25:74–5.
23. Zuily S, Clerc-Urmes I, Chighizola C, Baumann C, Wahl D, Meroni P, et al. Antiphospholipid Syndrome Alliance for Clinical Trials and International Networking (APS ACTION) clinical database and repository cluster analysis for the identification of different clinical phenotypes among antiphospholipid antibody-positive female patients with a history of pregnancy. Lupus 2016;25:61.
24. Garcia D, Erkan D. Diagnosis and management of the antiphospholipid syndrome. N Engl J Med 2018;378:2010–21.
25. Finazzi G, Brancaccio V, Moia M, Ciavarella N, Mazzucconi MG, Schinco P, et al. Natural history and risk factors for thrombosis in 360 patients with antiphospholipid antibodies: a four-year prospective study from the Italian registry. Am J Med 1996;100:530–6.
26. Gresele P, Migliacci R, Vedovati MC, Ruffatti A, Becattini C, Facco M, et al. Patients with primary antiphospholipid antibody syndrome and without associated vascular risk factors present a normal endothelial function. Thromb Res 2009;123:444–51.
27. Zuily S, de Laat B, Mohamed S, Kelchtermans H, Shums Z, Albesa R, et al. Validity of the global anti-phospholipid syndrome score to predict thrombosis: a prospective multicentre cohort study. Rheumatology (Oxford) 2015;54:2071–5.
28. Mustonen P, Lehtonen KV, Javela K, Puurunen M. Persistent antiphospholipid antibody (aPL) in asymptomatic carriers as a risk factor for future thrombotic events: a nationwide prospective study. Lupus 2014;23:1468–76.
29. Wahl DG, Guillemin F, Maistre E de, Perret C, Lecompte T, Thibaut G. Risk for venous thrombosis related to antiphospholipid antibodies in systemic lupus erythematosus: a meta-analysis. Lupus 1997;6:467–73.
30. Zuily S, Regnault V, Selton-Suty C, Eschwege V, Bruntz J-F, Bode-Dotto E, et al. Increased risk for heart valve disease associated with antiphospholipid antibodies in patients with systemic lupus erythematosus: meta-analysis of echocardiographic studies. Circulation 2011;124:215–24.
31. Domingues V, Magder LS, Petri M. Assessment of the independent associations of IgG, IgM and IgA isotypes of anticardiolipin with thrombosis in SLE. Lupus Sci Med 2016;3:e000107.
32. Lockshin M, Kim M, Laskin C, Guerra M, Branch D, Merrill J, 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.
33. Erkan D, Aguiar CL, Andrade D, Cohen H, Cuadrado MJ, Danowski A, et al. 14th International Congress on Antiphospholipid Antibodies Task Force Report on Antiphospholipid Syndrome Treatment Trends. Autoimmunity Reviews 2014;13:685–696.
34. Edwards MH, Pierangeli S, Liu X, Barker JH, Anderson G, Harris EN. Hydroxychloroquine reverses thrombogenic properties of antiphospholipid antibodies in mice. Circulation 1997;96:4380–4.
35. Nuri E, Taraborelli M, Andreoli L, Tonello M, Gerosa M, Calligaro A, et al. Long-term use of hydroxychloroquine reduces antiphospholipid antibodies levels in patients with primary antiphospholipid syndrome. Immunol Res 2017;65:17–24.
36. Erkan D, Vega J, Ramón G, Kozora E, Lockshin MD. A pilot open-label phase II trial of rituximab for non-criteria manifestations of antiphospholipid syndrome. Arthritis Rheum 2013;65:464–71.
37. Elazary AS, Klahr P, Hershko A, Dranitzki Z, Rubinow A, Naparstek Y. Rituximab induces resolution of recurrent diffuse alveolar hemorrhage in a patient with primary antiphospholipid antibody syndrome. Lupus 2011;21:438–40.
38. Lonze BE, Zachary AA, Magro CM, Desai NM, Orandi BJ, Dagher NN, et al. Eculizumab prevents recurrent antiphospholipid antibody syndrome and enables successful renal transplantation. Am J Transplant 2014;14:459–65.
39. Zapantis E, Furie R, Horowitz D. Response to eculizumab in the antiphospholipid antibody syndrome. Ann Rheum Dis 2015;74:341.
40. Rai R, Cohen H, Dave M, Regan L. Randomised controlled trial of aspirin and aspirin plus heparin in pregnant women with recurrent miscarriage associated with phospholipid antibodies (or antiphospholipid antibodies). BMJ 1997;314:253–7.

41. Kutteh WH. Antiphospholipid antibody-associated recurrent pregnancy loss: treatment with heparin and low-dose aspirin is superior to low-dose aspirin alone. Am J Obstet Gynecol 1996;174: 1584–9.
42. Farquharson RG, Quenby S, Greaves M. Antiphospholipid syndrome in pregnancy: a randomized, controlled trial of treatment. Obstet Gynecol 2002;100:408–13.
43. Laskin CA, Spitzer KA, Clark CA, Crowther MR, Ginsberg JS, Hawker GA, et al. Low molecular weight heparin and aspirin for recurrent pregnancy loss: results from the randomized, controlled HepASA Trial. J Rheumatol 2009;36:279–87.
44. Munoz-Rodriguez FJ, Font J, Cervera R, Reverter JC, Tassies D, Espinosa G, et al. Clinical study and follow-up of 100 patients with the antiphospholipid syndrome. Semin Arthritis Rheum 1999;29: 182–90.
45. Zuo Y, Barbhaiya M, Erkan D. Primary thrombosis prophylaxis in persistently antiphospholipid antibody-positive individuals: where do we stand in 2018? Curr Rheumatol Rep 2018;20:66.
46. Tektonidou MG, Andreoli L, Limper M, Amoura Z, Cervera R, Costedoat-Chalumeau N, et al. EULAR recommendations for the management of antiphospholipid syndrome in adults. Ann Rheum Dis 2019;78:1296.
47. Ugolini-Lopes M, Rosa R, Nascimento I, R de Jesus G, Levy R, Erkan D, et al, on behalf of APS ACTION registry. APS ACTION clinical database and repository analysis: primary antiphospholipid syndrome in Brazil versus other regions of the world [abstract]. Lupus 2016;25 Supp 1S:76.
48. Sciascia S, Murru V, Sanna G, Roccatello D, Khamashta MA, Bertolaccini ML. Clinical accuracy for diagnosis of antiphospholipid syndrome in systemic lupus erythematosus: evaluation of 23 possible combinations of antiphospholipid antibody specificities. J Thromb Haemost 2012;10:2512–8.
49. Pengo V, Ruffatti A, Tonello M, Cuffaro S, Banzato A, Bison E, et al. Antiphospholipid syndrome: antibodies to Domain 1 of beta2-glycoprotein 1 correctly classify patients at risk. J Thromb Haemost 2015;13:782–7.
50. Unlu O, Erkan D, Barbhaiya M, Andrade D, Nascimento I, Rosa R, et al. The impact of systemic lupus erythematosus on the clinical phenotype of antiphospholipid antibody-positive patients: results from the AntiPhospholipid Syndrome Alliance for Clinical Trials and InternatiOnal clinical database and repository. Arthritis Care Res (Hoboken) 2019;71:134–41.

APPENDIX A: THE APS ACTION INVESTIGATORS

Guillermo Pons-Estel (Santa Fe, Argentina), Bill Giannakopoulos, Steve Krilis (Sydney, Australia), Guilherme de Jesus, Roger Levy (Rio de Janeiro, Brazil), Danieli Andrade (São Paulo, Brazil), Paul R. Fortin (Quebec City, Canada), Lanlan Ji, Zhouli Zhang (Beijing, China); Stephane Zuily, Denis Wahl (Nancy, France); Maria G. Tektonidou (Athens, Greece) Cecilia Nalli, Laura Andreoli, Angela Tincani (Brescia, Italy), Cecilia B. Chighizola, Maria Gerosa, Pierluigi Meroni (Milan, Italy), Vittorio Pengo (Padova, Italy), Savino Sciascia (Tulin, Italy), Karel De Ceulaer, Stacy Davis (Kingston, Jamaica), Olga Amengual, Tatsuya Atsum (Sapporo, Japan), Imad Uthman (Beirut, Lebanon) Maarten Limper, Philip de Groot (Utrecht, The Netherlands), Guillermo Ruiz Irastorza, Amaia Ugarte (Barakaldo, Spain), Ignasi Rodriguez-Pinto, Ricard Cervera (Barcelona, Spain), Esther Rodriguez, Maria Jose Cuadrado (Madrid, Spain), Maria Angeles Aguirre Zamorano, Rosario Lopez-Pedrera (Cordoba, Spain), Bahar Artim-Esen, Murat Inanc (Istanbul, Turkey), Maria Laura Bertolaccini, Hannah Cohen, Maria Efthymiou, Munther Khamashta, Ian Mackie, Giovanni Sanna (London, UK), Jason Knight (Ann Arbor, Michigan, US), Michelle Petri (Baltimore, Maryland, US), Robert Roubey (Chapel Hill, North Carolina, US), Tom Ortel (Durham, North Carolina, US), Emilio Gonzalez, Rohan Willis (Galveston, Texas, US), Michael Belmont, Steven Levine, Jacob Rand, Medha Barbhaiya, Doruk Erkan, Jane Salmon, Michael Lockshin (New York City, New York, US), and Ware Branch (Salt Lake City, Utah, US).

ARP Announcements

Association of Rheumatology Professionals
2200 Lake Boulevard NE, Atlanta, Georgia 30319
www.rheumatology.org

Submissions are invited for the AC&R 2023 Themed Issue: Health Disparities and Health Equity in the Rheumatic Diseases

Arthritis Care & Research is soliciting manuscripts for a Themed Issue addressing health disparities and health equity, as related to outcomes and issues in the rheumatic diseases. This theme issue is part of an initiative across the American College of Rheumatology journals to better address these important issues in our field. Manuscripts may focus on health disparities or health equity related to race and ethnicity, sex and gender, socioeconomic status, sexual orientation, or other characteristics. Exploration of health disparities among understudied patient groups, as well as the intersection of multiple patient characteristics, are encouraged.

Manuscripts covering a broad range of topics related to the major theme are invited. Examples include observational studies that elucidate factors underlying disparities in health care quality or access; intervention studies that address health disparities; studies of differential impacts of treatments or behavioral interventions; studies describing mechanisms underlying disparities in key outcomes in rheumatic diseases (e.g., pain, function). Manuscripts addressing research related to disparities in rheumatology training and work force are also of interest. Both Original Research and Review articles will be considered.

The 2023 Themed Issue will include regular submissions as well, but a certain number of pages will be reserved for manuscripts accepted in response to this solicitation. All manuscripts will be peer reviewed. The Editor will select papers for publication in the Themed Issue based on reviewer ratings and the balance of subject matter. It is possible that manuscripts submitted for the themed issue may be accepted for publication in a regular issue of *Arthritis Care & Research*, rather than the themed issue.

Please follow the formatting requirements found in the Author Guidelines section at https://onlinelibrary.wiley.com/page/journal/21514658/homepage/ForAuthors.html. The deadline for submission is March 31, 2022. For further information, contact the Editor of Arthritis Care & Research, Dr. Kelli D. Allen; email: kdallen@email.unc.edu.

Download the ACR Publications Mobile App

The ACR Publications app can be downloaded for free from the Apple store or Google Play. ACR members can log in for full-text access to all articles in *Arthritis Care & Research* and *Arthritis & Rheumatology*. Nonmembers can access abstracts of all *AC&R* and *A&R* articles, the full text of articles published more than one year ago, and select open-access articles published recently, as well as the full text of all articles from *ACR Open Rheumatology* and *The Rheumatologist*.

ARP Membership

The Association of Rheumatology Professionals (ARP), a division of the American College of Rheumatology, appreciates your continued membership and looks forward to serving you another year. Membership costs range from $30 to $140. ARP welcomes nurse practitioners, nurses, physician assistants, office staff, researchers, physical therapists, occupational therapists, assistants, and students. Student membership is complimentary; the Annual Meeting registration fee is waived for students who submit the required student verification letter. For information, go to www.rheumatology.org and select “Membership” or call 404-633-3777 and ask for an ARP staff member.

ACR Journal Twitter Account (@ACR_Journals) and Social Media Editor

The ACR journals are heightening our focus on social media, to benefit authors and readers. Among our first activities is the introduction of an official ACR Journals Twitter account: @ACR_Journals. Followers will enjoy special features and the opportunity to engage with authors and other fellow professionals about studies published in *Arthritis Care & Research, Arthritis & Rheumatology,* and *ACR Open Rheumatology*. Authors of published articles will have the opportunity to use @ACR_Journals to share their work and engage in dialogue with others interested in the research. The journals welcome Dr. Paul Sufka of Minneapolis as our first Social Media Editor.