Arthritis Care & Research

Aims and Scope

Arthritis Care & Research is an official journal of the American College of Rheumatology and the Association of Rheumatology Professionals, a division of the College. *Arthritis Care & Research* is a peer-reviewed journal that publishes both original research and review articles that promote excellence in the clinical practice of rheumatology. Relevant to the care of individuals with arthritis and related disorders, major topics are evidence-based practice studies, clinical problems, practice guidelines, health care economics, health care policy, educational, social, and public health issues, and future trends in rheumatology practice.

Volume 77 No. 12 December 2025

Arthritis Care & Research

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

Cover design: Sandra Pulmano

♾ This journal is printed on acid-free paper.

Arthritis Care & Research

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

VOLUME 77 • DECEMBER 2025 • NO. 12

Cover image: The image on the cover (Puche-Larrubia et al; pages 1395–1402) shows bilateral camptodactyly on radiograph.

Arthritis Care & Research
Vol. 77, No. 12, December 2025, pp 1395–1402
DOI 10.1002/acr.25597

CLINICOPATHOLOGIC CONFERENCE

A Twist in the Diagnosis: Chronic Arthropathy Without Inflammation

María Á. Puche-Larrubia,[1,2,3] Inmaculada C. Aranda-Valera,[1,2,3] Alejandro Escudero-Contreras,[1,2,3] and Rosa Roldán-Molina[1,2,3]

CASE PRESENTATION

Reason for consultation and medical history

A six-year-old boy was referred to pediatric rheumatology in 2017 due to recurrent episodes of joint swelling affecting his knees, elbows, wrists, and ankles. These episodes began approximately one year earlier and were intermittent, lasting several days to weeks. The swelling was symmetric, nonpainful, and not associated with significant morning stiffness or functional limitation. Between episodes, the patient returned to his baseline status, with no residual symptoms, and remained physically active, participating in age-appropriate activities without restrictions.

There was no reported fever, rash, red or painful eyes, oral ulcers, or constitutional symptoms such as weight loss or fatigue. No history of trauma or infection was identified as a trigger.

Before rheumatologic evaluation, no disease-modifying treatments had been initiated. Nonsteroidal anti-inflammatory drugs (NSAIDs) were prescribed occasionally, with mild improvement in swelling. Additionally, the patient received intra-articular glucocorticoid injections in both knees, without significant or sustained clinical improvement.

He was the first child of two White parents. Although no known consanguinity was reported, both parents originated from the same small, low-population town. His mother had autoimmune hypothyroidism, whereas his father was healthy, and no other significant familial conditions were reported.

Regarding his personal medical history, the patient had undergone surgery at age two years for bilateral trigger fingers affecting the third digit of both hands—an unusual presentation at that age and in the absence of trauma. Unfortunately, detailed surgical records were not available, although the postoperative course had been uneventful. Additionally, he had a diagnosis of lactose intolerance, with recurrent abdominal discomfort triggered by dairy.

Consent for the use of the patient's data and images was obtained from his parents. This case report was exempt from institutional review board approval.

Physical examination

Physical examination revealed synovial hypertrophy in the wrists and elbows, without associated pain or functional limitation. Both knees and ankles also showed synovial thickening, and hip mobility was slightly restricted in internal and external rotation. The second phalanges of both hands were fixed in flexion consistent with a diagnosis of camptodactyly (Figure 1A). The feet exhibited a cavus deformity with an equinus position.

Laboratory evaluation

Laboratory investigations revealed normal markers of inflammation, including C-reactive protein (CRP) and erythrocyte sedimentation rate (ESR). Autoimmune testing showed negative results for antinuclear antibodies (ANAs), anti–double-stranded DNA (anti-dsDNA), extractable nuclear antigens, anti–cyclic citrullinated peptide antibodies (ACPAs), rheumatoid factor (RF), anti-transglutaminase antibodies, and thyroid peroxidase antibodies. HLA-B typing was positive for B13/B35.

To evaluate potential infectious etiologies, serologic screening (IgM and IgG) was performed for a wide panel of common pathogens, including Epstein–Barr virus, cytomegalovirus, HIV, hepatitis B virus, hepatitis C virus, herpes simplex virus, human parvovirus B19, *Rickettsia* species, *Coxiella burnetii*, and *Toxoplasma gondii*. All results were negative for acute infection. Additionally, blood cultures obtained during the diagnostic process were sterile, with no bacterial or mycobacterial growth.

[1]Reina Sofia University Hospital, Cordoba, Spain; [2]Maimonides Institute for Biomedical Research of Cordoba, Cordoba, Spain; [3]Medical and Surgical Sciences, University of Cordoba, Cordoba, Spain.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25597.

Address correspondence via email to María Á. Puche-Larrubia, MD, PhD, at mangeles.puche@gmail.com.

Submitted for publication March 25, 2025; accepted in revised form June 26, 2025.

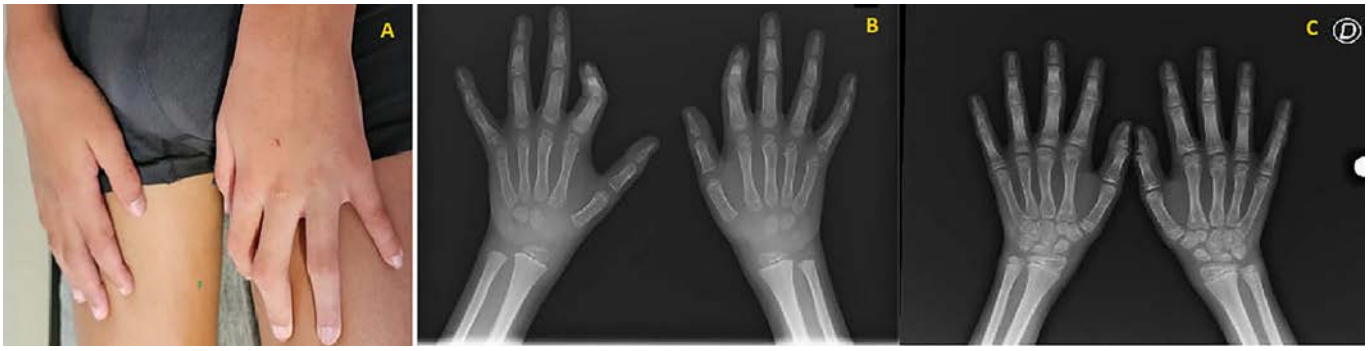

Figure 1. Clinical and radiographic evolution of camptodactyly. (A) Clinical photograph showing bilateral fixed flexion deformities of the fourth fingers at age 13 years. (B and C) Anteroposterior hand radiographs at ages 6 and 13 years, respectively. The initial image (B) reveals bilateral camptodactyly. The follow-up image (C) shows persistence of the deformities without evidence of progression, joint space narrowing, or erosive changes.

Synovial fluid analysis

An arthrocentesis of the left knee was performed, yielding approximately 15 mL of clear, yellow, and viscous synovial fluid. The sample was sent for cytologic, microbiologic, and biochemical analysis. Cytology of the synovial fluid from the left knee was negative for malignancy. No crystals were identified under polarized light microscopy. Microbiologic studies revealed a negative Gram stain, sterile aerobic and anaerobic bacterial cultures, and no growth on *Mycobacterium tuberculosis*–specific culture media. Acid-fast bacilli staining was also negative. The synovial fluid contained 345 leukocytes/mm^3, with a differential of 57% neutrophils and 43% mononuclear cells, consistent with a noninflammatory profile.

Imaging studies

The first hand radiograph, obtained at age six years, showed bilateral flexion deformities of the proximal interphalangeal joints of the fourth fingers, consistent with congenital camptodactyly. A follow-up radiograph at age 13 years showed persistence of the deformities, without radiographic progression, joint space narrowing, or erosive changes (Figure 1B and C).

Knee radiographs revealed epiphyseal enlargement and hypoplastic patellae (Figure 2A and B). Magnetic resonance imaging (MRI) confirmed bilateral joint effusion and demonstrated a well-defined intraosseous cyst in the medial metaphysis of the right proximal tibia (Figure 2C and D). No evidence of erosions or cartilage loss was observed.

A pelvic radiograph demonstrated coxa vara with smooth flattening of the femoral heads and well-defined bilateral subchondral radiolucencies in the acetabula suggestive of intraosseous cysts (Figure 3A). MRI of the hips revealed large bilateral joint effusions and intraosseous acetabular cysts without signs of erosive damage (Figure 3B). An incidental os odontoideum was identified on cervical spine MRI in the absence of neurologic symptoms.

Additional systemic evaluation

Results of cardiac, pulmonary, ophthalmologic, and dermatologic evaluations were normal. The cardiac workup included an echocardiogram, which revealed no abnormalities.

CASE SUMMARY

A six-year-old boy presented with painless, symmetrical joint swelling and camptodactyly. Laboratory studies revealed normal markers of inflammation and negative autoantibody panel results (ANA, RF, ACPA). Imaging showed hypoplastic patellae, coxa vara, an intraosseous femoral cyst, and large acetabular cysts, along with nonerosive joint effusions.

DIFFERENTIAL DIAGNOSIS

Given the patient's clinical progression, the differential diagnosis included juvenile idiopathic arthritis (JIA), mucopolysaccharidosis type I (MPS I), and the rare condition known as camptodactyly-arthropathy–coxa vara–pericarditis (CACP) syndrome. The differences between these conditions are summarized in Table 1.

JIA. JIA is the most common chronic rheumatic disease of childhood, with an estimated prevalence of 1 to 2 per 1,000 children worldwide.[1] It encompasses a heterogeneous group of disorders characterized by arthritis of unknown etiology lasting more than six weeks in children under 16 years of age. JIA is classified into several subtypes according to the criteria established by the International League of Associations for Rheumatology.[2] These subtypes include oligoarticular, polyarticular (RF-positive and RF-negative), systemic, enthesitis-related, psoriatic, and undifferentiated arthritis, which are defined by clinical presentation and specific laboratory features.[3,4]

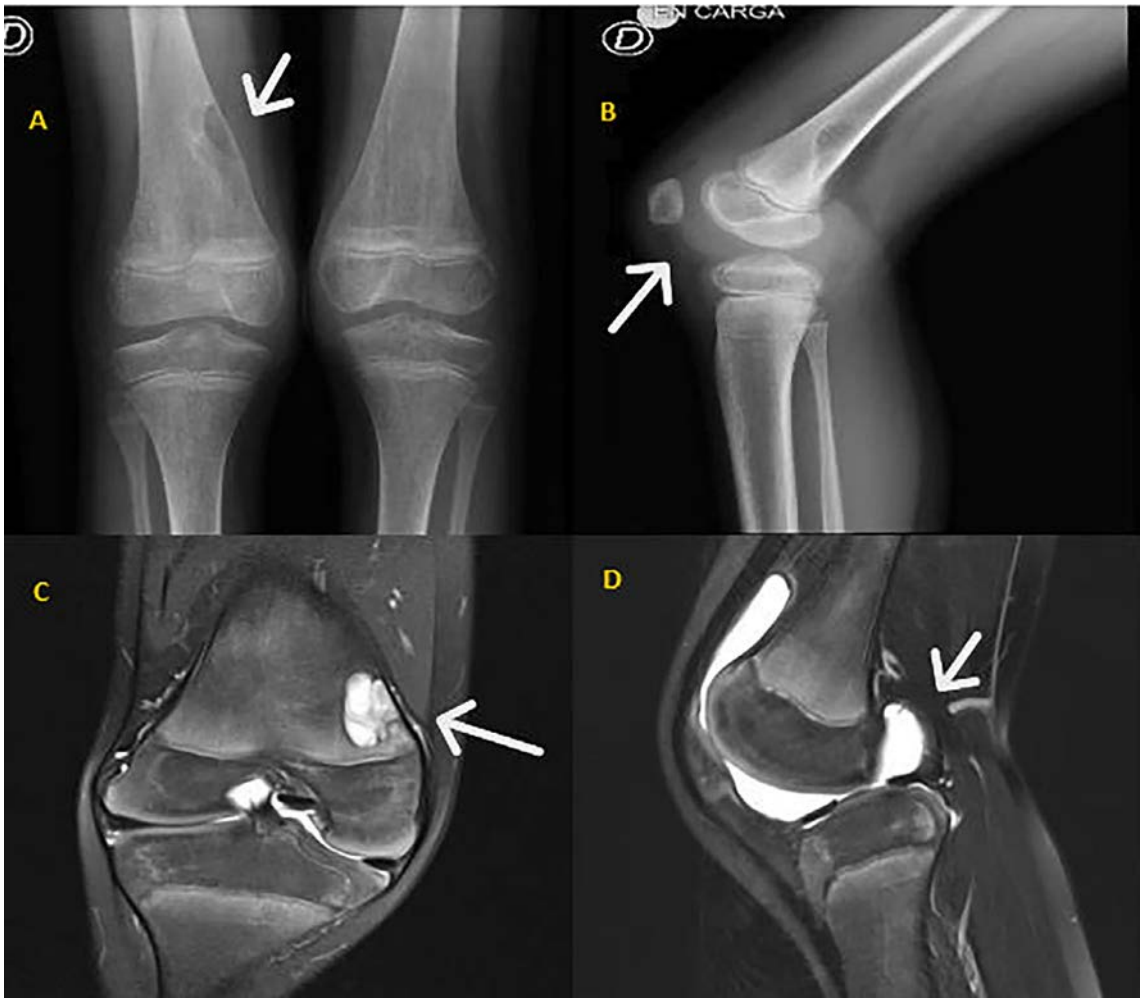

Figure 2. Radiographic and MRI findings of the knees. (A) Anteroposterior plain radiograph of the knees showing joint space widening and periarticular osteopenia without erosions. A well-defined intraosseous cyst is visible in the metaphyseal region of the right distal femur (white arrow). (B) Lateral radiograph of the knees demonstrating hypoplastic patellae (white arrow) and maintained joint spaces without destructive changes. (C) Coronal MRI view confirming the intraosseous cyst in the right femoral metaphysis (white arrow). (D) Sagittal MRI view showing a prominent joint effusion with synovial distension (white arrow). MRI, magnetic resonance imaging.

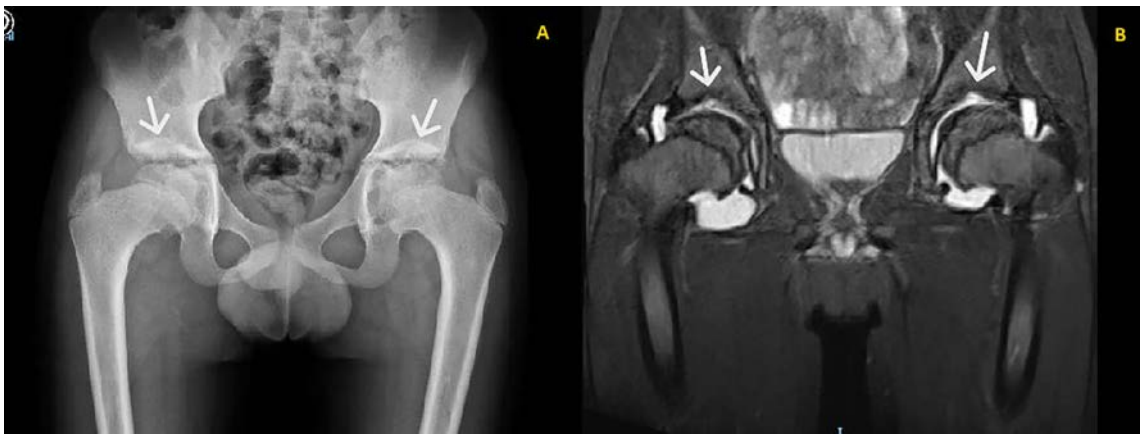

Figure 3. Pelvic radiograph and MRI of the hips. (A) Anteroposterior pelvic radiograph shows bilateral coxa vara with short femoral necks and mild irregularity of the acetabular contours. Bilateral subchondral radiolucent lesions consistent with acetabular cysts are also noted (white arrows). (B) Coronal T2-weighted MRI demonstrates large bilateral acetabular cysts (white arrows) communicating with joint effusions. No erosive changes or pannus formation is identified. MRI, magnetic resonance imaging. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25597/abstract.

Table 1. Comparative table: JIA vs MPS I vs CACP*

Feature	JIA	MPS I	CACP
Age at onset	Before age 16 y	Infancy to early childhood	Early childhood (often <2 y)
Pain	Usually present	Absent or mild	Absent
Morning stiffness	Common	Absent	Absent
Joint involvement	Asymmetric or symmetric, variable	Joint contractures, limited motion	Symmetric, large joints
Synovial fluid	Inflammatory (>5,000 WBCs/mm^3, neutrophilic)	Noninflammatory	Noninflammatory (<1,000 WBCs/mm^3, few neutrophils), presence of multinucleated giant cells on synovial biopsy sample
Markers of inflammation (ESR/CRP)	Often elevated	Normal	Normal
Autoantibodies (ANA, RF, ACPA)	Variable (ANA+, RF+, ACPA+ in subtypes)	No diagnostic value	No diagnostic value
Radiographic findings	Joint space narrowing, erosions in late stages	Dysostosis multiplex, bullet-shaped vertebrae	Coxa vara, acetabular and epiphyseal cysts, hypoplastic patellae
Extra-articular features	Uveitis, systemic symptoms (fever, rash, etc) in the systemic subtype	Coarse facies, hepatosplenomegaly, corneal clouding	Noninflammatory pericarditis
Camptodactyly	No	No	Yes
Genetic testing	Not diagnostic, HLA typing may support	*IDUA* gene mutation, GAGs in urine, enzyme assay	Biallelic *PRG4* mutations

* ACPA, anti–cyclic citrullinated peptide antibody; ANA, antinuclear antibody; CACP, camptodactyly-arthropathy–coxa vara–pericarditis; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; GAG, glycosaminoglycan; JIA, juvenile idiopathic arthritis; MPS I, mucopolysaccharidosis type I; RF, rheumatoid factor; WBC, white blood cell.

Common manifestations include joint pain, swelling, stiffness (especially morning stiffness), and decreased range of motion. Depending on the subtype, systemic features such as fever, rash, lymphadenopathy, or organomegaly may also be present.[5]

Among the subtypes, polyarticular JIA is the most relevant to the present case. It is defined by arthritis affecting five or more joints within the first six months of disease onset and typically presents with symmetrical involvement of both large and small joints. Hip involvement is relatively common in JIA, particularly in systemic and polyarticular forms, with prevalence estimates ranging from 20% to 63% depending on the subtype and disease severity.[6–8] Nevertheless, bilateral hip disease is more frequently a feature of long-standing or severe disease rather than an early finding. For example, in systemic JIA, hip arthritis develops in approximately 51% of patients, with a median onset around 24 months after diagnosis.[6] In polyarticular JIA, bilateral hip involvement has been associated with extended disease duration and poorer long-term outcomes.[9]

Although elevated markers of inflammation such as ESR and CRP are frequently observed in systemic and polyarticular JIA, they are not universally present. Up to 30% to 40% of children with oligoarticular or RF-negative polyarticular JIA may present with normal acute phase reactants at diagnosis, reflecting a lower-grade inflammatory phenotype not always captured by conventional biomarkers.[10,11]

Autoantibodies such as ANA, RF, and ACPA contribute to subclassification and prognostication in JIA but are not required for diagnosis. ANA is positive in up to 60% to 80% of patients with oligoarticular JIA. RF is detected in approximately 5% to 10% of polyarticular JIA cases and is associated with a more aggressive disease course. ACPA is positive in fewer than 15% of patients, most commonly in RF-positive polyarticular JIA.[12–15] Although these autoantibodies can support diagnosis and risk stratification, they lack sufficient sensitivity or specificity to rule out JIA when absent. For instance, ANA is more relevant as a predictor of uveitis than for the diagnosis of arthritis itself.[12]

Synovial fluid analysis in JIA is typically inflammatory, with white blood cell counts ranging from 5,000 to more than 50,000 cells/mm^3, predominantly neutrophils.[15] Histopathologic examination of the synovium often reveals synovial hyperplasia, increased vascularity, and dense inflammatory infiltrates, sometimes with pannus formation.[16,17]

In the present case, JIA was initially considered due to the presence of symmetrical polyarticular joint swelling involving large joints. However, several clinical and laboratory findings argued against this diagnosis. Notably, the episodes were painless, were not associated with morning stiffness or functional impairment, and resolved with return to baseline function between flares, an atypical pattern for active JIA. Markers of inflammation remained within normal limits throughout follow-up, and autoantibodies (ANA, RF, ACPA) were negative. Although these findings do not exclude the diagnosis of JIA, their absence (particularly in the context of clinically silent arthritis) reduces the overall probability of this diagnosis and prompted consideration of noninflammatory mimics.

Additionally, synovial fluid analysis revealed no inflammatory findings, which is inconsistent with JIA. Structural bone abnormalities, such as coxa vara, hypoplastic patellae, epiphyseal

enlargement, and intraosseous cysts, are not typically seen in JIA, especially in early disease stages. Finally, the lack of response to NSAIDs and intra-articular glucocorticoid injections, along with the absence of systemic features, further reduced the likelihood of a JIA diagnosis.

MPS I. MPS I is a lysosomal storage disorder caused by a deficiency of the enzyme α-L-iduronidase, leading to accumulation of glycosaminoglycans (GAGs) such as heparan sulfate and dermatan sulfate in various tissues. It is inherited in an autosomal recessive pattern and presents with a broad phenotypic spectrum that includes severe (Hurler), intermediate (Hurler–Scheie), and attenuated (Scheie) forms.[18]

The severe phenotype typically manifests in infancy with developmental delay, coarse facial features, corneal clouding, hepatosplenomegaly, skeletal dysplasia (dysostosis multiplex), joint contractures, and progressive neurodegeneration. In contrast, the attenuated forms may present later in childhood, with joint stiffness and skeletal deformities as predominant symptoms, sometimes mimicking rheumatologic conditions such as JIA.[19]

Articular manifestations in MPS I can include joint swelling, reduced range of motion, and bony deformities. However, unlike inflammatory arthritides, joint pain and true synovitis are typically absent. Radiographic features often show characteristic findings such as bullet-shaped phalanges, anterior beaking of vertebral bodies, and thickened ribs. These skeletal abnormalities, collectively termed dysostosis multiplex, are key to diagnosis.[20]

Laboratory diagnosis of MPS I involves measurement of urinary GAGs, enzymatic activity of α-L-iduronidase in leukocytes or fibroblasts, and genetic testing of the *IDUA* gene. Diagnosis can be challenging in attenuated forms, in which systemic features are subtle or absent and musculoskeletal symptoms may predominate.[21]

In the present case, MPS I was considered in the differential diagnosis due to the presence of early-onset joint symptoms, skeletal deformities, and the absence of pain or inflammatory signs. Although true camptodactyly is not a hallmark of MPS disorders, joint contractures and flexion deformities, particularly affecting the hands, are common and can mimic a camptodactylous appearance. However, the lack of coarse facial features, corneal clouding, hepatosplenomegaly, or neurodevelopmental impairment made the diagnosis less likely.

CACP. CACP syndrome is a rare autosomal recessive disorder caused by biallelic pathogenic mutations in the *PRG4* gene, which encodes lubricin, a glycoprotein essential for joint lubrication and synovial homeostasis. Loss of lubricin leads to mechanical friction in synovial joints and subsequent synovial hyperplasia without inflammation.[22]

Clinically, CACP is characterized by a triad of noninflammatory arthropathy, congenital or early-onset camptodactyly, and coxa vara. Some patients also develop noninflammatory pericardial effusions, typically presenting later in the disease course.[23] The exact number of pathogenic variants in the *PRG4* gene remains uncertain, but approximately 25 to 35 distinct mutations have been described to date across several cohorts worldwide. Although the true prevalence of CACP syndrome is unknown, these variants account for a limited number of genetically confirmed cases, likely underreported due to limited access to genetic testing and underrecognition of the syndrome.[24]

Camptodactyly, often present at birth or within the first two years of life, is typically bilateral and affects the proximal interphalangeal joints. It is considered a hallmark feature, seen in the majority of reported cases.[25,26] Coxa vara, present in a great number of patients, usually develops gradually and contributes to gait disturbances, leg length discrepancy, and pelvic tilt.[26] Pericardial effusion is less frequent and typically noninflammatory, with serous or serosanguinous content and absence of elevated markers of inflammation or systemic symptoms. This is attributed to the lack of lubricin in the pericardium, which is normally expressed on mesothelial surfaces.[25–27]

Imaging findings are particularly important in guiding diagnosis. One of the most distinctive radiologic features is the presence of large acetabular cysts, which are considered highly suggestive and potentially pathognomonic of CACP when observed in the appropriate clinical setting. These intraosseous cysts represent fluid-filled herniations of the joint capsule into the acetabulum and are rarely, if ever, seen in JIA, except in cases with severe erosive disease. Their presence should prompt strong consideration of CACP, particularly in a child with noninflammatory joint findings.[28] Other characteristic findings include epiphyseal enlargement,[27] hypoplastic patellae,[25] and nonerosive joint spaces with smooth bone contours.[27]

Laboratory tests in CACP are characteristically noninflammatory. Acute phase reactant (ESR and CRP) levels are normal, autoantibodies such as ANA, RF, and ACPA are negative, and synovial fluid analysis reveals low leukocyte counts without signs of inflammation. Histologically, synovial biopsies demonstrate synovial hyperplasia with minimal inflammatory infiltrate and may contain multinucleated giant cells.[28]

In our patient, CACP was considered in the differential diagnosis due to the constellation of features: symmetrical, painless joint swelling beginning in early childhood; congenital camptodactyly; radiographic evidence of coxa vara and dysplastic patellae; persistently normal markers of inflammation; and a noninflammatory synovial profile both cytologically and histologically. The presence of both acetabular and interosseous knee cysts further increased clinical suspicion. The absence of systemic features, negative autoantibodies, and lack of response to anti-inflammatory therapies supported this hypothesis.

Although CACP is most often reported in populations with a high degree of consanguinity, it can occur in nonconsanguineous families and has been described in various ethnic backgrounds.[18]

In our case, although no consanguinity was reported, both parents originated from the same small town, raising the possibility of a shared ancestral mutation.

Nevertheless, some features were not entirely typical. The patient did not initially present with pericardial involvement, a finding that, although not universally present, can provide a key diagnostic clue. Additionally, although coxa vara and patellar abnormalities were identified, other radiographic signs such as advanced joint space narrowing or vertebral changes were absent at early stages.

PATIENT'S COURSE

A chronological summary of the patient's clinical evolution is shown in Table 2. At age six years, the patient was referred due to recurrent joint swelling affecting large joints. Given the early onset, noninflammatory pattern, congenital camptodactyly, progressive coxa vara, persistently normal markers of inflammation, and absence of systemic symptoms, a noninflammatory genetic arthropathy was considered early in the diagnostic process. Initial investigations included screening for metabolic arthropathies. The patient underwent enzymatic testing of α-L-iduronidase activity in leukocytes, urinary GAGs quantification, and genetic testing of the *IDUA* gene. All results were within normal limits or negative, effectively ruling out MPS I.

Simultaneously, the clinical and radiologic phenotype, particularly the camptodactyly, large acetabular and knee cysts, dysplastic patellae, and absence of inflammatory findings in synovial fluid, raised clinical suspicion of CACP syndrome. Targeted genetic testing of the *PRG4* gene was pursued. This revealed a homozygous pathogenic variant, *c.2897_2898del* (*p.Gln966Argfs*11*), confirming the diagnosis of CACP, and the patient began a rehabilitation program.

At age seven years, diagnostic arthroscopy of the left knee was performed, allowing synovial fluid analysis and biopsy. Synovial histology revealed chronic synovial hyperplasia with minimal mononuclear infiltrate and the presence of multinucleated giant cells, supporting a noninflammatory arthropathy. At age eight years, due to persistent joint effusion, the patient underwent arthroscopic synovectomy of both knees, achieving moderate improvement in range of motion and reduction of swelling. At age 11 years, owing to ongoing symptoms, an open synovectomy of both knees was performed by the pediatric traumatology unit. No further episodes of joint effusion have occurred during follow-up. The patient remains under ongoing follow-up with rheumatology, traumatology, cardiology, and physiotherapy specialists. Notably, immunosuppressive treatment was not initiated at any point because it is ineffective in managing this condition, avoiding iatrogenesis.

Table 2. Clinical timeline of the patient*

Age, y	Event
2	Surgery for bilateral trigger fingers.
6	Onset of symmetrical, painless joint swelling. Referred to pediatric rheumatology. Initial laboratory workup reveals normal markers of inflammation and negative autoantibody profile. Synovial fluid analysis is noninflammatory. MRI shows coxa vara and intraosseous cysts.
7	Arthroscopic synovial biopsy of the left knee was performed. Histology shows synovial hyperplasia with multinucleated giant cells. Genetic testing: *IDUA* gene sequencing (to evaluate MPS I) and targeted *PRG4* gene analysis. *IDUA* test result normal. *PRG4* test confirms homozygous pathogenic mutation. Diagnosis of CACP established. Rehabilitation program initiated.
8	First open synovectomy of both knees due to persistent effusion.
9–10	Clinical and functional improvement. Continued follow-up. No immunosuppressive treatment required.
11	Second open synovectomy performed on both knees due to clinical worsening.
12–14	Stable condition under multidisciplinary follow-up (rheumatology, orthopedics, cardiology, and physical therapy). No evidence of pericardial involvement during follow-up.

* CACP, camptodactyly-arthropathy–coxa vara–pericarditis; MPS I, mucopolysaccharidosis type I; MRI, magnetic resonance imaging.

DISCUSSION

Clinical relevance. This case underscores the importance of recognizing characteristic features of CACP syndrome early in the diagnostic process, not only to guide appropriate management but also to avoid unnecessary immunosuppressive treatments and potential iatrogenesis. Although the initial presentation with symmetric joint swelling suggested JIA, the absence of pain, morning stiffness, or functional limitation, together with the lack of systemic symptoms, stable markers of inflammation, and no clinical response to NSAIDs or intra-articular glucocorticoid injections, shifted suspicion toward a noninflammatory genetic arthropathy. The presence of congenital camptodactyly and progressive coxa vara provided early diagnostic clues. However, the most compelling radiologic features were the intraosseous acetabular and knee cysts, which—unlike in JIA—are highly specific and potentially pathognomonic for CACP in the appropriate clinical context. These imaging findings should prompt early consideration of CACP and facilitate timely genetic confirmation, especially in patients with compatible clinical and histologic profiles.

Identifying the syndrome is crucial not only to establish a correct diagnosis but also to avoid ineffective and potentially harmful immunosuppressive treatments. In our patient, diagnosis led to targeted supportive management, including physiotherapy and surgical interventions, which contributed to functional gains and symptom control.

The role of genetic testing in CACP. CACP is an autosomal recessive disease that results from biallelic pathogenic variants in *PRG4*, located on chromosome 1. The *PRG4* gene encodes lubricin, a glycoprotein secreted by synovial fibroblasts and articular chondrocytes that plays a critical role in joint lubrication and synovial homeostasis.[25]

The definitive diagnosis in our patient was established through the identification of a homozygous pathogenic variant in *PRG4* (*c.2897_2898del*; *p.Gln966Argfs*11*), a frameshift mutation predicted to result in loss of protein expression. Most pathogenic *PRG4* variants reported in the literature are nonsense, frameshift, or splice-site mutations, leading to complete absence of functional lubricin.[29]

The absence of lubricin results in increased mechanical friction and impaired protection of articular surfaces, which secondarily leads to synovial thickening and joint degeneration.[30,31] *PRG4* is also expressed on mesothelial surfaces, including the pericardium, which explains the noninflammatory pericardial effusions sometimes seen in affected individuals.[32] Genetic testing is crucial in CACP to confirm the diagnosis, guide clinical management, and avoid unnecessary and ineffective immunosuppressive treatments often initiated under the assumption of JIA.

Implications for clinical management and treatment. The management of CACP focuses on rehabilitation therapy to maintain mobility and prevent joint contractures, orthotic devices to improve biomechanical function and joint stability, and surgical interventions.[33] Studies have suggested that recombinant lubricin could be a promising therapeutic option for CACP. However, clinical trials are required, and challenges regarding localized delivery and bioavailability must be addressed before this becomes a viable treatment option.[33–35] Multidisciplinary follow-up with rheumatology, orthopedics, and cardiology was maintained to monitor disease progression and manage potential complications.

Prognosis and long-term evolution. After the patient underwent rehabilitation and surgical interventions, no new episodes of joint effusion were observed. However, biomechanical alterations (pes cavus, equinus posture, and coxa vara) persisted, requiring continued monitoring and potential future orthopedic interventions. The clinical prognosis remains uncertain for the future.

Teaching points:

- CACP is a rare, noninflammatory arthropathy that should be considered in the differential diagnosis of chronic arthritis in children.
- Genetic testing is crucial for confirming the diagnosis and preventing unnecessary immunosuppressive therapy, which is ineffective in CACP.
- Management requires a multidisciplinary approach, focusing on physiotherapy, orthopedic interventions, and long-term monitoring, rather than pharmacological immunosuppression.

This case highlights the importance of recognizing rare genetic arthropathies in pediatric rheumatology and tailoring treatment strategies accordingly.

FINAL DIAGNOSIS

CACP (camptodactyly-arthropathy–coxa vara–pericarditis syndrome)

ACKNOWLEDGMENT

Funding for open access charge: Universidad de Córdoba/CBUA.

AUTHOR CONTRIBUTIONS

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

REFERENCES

1. Zaripova LN, Midgley A, Christmas SE, et al. Juvenile idiopathic arthritis: from aetiopathogenesis to therapeutic approaches. Pediatr Rheumatol Online J 2021;19(1):135.
2. Petty RE, Southwood TR, Manners P, et al. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. J Rheumatol. 2004;31(2):390–392.
3. Martini A, Lovell DJ, Albani S, et al. Juvenile idiopathic arthritis. Nat Rev Dis Primers 2022;8(1):5.
4. Martini A, Ravelli A, Avcin T, et al; Pediatric Rheumatology International Trials Organization (PRINTO). Toward new classification criteria for juvenile idiopathic arthritis: first steps, Pediatric Rheumatology International Trials Organization International consensus. J Rheumatol 2019;46(2):190–197.
5. Lee JJY, Schneider R. Systemic juvenile idiopathic arthritis. Pediatr Clin North Am 2018;65(4):691–709.
6. Batthish M, Feldman BM, Babyn PS, et al. Predictors of hip disease in the systemic arthritis subtype of juvenile idiopathic arthritis. J Rheumatol 2011;38(5):954–958.
7. Shelmerdine SC, Di Paolo PL, Tanturri de Horatio L, et al. Imaging of the hip in juvenile idiopathic arthritis. Pediatr Radiol 2018;48(6):811–817.
8. Rostom S, Amine B, Bensabbah R, et al. Hip involvement in juvenile idiopathic arthritis. Clin Rheumatol 2008;27(6):791–794.
9. Elhai M, Bazeli R, Freire V, et al. Radiological peripheral involvement in a cohort of patients with polyarticular juvenile idiopathic arthritis at adulthood. J Rheumatol 2013;40(4):520–527.

10. Nordal EB, Zak M, Aalto K, et al. Validity and predictive ability of the juvenile arthritis disease activity score based on CRP versus ESR in a Nordic population-based setting. Ann Rheum Dis 2012;71(7): 1122–1127.
11. Alberdi-Saugstrup M, Zak M, Nielsen S, et al; Nordic Study Group of Pediatric Rheumatology (NoSPeR). High-sensitive CRP as a predictive marker of long-term outcome in juvenile idiopathic arthritis. Rheumatol Int 2017;37(5):695–703.
12. Storwick JA, Brett AC, Buhler K, et al. Prevalence and titres of antinuclear antibodies in juvenile idiopathic arthritis: a systematic review and meta-analysis. Autoimmun Rev 2022;21(6):103086.
13. Peckham H, Cambridge G, Bourke L, et al. Antibodies to cyclic citrullinated peptides in patients with juvenile idiopathic arthritis and patients with rheumatoid arthritis: shared expression of the inherently autoreactive 9G4 idiotype. Arthritis Rheumatol 2017;69(7):1387–1395.
14. Spârchez M, Miu N, Bolba C, et al. Evaluation of anti-cyclic citrullinated peptide antibodies may be beneficial in RF-negative juvenile idiopathic arthritis patients. Clin Rheumatol 2016;35(3):601–607.
15. Wang Y, Pei F, Wang X, et al. Meta-analysis: diagnostic accuracy of anti-cyclic citrullinated peptide antibody for juvenile idiopathic arthritis. J Immunol Res 2015;2015:915276.
16. Metzemaekers M, Malengier-Devlies B, Yu K, et al. Synovial fluid neutrophils from patients with juvenile idiopathic arthritis display a hyperactivated phenotype. Arthritis Rheumatol 2021;73(5):875–884.
17. Finnegan S, Clarke S, Gibson D, et al. Synovial membrane immunohistology in early untreated juvenile idiopathic arthritis: differences between clinical subgroups. Ann Rheum Dis 2011;70(10):1842–1850.
18. Hampe CS, Eisengart JB, Lund TC, et al. Mucopolysaccharidosis type I: a review of the natural history and molecular pathology. Cells 2020;9(8):1838.
19. Morishita K, Petty RE. Musculoskeletal manifestations of mucopolysaccharidoses. Rheumatology (Oxford) 2011;50(suppl 5):v19–v25.
20. White KK. Orthopaedic aspects of mucopolysaccharidoses. Rheumatology (Oxford) 2011;50(suppl 5):v26–v33.
21. Lehman TJ, Miller N, Norquist B, et al. Diagnosis of the mucopolysaccharidoses. Rheumatology (Oxford) 2011;50(suppl 5):v41–v48.
22. Rhee DK, Marcelino J, Baker M, et al. The secreted glycoprotein lubricin protects cartilage surfaces and inhibits synovial cell overgrowth. J Clin Invest 2005;115(3):622–631.
23. Waller KA, Zhang LX, Jay GD. Friction-induced mitochondrial dysregulation contributes to joint deterioration in Prg4 knockout mice. Int J Mol Sci 2017;18(6):1252.
24. Nandagopalan RS, Phadke SR, Dalal AB, et al. Novel mutations in PRG4 gene in two Indian families with camptodactyly-arthropathy-coxa vara-pericarditis (CACP) syndrome. Indian J Med Res 2014;140(2):221–226.
25. Albuhairan I, Al-Mayouf SM. Camptodactyly-arthropathy-coxavara-pericarditis syndrome in Saudi Arabia: clinical and molecular genetic findings in 22 patients. Semin Arthritis Rheum 2013;43(2):292–296.
26. El-Garf A, Mahmoud G, Gheith R, et al. Camptodactyly, arthropathy, coxa vara, and pericarditis syndrome among Egyptians. J Rheumatol 2003;30(5):1081–1086.
27. Johnson N, Chaudhary H, Kumrah R, et al. Syndrome of progressive deforming non-inflammatory arthritis of childhood: two patients of camptodactyly-arthropathy-coxa vara-pericarditis syndrome. Rheumatol Int 2021;41(10):1875–1882.
28. Alazami AM, Al-Mayouf SM, Wyngaard CA, et al. Novel PRG4 mutations underlie CACP in Saudi families. Hum Mutat 2006;27(2):213.
29. Ciullini Mannurita S, Vignoli M, Bianchi L, et al. CACP syndrome: identification of five novel mutations and of the first case of UPD in the largest European cohort. Eur J Hum Genet 2014;22(2):197–201.
30. Maenohara Y, Chijimatsu R, Tachibana N, et al. Lubricin contributes to homeostasis of articular cartilage by modulating differentiation of superficial zone cells. J Bone Miner Res 2021;36(4):792–802.
31. Marcelino J, Carpten JD, Suwairi WM, et al. CACP, encoding a secreted proteoglycan, is mutated in camptodactyly-arthropathy-coxa vara-pericarditis syndrome. Nat Genet 1999;23(3): 319–322.
32. Murphy JM, Vanderhave KL, Urquhart AG. Total hip arthroplasty in adolescents with severe hip arthropathy and dysplasia associated with camptodactyly-arthropathy-coxa vara-pericarditis syndrome. J Arthroplasty 2012;27(8):1581.e5–1581.e8.
33. Bao JP, Chen WP, Wu LD. Lubricin: a novel potential biotherapeutic approaches for the treatment of osteoarthritis. Mol Biol Rep 2011; 38(5):2879–2885.
34. Ai M, Cui Y, Sy MS, et al. Anti-lubricin monoclonal antibodies created using lubricin-knockout mice immunodetect lubricin in several species and in patients with healthy and diseased joints. PLoS One 2015; 10(2):e0116237.
35. Yuan H, Mears LLE, Liu X, et al. Recombinant lubricin improves antiadhesive, wear protection, and lubrication of collagen II surface. Colloids Surf B Biointerfaces 2022;220:112906.

Arthritis Care & Research
Vol. 77, No. 12, December 2025, pp 1403–1408
DOI 10.1002/acr.25602

American College of Rheumatology
Empowering Rheumatology Professionals

REVIEW ARTICLE

Reconsidering Race-Based Medicine in Pediatric Rheumatology: Challenges and Opportunities for Equitable Care

Nayimisha Balmuri,[1] Alisha Akinsete,[2] Laura B. Lewandowski,[3] Mallet Reid,[4] Jordan E. Roberts,[5] and Jennifer M. P. Woo[6]

Despite growing evidence of their limitations, race-based practices in pediatric rheumatology—those that rely on race or ethnicity to influence diagnosis and treatment—continue to shape care, often reinforcing health disparities. The assumption that biologic or genetic differences exist between racial groups oversimplifies complex health issues and perpetuates health inequities. This article examines persistent race-based practices in pediatric rheumatology, particularly in the interpretation of laboratory results and clinical decision-making, and highlights their clinical limitations. For example, the use of race-adjusted formulas in evaluating estimated glomerular filtration rate, pulmonary function tests, and creatine kinase levels can lead to misdiagnoses and delayed interventions, particularly in Black and Asian populations. Additionally, race-based assumptions in diseases like Kawasaki disease and multisystem inflammatory syndrome in children can lead to incorrect conclusions about disease severity and treatment efficacy. This article advocates for a shift toward race-conscious practices that consider the role of social determinants of health and biases in clinical care. It also emphasizes the need for more inclusive research methodologies and diverse representation in clinical trials to enhance the generalizability of findings. By moving away from race-based practices and adopting equity-oriented frameworks, pediatric rheumatologists can better address the needs of marginalized populations and improve health outcomes. This shift is crucial in dismantling systemic disparities and advancing health equity in clinical and research settings.

Introduction

Race-based medicine refers to using a patient's race or ethnicity to guide medical treatment or diagnosis, often based on the assumption that distinct biologic or genetic differences exist among different racial or ethnic groups. This approach is problematic for several reasons. First, it oversimplifies complex health issues by treating race as a biologic determinant, when in fact, race is a social construct that fails to account for individual patient variations. Additionally, relying on race can perpetuate stereotypes and biases, leading to unequal treatment and reinforcing health disparities rather than addressing their root causes.

Furthermore, focusing on race can divert attention from critical social determinants of health, such as socioeconomic status, access to care, and environmental factors, which are often more influential in determining health outcomes.[1–3] Epidemiologic data demonstrating varying rates in those of different races and ethnicities are frequently highlighted in pediatric rheumatology textbooks.[4] Such demographic data may be pertinent in understanding the unequal burden of disease. However, particularly when derived from single-center experiences, which may reflect regional care practices or disparities in access to care, it is too often presented due to genetic predisposition. This misleading data collection and interpretation ultimately hinders the development of effective treatments and public health strategies.

The opinions and assertions contained herein are those of the authors and do not necessarily represent the view of the NIH, the Department of Health and Human Services, or the US government.

Dr Lewandowski's work was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases Intramural Program, NIH (grant ZIA-AR-041214-04). Dr Woo's work was supported by the Intramural Research Program of the National Institute of Environmental Health Sciences, NIH.

[1]Nayimisha Balmuri, MD: Johns Hopkins School of Medicine, Baltimore, Maryland; [2]Alisha Akinsete, MD: Weill Cornell Medicine and Hospital for Special Surgery, New York, New York; [3]Laura B. Lewandowski, MD, MS: Bethesda, Maryland; [4]Mallet Reid, PhD, LMSW: Michigan State University College of Human Medicine, East Lansing, Michigan; [5]Jordan E. Roberts, MD, MPH: University of Washington School of Medicine, Seattle; [6]Jennifer M. P. Woo, PhD, MPH: Durham, North Carolina.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25602.

Address correspondence via email to Nayimisha Balmuri, MD, at nbalmur1@jh.edu.

Submitted for publication January 7, 2025; accepted in revised form June 23, 2025.

Steeping our clinical teaching in supposed differences between children of different races and ethnicities may contribute to the development of clinical practices and a research agenda in which race is inappropriately deemed an important driver of developing a disease and its severity. This article describes common race-based practices in pediatric rheumatology that have persisted despite evidence suggesting that they are not clinically appropriate and calls for a re-evaluation of standard-of-care practices and to inform the development of more equitable practices.[5,6]

Examples in laboratory result interpretation

The estimated glomerular filtration rate (eGFR) uses creatine for its calculations; however, Black patients are assigned higher baseline normal levels of creatinine, leading to an overestimation of renal function using the Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) eGFR calculation or the Modification of Diet in Renal Disease eGFR equation[7–10] and thus a delayed diagnosis of kidney disease or incorrect chronic kidney disease classification and contributing to delayed renal transplant. This may particularly impact Black youth with childhood-onset systemic lupus erythematosus, as higher prevalence of renal disease, worse renal outcomes, and a lower likelihood of receiving a renal transplant have been reported for Black children with lupus nephritis.[11–16] In 2020, the National Kidney Foundation and the American Society of Nephrology announced a task force that discussed this issue and identified a patient-centered solution that did not include race within the calculation equation. The task force supports using the CKD-EPIcr_R equation in all laboratory tests and recommends national efforts to increase the routine and timely use of cystatin C as an additional filtration marker.[10]

Laboratory reference values that have higher creatine kinase (CK) reference values in Black patients in comparison with other populations due to an assumption that Black individuals as a whole have a higher skeletal muscle mass and a higher proportion of fast-twitch muscle fibers compared with other groups of individuals[17,18] may also be detrimental to pediatric rheumatology patients if clinicians potentially dismiss early active myositis or other causes of elevation of CK levels due to race reference values.

Pulmonary function tests (PFTs) use a race correction for Black and Asian patients' reported values,[19] which can result in the misdiagnosis and misclassification of patients, as seen in an observational asthma study in children.[20] In a single-center, cross-sectional study, the race-neutral global reference equations were compared with the race- and ethnicity-specific race equation, which showed that using the race-neutral equation increased the prevalence of respiratory impairment in Black patients.[21] Many rheumatologic conditions such as systemic juvenile idiopathic arthritis, dermatomyositis, vasculitis, sarcoidosis, mixed connective tissue disease, and systemic sclerosis can significantly impair a patient's pulmonary function. It is possible that ensuring the use of PFT laboratory tests that use race-neutral global reference equations in the analysis of PFT results could lead to earlier detection of pulmonary disease.

Differences seen in hemoglobin reference levels also potentially underestimate anemia. In an end-stage renal disease management study, Black and Hispanic patients had lower hemoglobin levels and were less likely to receive erythropoietin-stimulating agents.[22,23] Black patients' reference values for hemoglobin baseline tend to be lower than those of White age- and sex-matched cohorts. The lower reference value seen in Black patients has been attributed to higher rates of diabetes, hypertension, kidney disease, beta or alpha thalassemia trait, sickle cell trait, fibroids, heavy menstrual bleeding, chronic stress, and poor access to nutrition-rich foods (iron and B12).[22]

Examples in clinical medicine

Kawasaki disease was first described in Japan. Despite a very high level of randomized clinical trial evidence for glucocorticoids in Japanese children with Kawasaki disease,[24–26] these practices have not been widely implemented in the United States and elsewhere due to an assumption that non-Asian races may not benefit from this same treatment. Currently, several clinical algorithms include Asian race in determining whether a child receives glucocorticoids when other risk factors for severe disease and coronary artery aneurysm are present.[27,28] This has been extrapolated to infer that those of non-Asian race are at lower risk, including the suggestion that Black race may be a protective factor for coronary artery aneurysms among youths with Kawasaki disease.[29] A small single-center study, which took place at a hospital that served predominantly Black patients, reported lower rates of coronary artery aneurysms in Black children.[29] Black children in this cohort did not experience diagnostic delays, as have been reported elsewhere, suggesting that timely treatment may drive better outcomes.[30]

This tendency to prioritize genetic differences as a potential driver of observed differences in the racial burden of disease was highlighted in the pediatric rheumatology community's response to the development of multisystem inflammatory syndrome in children (MIS-C). Rates of SARS-CoV-2 infection were higher in historically marginalized groups across the globe, including higher rates in Black, Hispanic, and Indigenous populations.[31–33] Social determinants of health including essential worker status, which vary by social class, income, and race, were also shown to impact SARS-CoV-2 infection rates.[34] However, the higher rates of MIS-C in non-White racial groups were repeatedly reported as disproportionate and speculated to have a biologic origin, despite the lack of plausible shared genetic ancestry among such geographic and ancestrally disparate groups.[35,36] Much of the pediatric rheumatology literature continued to speculate about genetic drivers long after epidemiologic data were available to suggest that rates of MIS-C were proportionate to

the burden of acute SARS-CoV-2 infection in these communities.[37] Perhaps in response to this disproportionate focus on biologic differences underlying disparities in outcomes, tremendous resources were devoted to the research of genetic risk factors, in comparison with resources dedicated to addressing inequities in SARS-CoV-2 exposure, vaccination, and testing and treatment of acute infection.

Clinical solutions

The pediatric rheumatology clinic has served as the primary stage on which race-based practices have been implemented without critically questioning their source and scientific rigor. To rectify this, clinical practice in pediatric rheumatology must undergo an intentional shift in redefining standard-of-care practices to ensure that the needs of the diverse populations affected by pediatric rheumatologic diseases are adequately and equitably addressed.

Within the clinical sphere, clinicians can serve as race-equity champions and question when persistent practices may have limited evidence and be based in historically unfounded beliefs and systemic practices. A policy statement by the American Academy of Pediatrics proposed that pediatricians (1) integrate health equity in their continuing medical education, as well as content on the elimination of race-based medicine in lifelong learning and maintenance of certification efforts, and (2) assess their clinical practices for race-based practices and eliminate race-based medicine from their care delivery.[6]

We encourage clinicians and clinical researchers to consider the following points when assessing current practices or developing or applying new or revised clinical algorithms:

1. Whether the need for race correction is based on robust evidence and statistical analyses (ie, with consideration of internal and external validity, potential confounders, and bias), including study samples that are representative of the patient populations for which the clinical algorithm is to be applied.
2. Does the causal mechanism for the proposed "racial" difference justify the application of the race correction or are there other criteria, such as other social determinants, that may better explain the distribution of hypothesized efficacy?
3. What are the health equity implications of implementing the race correction? Will it attenuate or exacerbate existing health inequities?

Examples in research

Although diverse representation in clinical trials is necessary for equitable access to new therapeutics and inclusion throughout the spectrum of research, researchers and clinicians must also be cautious of the use of race and ethnicity in reporting medication efficacy. In the phase 3 Aspreva Lupus Management Study comparing intravenous cyclophosphamide with mycophenolate, a small sample of Black patients who had more renal scarring and hypertension were used to show inferiority of cyclophosphamide compared with mycophenolate among Black patients, but not among other races.[38,39] Within the article, the authors state that "the trial was not designed to be powered to detect an effect of a specific region, race or ethnicity"[39]; however, clinically, rheumatologists may have misinterpreted the data and made the misleading conclusion that Black patients do not respond as well to cyclophosphamide treatments.

Belimumab, one of the first targeted therapies approved for lupus in decades, included relatively few Black patients in its initial successful trial.[40] Two post hoc analyses of these small samples showed conflicting efficacy results. Therefore, due to the lack of clear scientific rationale for why a successful trial in a mixed racial group would not demonstrate similar efficacy among Black patients, a second, smaller trial was undertaken with the same endpoints and exclusively enrolled Black patients. However, this trial failed to meet reductions in disease activity, safety, and efficacy, although overall numeric trends toward improvement were similar to prior studies.[40,41]

When subgroups of racial groups are very small, analyses may be underpowered and researchers should therefore avoid making unsupported inferences. This is particularly worrisome when a non–statistically significant trend, in the absence of a plausible biologic mechanism, is interpreted as evidence that a drug may be less efficacious in patients of some races (typically those from historically marginalized groups, as White patients are rarely underrepresented in clinical trials). Indeed, care must be taken to decide whether to compare across racial and ethnic subgroups at all, as a perception of lower efficacy may contribute to differential access to new therapeutics among patients belonging to minoritized racial groups. Power calculations should guide analyses to minimize spurious and misleading results.

Research solutions

The development of revised standard-of-care practices will need to be based on rigorous research practices to prevent additional racial, ethnic, and other identity-based disparities. There is a rich body of scholarly work that can be used to frame efforts to enhance equity from a research perspective.

Methodologies. The development of alternatives to race-based practices in pediatric rheumatology will require input from diverse research and clinical teams to increase representativeness across participants and researchers. Research methods that focus on the voice and participation of individuals from marginalized communities should be implemented to ensure that the resulting research is representative of the lived experiences and

disease progression of the most affected populations. These could be fundamental in identifying alternatives to current race-based practices in pediatric rheumatology.

- Community-based participatory approaches: community engagement with those who are the focus of the research to aid in the development of, implementation of, and insight generation into the research process.[42]
- Explanatory- or exploratory-designed mixed-methods approach: define a problem either quantitatively or qualitatively first, then use results to implement a second-phase study that addresses the problems defined in the first phase.

Equity-oriented frameworks. Equity-oriented frameworks are approaches to knowledge production that ask questions, develop research projects, and analyze and disseminate data from a social justice value orientation. These frameworks are useful for countering science's complicity with racist oppression and assist researchers in identifying and correcting agendas in research that support white supremacy.[43] Although many equity-oriented frameworks have been developed, feminist standpoint theory and intersectionality represent two valuable frameworks for enhancing racial equity in pediatric rheumatology practice and research.

- Intersectionality in medical research is a framework that recognizes how various aspects of a person's identity—such as race, gender, age, sexual orientation, socioeconomic status, and disability—interact to shape their health experiences and outcomes. For example, structural-level issues such as racism and sexism influence individual-level outcomes in health in impacted people.[44,45]
- Feminist standpoint theory in medical research focuses on understanding health issues from the perspectives of women and marginalized groups. It argues that people's social backgrounds, such as gender, race, and economic status, shape experiences and insights about health and challenges the idea of objective knowledge in medicine, suggesting that traditional medical research often reflects the biases of those in power, primarily White, male researchers.[43,46–48]

Publication. Publication of scientific work can be a space in which race-based practices are perpetuated and therefore serves as a point of intervention that can be leveraged to support the dissemination and adoption of equitable clinical practices in pediatric rheumatology. A recent commentary by editors of multiple biomedical journals has highlighted the need for systematic changes in the reporting of race and ethnicity in genetic and genomic research.[49] Multiple leading biomedical journals have implemented protocols around the acknowledgment of race and ethnicity as social constructs and the reporting of race and ethnicity in their publication guidelines.[50,51] A review by Gilliam et al reported that almost 50% of 126 identified published US pediatric clinical practice guidelines contained language around race that could negatively affect health care inequities.[52] A primary barrier to the dissemination of equitable practices is the inconsistent practices around reporting and interpreting results based on race and ethnicity. Researchers should carefully consider how race and ethnicity are presented, how results stratified by race are interpreted, the representativeness of the participant sample, the impact of missing data, and other indicators that may reduce the generalizability of results.[53,54] Researchers should be careful not to conclude that outcomes are based on biology when systemic and structural racism may be driving the disparities noted in research studies. Ensuring rigorous research and publication practices in the effort to inform clinical practice can support equitable practices in the clinic. However, elimination of race-based medicine does not mean that we should not continue to collect race and ethnicity data and highlight and increase diversity in research.

Conclusions

In this article, we promote alternatives to race-based practices—namely, race-conscious practices in which clinicians consider the racialized systems that contribute to disease presentation and racialized biases in care.[54,55] Addressing and eliminating racial and ethnic health inequities associated with race-based practices in pediatric rheumatology requires a willingness to dismantle practices ingrained in the pediatric rheumatology pedagogy and to design and conduct research that prioritizes participant representation as a means to promote scientific rigor. We hope that by taking steps toward race-conscious practices and away from race-based practice in the clinic and research, we will continue to advance health equity and the field of pediatric rheumatology.

ACKNOWLEDGMENT

Although we acknowledge the importance of inclusive language, we aim to communicate in a manner that prioritizes clarity and accessibility for a broad audience and utilizes terminology used in referenced studies. We recognize that terminology can evolve, and we encourage readers to engage critically with the content and consider the context in which language is used. Our commitment is to foster understanding and respect for all individuals, regardless of their racial or ethnic backgrounds.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Dr Balmuri confirms that all authors have provided the final approval of the version to be published and takes responsibility for the

affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. Amutah C, Greenidge K, Mante A, et al. Misrepresenting race - the role of medical schools in propagating physician bias. N Engl J Med 2021;384(9):872–878.
2. Lujan HL, DiCarlo SE. The racist "one drop rule" influencing science: it is time to stop teaching "race corrections" in medicine. Adv Physiol Educ 2021;45(3):644–650.
3. Smedley A, Smedley BD. Race as biology is fiction, racism as a social problem is real: anthropological and historical perspectives on the social construction of race. Am Psychol 2005;60(1):16–26.
4. Petty RE, Laxer RM, Lindsley CB, et al. Textbook of Pediatric Rheumatology. Elsevier Health Sciences; 2020.
5. Cerdeña JP, Plaisime MV, Tsai J. From race-based to race-conscious medicine: how anti-racist uprisings call us to act. Lancet 2020; 396(10257):1125–1128.
6. Wright JL, Davis WS, Joseph MM, et al; AAP Board Committee on Equity. Eliminating race-based medicine. Pediatrics 2022;150(1): e2022057998.
7. Levey AS, Stevens LA, Schmid CH, et al; CKD-EPI (Chronic Kidney Disease Epidemiology Collaboration). A new equation to estimate glomerular filtration rate. Ann Intern Med 2009;150(9):604–612.
8. Eneanya ND, Yang W, Reese PP. Reconsidering the consequences of using race to estimate kidney function. JAMA 2019;322(2):113–114.
9. Yan G, Nee R, Scialla JJ, et al. Estimation of Black-White disparities in CKD outcomes: comparison using the 2021 versus the 2009 CKD-EPI creatinine equations. Am J Kidney Dis 2022;80(3):423–426.
10. Delgado C, Baweja M, Crews DC, et al. A unifying approach for GFR estimation: recommendations of the NKF-ASN Task Force on reassessing the inclusion of race in diagnosing kidney disease. Am J Kidney Dis 2022;79(2):268–288.e1.
11. Hiraki LT, Feldman CH, Liu J, et al. Prevalence, incidence, and demographics of systemic lupus erythematosus and lupus nephritis from 2000 to 2004 among children in the US Medicaid beneficiary population. Arthritis Rheum 2012;64(8):2669–2676.
12. Hiraki LT, Benseler SM, Tyrrell PN, et al. Ethnic differences in pediatric systemic lupus erythematosus. J Rheumatol 2009;36(11):2539–2546.
13. Alarcón GS, McGwin G Jr, Bartolucci AA, et al; LUMINA Study Group. Lupus in Minority Populations, Nature versus Nurture. Systemic lupus erythematosus in three ethnic groups. IX. Differences in damage accrual. Arthritis Rheum 2001;44(12):2797–2806.
14. Chang JC, Sears C, Torres V, et al. Racial disparities in renal outcomes over time among hospitalized children with systemic lupus erythematosus. Arthritis Rheumatol 2022;74(8):1430–1439.
15. Tucker LB, Uribe AG, Fernández M, et al. Adolescent onset of lupus results in more aggressive disease and worse outcomes: results of a nested matched case-control study within LUMINA, a multiethnic US cohort (LUMINA LVII). Lupus 2008;17(4):314–322.
16. Hiraki LT, Lu B, Alexander SR, et al. End-stage renal disease due to lupus nephritis among children in the US, 1995-2006. Arthritis Rheum 2011;63(7):1988–1997.
17. George MD, McGill NK, Baker JF. Creatine kinase in the U.S. population: impact of demographics, comorbidities, and body composition on the normal range. Medicine (Baltimore) 2016;95(33): e4344.
18. Black HR, Quallich H, Gareleck CB. Racial differences in serum creatine kinase levels. Am J Med 1986;81(3):479–487.
19. Braun L. Race, ethnicity and lung function: a brief history. Can J Respir Ther 2015;51(4):99–101.
20. Burbank AJ, Atkinson CE, Espaillat AE, et al. Race-specific spirometry equations may overestimate asthma control in Black children and adolescents. Respir Res 2023;24(1):203.
21. Moffett AT, Bowerman C, Stanojevic S, et al. Global, race-neutral reference equations and pulmonary function test interpretation. JAMA Netw Open 2023;6(6):e2316174.
22. Merz LE, Siad FM, Creary M, et al. Laboratory-based inequity in thrombosis and hemostasis: review of the evidence. Res Pract Thromb Haemost 2023;7(2):100117.
23. Weisbord SD, Fried LF, Mor MK, et al. Associations of race and ethnicity with anemia management among patients initiating renal replacement therapy. J Natl Med Assoc 2007;99(11):1218–1226.
24. Kobayashi T, Saji T, Otani T, et al; RAISE study group investigators. Efficacy of immunoglobulin plus prednisolone for prevention of coronary artery abnormalities in severe Kawasaki disease (RAISE study): a randomised, open-label, blinded-endpoints trial. Lancet 2012; 379(9826):1613–1620.
25. Kobayashi T, Inoue Y, Otani T, et al. Risk stratification in the decision to include prednisolone with intravenous immunoglobulin in primary therapy of Kawasaki disease. Pediatr Infect Dis J 2009;28(6): 498–502.
26. Inoue Y, Okada Y, Shinohara M, et al. A multicenter prospective randomized trial of corticosteroids in primary therapy for Kawasaki disease: clinical course and coronary artery outcome. J Pediatr 2006; 149(3):336–341.
27. Gorelik M, Chung SA, Ardalan K, et al. 2021 American College of Rheumatology/Vasculitis Foundation guideline for the management of Kawasaki disease. Arthritis Care Res (Hoboken) 2022;74(4): 538–548.
28. Son MBF, Gauvreau K, Tremoulet AH, et al. Risk model development and validation for prediction of coronary artery aneurysms in Kawasaki disease in a North American population. J Am Heart Assoc 2019; 8(11):e011319.
29. Porcalla AR, Sable CA, Patel KM, et al. The epidemiology of Kawasaki disease in an urban hospital: does African American race protect against coronary artery aneurysms? Pediatr Cardiol 2005;26(6): 775–781.
30. Padilla LA, Collins JL, Idigo AJ, et al. Kawasaki disease and clinical outcome disparities among Black children. J Pediatr 2021;229:54–60.e2.
31. Wu SL, Mertens AN, Crider YS, et al. Substantial underestimation of SARS-CoV-2 infection in the United States. Nat Commun 2020; 11(1):4507.
32. Gaffney AW, Himmelstein D, Bor D, et al. Home sick with coronavirus symptoms: a national study, April-May 2020. J Gen Intern Med 2020; 35(11):3409–3412.
33. Acosta AM, Garg S, Pham H, et al. Racial and ethnic disparities in rates of COVID-19-associated hospitalization, intensive care unit admission, and in-hospital death in the United States from March 2020 to February 2021. JAMA Netw Open 2021;4(10):e2130479.
34. Goyal MK, Simpson JN, Boyle MD, et al. Racial and/or ethnic and socio-economic disparities of SARS-CoV-2 infection among children. Pediatrics 2020;146(4):e2020009951.
35. Stierman B, Abrams JY, Godfred-Cato SE, et al. Racial and ethnic disparities in multisystem inflammatory syndrome in children in the United States, March 2020 to February 2021. Pediatr Infect Dis J 2021;40(11):e400–e406.
36. Feldstein LR, Rose EB, Horwitz SM, et al. Multisystem inflammatory syndrome in U.S. children and adolescents. N Engl J Med 2020; 383(4):334–346.

37. Encinosa W, Moon K, Figueroa J, et al. Complications, adverse drug events, high costs, and disparities in multisystem inflammatory syndrome in children vs COVID-19. JAMA Netw Open 2023;6(1): e2244975.

38. Appel GB, Contreras G, Dooley MA, et al; Aspreva Lupus Management Study Group. Mycophenolate mofetil versus cyclophosphamide for induction treatment of lupus nephritis. J Am Soc Nephrol 2009; 20(5):1103–1112.

39. Isenberg D, Appel GB, Contreras G, et al. Influence of race/ethnicity on response to lupus nephritis treatment: the ALMS study. Rheumatology (Oxford) 2010;49(1):128–140.

40. Furie R, Petri M, Zamani O, et al; BLISS-76 Study Group. A phase III, randomized, placebo-controlled study of belimumab, a monoclonal antibody that inhibits B lymphocyte stimulator, in patients with systemic lupus erythematosus. Arthritis Rheum 2011;63(12):3918–3930.

41. Ginzler E, Guedes Barbosa LS, D'Cruz D, et al. Phase III/IV, randomized, fifty-two-week study of the efficacy and safety of belimumab in patients of Black African ancestry with systemic lupus erythematosus. Arthritis Rheumatol 2022;74(1):112–123.

42. Brockman TA, Shaw O, Wiepert L, et al. Community engagement strategies to promote recruitment and participation in clinical research among rural communities: a narrative review. J Clin Transl Sci 2023; 7(1):e84.

43. Harding S. Science and Social Inequality: Feminist and Postcolonial Issues. University of Illinois Press; 2006.

44. Collins PH. Learning from the outsider within: the sociological significance of Black feminist thought. Soc Probl 1986;33(6):s14–s32.

45. Collins PH. Bilge S. Intersectionality. John Wiley & Sons; 2020.

46. Cho S, Crenshaw KW, McCall L. Toward a field of intersectionality studies: theory, applications, and praxis. Signs (Chic) 2013;38(4): 785–810.

47. Harding S. Whose Science? Whose Knowledge? Thinking From Women's Lives. Cornell University Press; 1991.

48. Harding SG. The Science Question in Feminism. Cornell University Press; 1986.

49. Feero WG, Steiner RD, Slavotinek A, et al. Guidance on use of race, ethnicity, and geographic origin as proxies for genetic ancestry groups in biomedical publications. JAMA 2024;331(15):1276–1278.

50. Woo JMP, Simmonds F, Dennos A, et al; CARRA Registry investigators. Health equity implications of missing data among youths with childhood-onset systemic lupus erythematosus: a proof-of-concept study in the Childhood Arthritis and Rheumatology Research Alliance Registry. Arthritis Care Res (Hoboken) 2023;75(11):2285–2294.

51. Flanagin A, Frey T, Christiansen SL; AMA Manual of Style Committee. Updated guidance on the reporting of race and ethnicity in medical and science journals. JAMA 2021;326(7):621–627.

52. Gilliam CA, Lindo EG, Cannon S, et al. Use of race in pediatric clinical practice guidelines: a systematic review. JAMA Pediatr 2022;176(8): 804–810.

53. Swilley-Martinez ME, Coles SA, Miller VE, et al. "We adjusted for race": now what? A systematic review of utilization and reporting of race in American Journal of Epidemiology and Epidemiology, 2020-2021. Epidemiol Rev 2023;45(1):15–31.

54. AMA Manual of Style Committee. AMA Manual of Style: A Guide for Authors and Editors. Oxford University Press; 2020.

55. Cerdeña J, Plaisime MV, Belcher HME, et al. A PLAN for race-conscious medicine in pediatrics. Pediatrics 2024;153(3): e2023061893.

Arthritis Care & Research
Vol. 77, No. 12, December 2025, pp 1409–1417
DOI 10.1002/acr.25588

American College of Rheumatology
Empowering Rheumatology Professionals

Development and External Validation of a Genetic Risk Score for Pain in Rheumatoid Arthritis

Katie J. McMenamin,[1] Thomas R. Riley IV ,[2] Kristin Wipfler,[3] Kaleb Michaud,[4] Austin Wheeler,[5] Bryant R. England,[5] Brian Sauer,[6] Katherine D. Wysham,[7] Rui Xiao,[8] Michael March,[9] Grant W. Cannon,[6] Sylvanus Toikumo,[2] Henry R. Kranzler,[2] Rachel L. Kember,[2] Ted R. Mikuls,[5] and Joshua F. Baker[2]

Objective. Several single-nucleotide polymorphisms (SNPs) have been associated with chronic pain syndromes. Our objective was to determine whether genetic variants are associated with pain and disease activity in rheumatoid arthritis (RA).

Methods. Participants were included from two independent RA cohorts: FORWARD (National Databank for Rheumatic Diseases, training data set) and the Veterans Affairs Rheumatoid Arthritis Registry (VARA; validation data set). Multivariable linear regression was used to estimate the relationship between cross-sectional pain scores and 36 fibromyalgia (FM)–associated SNPs in FORWARD. SNP alleles were summed and weighted by these regression coefficients to generate a genetic risk score (GRS) for pain for each participant in both cohorts. Linear regressions and generalized estimating equations were used to determine the relationship between this GRS, an existing pain intensity GRS, and pain and self-reported disease activity.

Results. The sample comprised 756 participants from FORWARD (mean age 56.8 years, 89.4% female) and 2,176 participants from VARA (mean age 64.3 years, 11.0% female) who had pain and genotyping data. Participants in the validation data set (VARA) with FM GRS in the highest quartile had more baseline pain than those in the lowest quartile (+0.55 [95% confidence interval 0.16–0.93], $P = 0.006$). This was also true for the existing pain intensity GRS. VARA participants in the highest quartile of both GRS had more pain throughout follow-up and higher disease activity scores.

Conclusion. GRS based on pain-related SNPs were associated with RA pain and disease activity, suggesting that the genetic risk of pain may have clinical impacts in RA, such as the likelihood of achieving remission.

INTRODUCTION

Rheumatoid arthritis (RA) is the most common autoimmune inflammatory joint disease, affecting approximately 1.3 million Americans.[1] People with RA often identify pain as their most significant symptom, with major impacts on quality of life and physical and psychosocial functioning.[2,3] Pain in RA is multifactorial and is not always correlated with levels of joint inflammation, making remission and disease activity targets difficult to achieve.[2,3] There is a greater burden of fibromyalgia (FM) among people with RA, with an estimated prevalence of 18% to 24%, compared to 2% to 4% in the general population.[4] Patients with RA with comorbid FM have more severe symptoms, higher disease activity scores and worse quality of life outcomes, and take more

The contents of this article are solely the responsibility of the authors and do not necessarily represent the views of the Department of Veterans Affairs or the United States Government.

Dr Baker's work was supported by a VA Clinical Science Research and Development Career Merit Award (grant I01-CX-001703) and a Rehabilitation Research and Development Merit Award (grant I01-RX-003644).

U.S. Department of Defense: PR200793; Rehabilitation Research and Development Service: I01 RX003644; U.S. Department of Veterans Affairs: I01 CX001703, IK2 CX002203; National Institutes of Health: 5T32GM075766-17; Clinical Science Research and Development: CX002351; Horizon Therapeutics; National Institute of General Medical Sciences: U54GM115458.

[1]Katie J. McMenamin, MD: Boston Medical Center, Boston, Massachusetts; [2]Thomas R. Riley IV, MD, Sylvanus Toikumo, PhD, Henry R. Kranzler, MD, Rachel L. Kember, MSc, PhD, Joshua F. Baker, MD, MSCE: University of Pennsylvania and Corporal Michael J. Crescenz Department of Veterans Affairs Medical Center, Philadelphia; [3]Kristin Wipfler, PhD: FORWARD, The National Databank for Rheumatic Diseases, Wichita, Kansas; [4]Kaleb Michaud, PhD: FORWARD, The National Databank for Rheumatic Diseases, Wichita, Kansas, and VA Nebraska-Western Iowa Health Care System and University of Nebraska Medical Center, Omaha; [5]Austin Wheeler, MD, Bryant R. England, MD, PhD, Ted R. Mikuls, MD, MSPH: VA Nebraska-Western Iowa Health Care System and University of Nebraska Medical Center, Omaha; [6]Brian Sauer, PhD, Grant W. Cannon, MD: Salt Lake City VA Medical Center and University of Utah, Salt Lake City; [7]Katherine D. Wysham, MD: VA Puget Sound Health Care System and University of Washington, Seattle; [8]Rui Xiao, PhD: University of Pennsylvania and Children's Hospital of Philadelphia, Philadelphia; [9]Michael March, PhD: Children's Hospital of Philadelphia, Philadelphia, Pennsylvania.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25588).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25588.

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

Submitted for publication September 5, 2024; accepted in revised form June 13, 2025.

SIGNIFICANCE & INNOVATIONS

- This study characterized associations between previously identified fibromyalgia-related genetic polymorphisms and pain in patients with rheumatoid arthritis (RA).
- Two genetic risk scores based on genetic polymorphisms correlated with pain and disease activity in an external population.
- These findings support a proof-of-concept that genetic predisposition to pain can affect outcomes in RA.

medications, including biologic therapies, than patients with RA without FM.[5,6] Recognition of the multifactorial sources of pain in RA, including comorbid FM, is important to management to prevent unnecessary escalation of antirheumatic therapies or, conversely, the discontinuation of effective drugs due to blunted patient response.

Many factors have been associated with the risk of developing FM, which, in patients with RA, include socioeconomic status, psychological distress, RA severity, and comorbid conditions.[7] Additionally, studies have suggested a genetic predisposition to FM.[8] Genetic differences in peripheral and central pain processing are thought to play key roles but remain poorly understood.[9] Several single-nucleotide polymorphisms (SNPs) have been reported to be associated with FM susceptibility, whereas others have been shown to alter pain thresholds and processing.[8] A recent meta-analysis summarized data on 30 genes related to the pathophysiology and clinical presentation of FM, including *MTHFR*, *SCN9A*, *CNR1*, and *COMT*.[8] However, the prevalence of these SNPs and their association with pain and disease severity among people with RA has not been studied to date. The ability to identify those patients with RA at highest risk of developing FM and/or experiencing higher levels of pain discordant with disease burden would enable clinicians to better predict clinical course, counsel patients, and identify treatment goals. Furthermore, understanding genetic susceptibility to pain in people with RA may elucidate the mechanisms of overlap between the two conditions.

To evaluate the relationship between FM-related SNPs and pain in people with RA, we conducted a novel analysis of two distinct existing longitudinal cohorts of patients with RA from the National Databank for Rheumatic Diseases (FORWARD) and the Veterans Affairs Rheumatoid Arthritis Registry (VARA). We hypothesized that patients with more FM-related SNPs would have higher levels of pain and RA disease activity scores. Specifically, we aimed to determine (1) whether any of 36 identified FM-related SNPs[8] are associated with pain and global disease activity scores in patients with RA and (2) whether a composite score of genetic risk from these SNPs can predict pain and global disease activity scores in an external cohort. We also compared our genetic risk score (GRS) to a recently published GRS from a genome-wide association study (GWAS) of pain intensity.[10]

MATERIALS AND METHODS

Study setting. FORWARD is a patient-facing rheumatic disease registry established in 1998. Patients are enrolled from community-based rheumatology practices across the United States and complete regularly scheduled questionnaires. RA is diagnosed by treating rheumatologists, and key patient data are validated with medical records. In 2010, the Arthritis Internet Registry and biorepository was initiated to provide opportunities to evaluate laboratory measures and correlate these values with detailed questionnaire data. A subset of existing FORWARD participants and newly recruited participants provided biologic samples that underwent genotyping, among other tests.[11]

VARA is a national biologic repository and longitudinal clinical database following veterans from Department of Veterans Affairs (VA) rheumatology clinics in 17 US cities. Clinical data, including demographics, disease features, and disease activity, are entered by clinicians during routine visits and are linked with biobanked serum, plasma, and DNA samples obtained at enrollment for each participant. Biobanked specimens are used for genotyping and to conduct standard laboratory assays, such as for markers of inflammation and anti–cyclic citrullinated protein antibodies.[12] For both cohorts, participants were included if they had visual analog pain scores from their index visit and genotype data. The study received local approval from the institutional review boards at the participating sites.

SNPs. In both FORWARD and VARA, genotyping was performed with the Illumina Infinium Global Screening Array platform, a high-density throughput array that includes 665,608 SNPs, by the Children's Hospital of Philadelphia Genomics Analytics Core.[13,14] Genotypes underwent quality control processing, including filtering out genotypes and samples with missingness >10% or minor allele frequency <1%.[14] Imputation was performed using the Michigan Imputation Server and the Haplotype Reference Consortium reference panel, filtered by $R^2 > 0.7$ to retain the highest quality imputed SNPs.[15] Genotypes underwent quality control processing, including filtering out genotypes with missingness >10%, samples with missingness >5%, or minor allele frequency <1%. From these genetic data, we examined 42 specific SNPs, which we prehypothesized to be associated with pain based on their potential associations with FM identified in a review by Janssen et al.[8] Of these 42 SNPs, five pairs were found to be in linkage disequilibrium with an R^2 of greater than 0.5. For each pair in linkage disequilibrium, the SNP with the least significant association with the baseline pain scale was dropped. One additional SNP was dropped because it was x-linked (rs6323), and sample sizes were inadequate to assess differential effects within each sex, leaving 36 SNPs in the analysis

(Supplemental Table 1). Most SNPs were analyzed as categorical variables with three levels: heterozygous, homozygous reference alleles, or homozygous alternate alleles. SNPs with minor allele frequencies <0.20 were analyzed as binary variables, with homozygous minor alleles and heterozygosity in the same category (presence of minor allele). All associations tested are presented within the article. We used the independent associations defined for these SNPs within linear regression models to develop a novel GRS for pain (FM GRS).

As a secondary analysis, we also evaluated the application of an existing GRS recently developed using a GWAS for pain intensity performed using all participants from the Million Veteran Program (MVP).[10] The GRS was derived using the effect estimates from the MVP pain intensity GWAS, and SNPs were selected using the PRS-Continuous shrinkage software in all participants from the Penn Medicine Biobank.[10,16] We calculated this pain intensity GRS with PLINK v1.9 using the imputed genotypes of participants in FORWARD and VARA. The pain intensity GRS included 39% of SNPs in the FM GRS, but these overlapping SNPs only accounted for 0.0025% of the pain intensity GRS (14 of 556,987). SNPs included in both the FM GRS and an existing pain intensity GRS[17,18] are shown in Supplemental Table 1.

Outcomes: Pain. In FORWARD, participants reported visual analog scale (VAS) pain scores on a range from 0 (no pain) to 10 (severe pain) at six-month intervals either by mail or online questionnaire according to participant preference. VAS pain was reported as discrete values with a precision of 0.5 by selecting one of 21 boxes on a visual scale. In VARA, patients reported VAS pain scores at their routine clinical visits by circling a whole number between 0 and 10 from a horizontal list. In separate analyses, we assessed VAS pain scores at baseline (upon study enrollment) and throughout the course of each study (using all follow-up observations) for both populations. In FORWARD, the presence of FM was defined as the presence of widespread pain for at least three months and either a widespread pain index (WPI) >6 and a symptom severity score (SSS) >4 or a WPI >3 with an SSS >8.

Outcomes: Disease activity scales. In FORWARD, RA disease activity was assessed with the Patient Activity Scale II (PAS-II). This is a validated patient-reported assessment of RA activity that combines the VAS pain score, patient assessment of global disease activity, and measures of physical disability from the Health Assessment Questionnaire Disability Index II onto a scale of 0 to 10 from lowest to highest disease activity.[19] In VARA, we used a similar patient-reported assessment of disease activity, the Routine Assessment of Patient Index Data 3 (RAPID3). This is a composite index of patient-reported function (Multidimensional Health Assessment Questionnaire), pain, and patient global assessment, each rated from 0 to 10, resulting in a score[20] from 0 to 30. We also explored associations with the Disease Activity Score in 28 joints (DAS28) among those with available data in VARA.[21] Each disease activity score was assessed at baseline (upon study enrollment) and throughout the course of follow-up (all observations) within each study.

Statistical analysis. To account for potential population stratification due to genetic ancestry, a principal component analysis on the preimputed genotypes was performed to identify major clusters of genetic similarity using PLINK version 1.9, as previously described.[17,18,22] All models were then adjusted for the top five principal components of genetic ancestry, which explained 90% of the variance in population structure in the FORWARD population.[23]

We used multivariable linear regression to assess the independent relationship between SNP variants and baseline VAS pain scores, adjusting for age, sex, and genetic ancestry in both FORWARD and VARA. We used 2df testing for genetic association, using the three aforementioned categories. Two categories were used for SNPs with low-frequency alternative alleles, as described previously. The reference category was the presence of homozygous reference alleles. The coefficients for all 36 SNPs from this regression in FORWARD were then used as weights to generate a GRS computing the sum of the risk alleles for each individual weighted by the effect size estimate (ie, the regression coefficient) from the test of association for each of the 36 SNPs within the multivariable model. We calculated this score for participants in both FORWARD (training set) and VARA (validation set).

We then used multivariable linear regression to assess the relationship between the novel FM GRS and an existing pain intensity GRS and baseline VAS pain scores in both datasets, adjusting for age, sex, and genetic ancestry. Similarly, we also used linear regression to assess the relationship between each GRS and baseline disease activity in FORWARD (PAS-II) and VARA (RAPID3), adjusting for the same covariates. We also assessed associations with repeated measures over all follow-up visits using linear regression incorporating generalized estimating equations, adjusting for age, sex, and genetic ancestry.

Differences in individual components of disease activity between GRS quartiles were assessed using Kruskal-Wallis tests. In addition to individual components of disease activity scores, we assessed differences in more objective measures of disease activity/inflammation using C-reactive protein (CRP) in both datasets and tender and swollen joint counts in VARA.

RESULTS

Study population. The study included 756 participants from FORWARD and 2,176 participants from VARA who had both genetic and VAS pain score data available at the time of the analysis. Participant characteristics for each cohort are shown in Table 1. FORWARD participants were mostly female (89.4%) and White (91.5%) and had a mean age of 56.8 years, whereas

Table 1. Baseline characteristics of FORWARD and VARA populations*

	FORWARD (N = 756)	VARA (N = 2,176)
Age, mean (SD), y	56.8 (11.3)	64.3 (11.0)
Female, n (%)	683 (89.4)	240 (11)
Race, n (%)		
American Indian or Alaskan Native	3 (0.4)	32 (1.4)[a]
Asian or Pacific Islander	10 (1.3)	
Black (not Hispanic)	17 (2.3)	371 (17.1)
Hispanic	24 (3.2)	104 (4.8)
White (not Hispanic)	696 (91.5)	1,653 (76.0)
Other or unknown	11 (1.45)	16 (0.5)
Disease duration, median (IQR), y	10 (3.70–19.0)	7.5 (2.2–16.6)
BMI, mean (SD)	28.3 (6.9)	28.8 (5.7)
Met fibromyalgia 2016[24] criteria, n (%)	118 (20.3)	N/A
ACPA status, n (%)	426 (55.7)	1,578 (76.8)
Biologic DMARD, n (%)	389 (54.6)	700 (32.2)
High CRP (≥0.8 mg/dL), n (%)	223 (29.2)	1,044 (50.6)
Conventional DMARD, n (%)	584 (77.4)	1,707 (78.5)
Baseline VAS pain score, mean (SD)	3.8 (2.8)	4.4 (2.8)

* Categorical variables are presented as n (percentage of population), and continuous variables are presented as mean (SD) or median (IQR) for skewed variables. ACPA, anti–citrullinated protein antibody; BMI, body mass index; CRP, C-reactive protein; DMARD, disease-modifying antirheumatic drug; IQR, interquartile range; N/A, not applicable; VARA, Veterans Affairs Rheumatoid Arthritis Registry; VAS, visual analog scale.
[a] Pacific Islander and Alaskan Native are categorized together in VARA.

VARA participants were mostly male (89%) and older (mean age of 64.3 years) and had more racial and ethnic diversity (76.0% White, 17.1% Black, 4.8% Hispanic). VARA participants also had shorter disease duration (7.5 vs 10 years), more seropositivity (76.8% vs 55.7%), higher CRP levels (>0.8 mg in 50.6% vs 29.2%), and less biologic use (32.2% vs 54.6%) at enrollment. In FORWARD, 20.3% of included participants met criteria for FM, whereas VARA did not have these data available. Baseline VAS pain scores were slightly higher for VARA participants than for FORWARD participants (mean 4.4 vs 3.8). The allele frequency in each population ranged from 0.06 to 0.95 and is included in detail in Supplemental Table 1.

Association of individual SNPs with baseline pain. The results of the multivariable model showing independent associations between individual SNP variants and baseline VAS pain scores in FORWARD and VARA are shown in Table 2. In FORWARD, only one SNP, rs25531, was independently associated with baseline pain. The presence of an alternate allele was associated with less baseline pain compared to homozygous reference alleles (β –0.68 [95% confidence interval (CI) –0.1.28 to –0.08], $P = 0.03$). In VARA, this SNP was not significantly associated with baseline pain, but four others were. For example, the heterozygosity of rs6280 was associated with less baseline pain compared to homozygous reference alleles (β –0.43 [95% CI –0.78 to –0.07] $P = 0.02$).

GRS. In FORWARD, the FM GRS (generated from coefficients from Table 2) had a mean of –0.92 and an SD of 0.74, with lower scores representing lower risk for pain and higher scores representing higher risk. In VARA, the mean and distribution were similar (as shown in Supplemental Figure 1), with a mean GRS of –0.93 and an SD of 0.77. The FM GRS correlated poorly with the existing pain intensity GRS in both FORWARD (ρ = 0.12) and VARA (ρ = 0.02).

GRS and pain. The novel FM GRS was strongly associated with baseline VAS pain scores in FORWARD participants (β 0.15 [95% CI 0.11–0.18], $P < 0.01$). In VARA participants, the FM GRS tended to correlate with baseline pain, but the association was modest (β 0.15 [95% CI –0.01 to 0.31], $P = 0.06$). The regression model including the FM GRS explained 7.9% of the variance in pain scales in FORWARD (compared to 1.7% without the GRS) and 1.9% of the variance in pain scales in VARA (compared to 1.7% without the GRS). As illustrated in Figure 1, participants in FORWARD in the highest quartile of the FM GRS had baseline VAS pain scores 1.96 units higher (95% CI 1.39–2.54, $P < 0.001$) compared to those in the lowest GRS quartile. VARA participants in the highest FM GRS quartile had 0.55 (95% CI 0.16–0.93, $P = 0.006$) greater baseline VAS pain scores than those in the lowest quartile.

Similarly, the FM GRS was associated with VAS pain scores across all follow-up time points in FORWARD participants (β 0.33 [95% CI 0.23–0.43], $P < 0.01$) and in VARA participants (β 0.13 [95% CI 0.02–0.24], $P = 0.02$). Participants in the highest quartile in FORWARD had pain scores that were, on average, 1.28 higher than those in the lowest quartile over all observations (Table 3). In VARA, participants in the highest quartile had 0.36 (95% CI 0.08–0.64, $P = 0.01$) higher pain scores. In FORWARD, the FM GRS was weakly correlated with polysymptomatic distress scores at baseline (correlation coefficient of 0.12). FM and pain sensitization data were not available in VARA.

The existing pain intensity GRS was strongly associated with pain in both FORWARD and VARA at baseline and over time (Table 3). For example, in VARA there were significantly higher pain scores over time among participants in the highest quartile of the pain intensity GRS (β 1.07 [95% CI 0.64–1.51], $P < 0.001$).

GRS and disease activity. As shown in Table 4, the FM GRS was also associated with baseline disease activity as quantified by the PAS-II score in FORWARD participants. Participants in the highest FM GRS quartile had 1.39 (95% CI 0.92–1.85, $P <$ 0.01) higher PAS-II scores (on a 10-point scale) than those in the lowest quartile. In VARA, the testing data set, the FM GRS was also significantly associated with baseline disease activity but to

Table 2. Baseline pain scale and individual SNPs in multivariable regression with adjustment for age, sex, and genetic ancestry*

			FORWARD		VARA	
SNP	Ref. Allele	Gene	Heterozygous, β (95% CI)	Homozygous alt. alleles, β (95% CI)	Heterozygous, β (95% CI)	Homozygous alt. alleles, β (95% CI)
rs1801133	A	*MTHFR*	0.04 (−0.40 to 0.48)	−0.21 (−0.93 to 0.52)	0.08 (−0.19 to 0.34)	0.14 (−0.30 to 0.58)
rs2805050[a]	C	*RGS4*	0.30 (−0.37 to 0.97)	–	0.16 (−0.20 to 0.53)	–
rs2842003	G	*RGS4*	−0.20 (−0.79 to 0.39)	0.30 (−0.31 to 0.92)	−0.18 (−0.51 to 0.15)	−0.18 (−0.51 to 0.15)
rs10799897	G	*RGS4*	−0.01 (−0.50 to 0.49)	0.14 (−0.43 to 0.71)	0.09 (−0.20 to 0.37	0.08 (−0.27 to 0.42)
rs11127292	T	*MYT1L*	−0.27 (−0.80 to 0.27)	0.52 (−1.06 to 2.10)	−0.10 (−0.41 to 0.21)	0.17 (−0.96 to 1.31)
rs3771863[a]	T	*TACR1*	0.11 (−0.350 to 0.57)	–	0.08 (−0.19 to 0.35)	–
rs6280	T	*SCN9A*	−0.06 (−0.81 to 0.69)	0.21 (−0.55 to 0.97)	−0.43 (−0.78 to −0.07)[b]	−0.32 (−0.71 to 0.06)
rs1042713	A	*DRD3*	0.01 (−0.52 to 0.54)	−0.20 (−1.02 to 0.63)	−0.03 (−0.36 to 0.30)	0.05 (−0.41 to 0.51)
rs1042714	C	*ADRB2*	−0.00 (−0.68 to 0.68)	−0.18 (−0.98 to 0.61)	−0.21 (−0.63 to 0.22)	−0.21 (−0.63 to 0.22)
rs1800541	G	*ADRB2*	−0.29 (−0.73 to 0.16)	−0.57 (−1.79 to 0.66)	−0.12 (−0.39 to 0.15)	−0.44 (−1.05 to 0.16)
rs10485171	G	*EDN1*	0.28 (−0.19 to 0.75)	0.06 (−0.55 to 0.66)	0.13 (−0.15 to 0.42)	0.11 (−0.24 to 0.47)
rs6454674	G	*CNR1*	0.04 (−0.40 to 0.48)	−0.20 (−0.94 to 0.54)	0.32 (0.06 to 0.58)[b]	−0.02 (−0.45 to 0.40)
rs4129256	A	*CNR1*	−0.09 (−0.56 to 0.39)	−0.78 (−2.04 to 0.49)	−0.05 (−0.34 to 0.24)	−0.14 (−0.84 to 0.55)
rs8192619[a]	A	*TAAR1*	−0.35 (−1.11 to 0.41)	–	0.32 (−0.13 to 0.77)	–
rs1799971	G	*TAAR1*	−0.25 (−0.73 to 0.23)	0.57 (−0.86 to 1.99)	−0.12 (−0.43 to 0.19)	0.05 (−0.92 to 1.02)
rs2097903	T	*OPRM1*	−0.45 (−0.96 to 0.06)	0.11 (−0.47 to 0.69)	−0.03 (−0.33 to 0.27)	0.031 (−0.31 to 0.38)
rs1048101	G	*ANDRA1A*	−0.03 (−0.50 to 0.44)	−0.33 (−0.92 to 0.28)	0.10 (−0.20 to 0.40)	0.03 (−0.33 to 0.40)
rs1383914	C	*ANDRA1A*	−0.13 (−0.62 to 0.35)	−0.11 (−0.72 to 0.49)	0.14 (−0.15 to 0.44)	−0.12 (−0.48 to 0.24)
rs574584[a]	T	*ANDRA1A*	−0.24 (−0.93 to 0.45)	–	0.34 (−0.04 to 0.72)	–
rs4994	G	*ADRB3*	0.22 (−0.35 to 0.78)	0.80 (−2.05 to 3.66)	−0.01 (−0.36 to 0.35)	1.39 (0.08 to 2.70)[b]
rs12273539[a]	T	*BDNF*	0.34 (−1.14 to 1.82)	–	0.55 (0.07 to 1.03)[b]	–
rs11030104[a]	G	*BDNF*	−0.18 (−0.62 to 0.26)	–	0.09 (−0.17 to 0.37)	–
rs2510177[a]	G	*GRIA4*	0.12 (−0.85 to 1.09)	–	0.19 (−0.29 to 0.66)	–
rs10895837[a]	C	*GRIA4*	−0.24 (−1.08 to 0.60)	–	−0.32 (−0.75 to 0.10)	–
rs642544	G	*GRIA4*	−0.00 (−0.47 to 0.46)	−0.26 (−0.90 to 0.38)	−0.01 (−0.29 to 0.26)	0.18 (−0.20 to 0.56)
rs17104711[a]	A	*GRIA4*	−0.39 (−0.95 to 0.16)	–	−0.10 (−0.47 to 0.27)	–
rs118162387[a]	A	*HTR3A*	0.74 (−0.14 to 1.63)	–	−0.14 (−0.70 to 0.42)	–
rs6313	A	*HTR2A*	−0.06 (−0.51 to 0.40)	0.39 (−0.24 to 1.01)	0.12 (−0.15 to 0.39)	0.21 (−0.16 to 0.57)
rs841	A	*GCH1*	0.10 (−0.35 to 0.56)	−0.26 (−1.17 to 0.65)	−0.07 (−0.33 to 0.19)	−0.19 (−0.78 to 0.40)
rs224222	T	*TRPV3*	0.12 (−0.32 to 0.57)	−0.25 (−1.09 to 0.58)	0.11 (−0.15 to 0.38)	0.18 (−0.37 to 0.74)
rs395357	T	*MEFV*	0.18 (−0.31 to 0.67)	0.22 (−0.37 to 0.81)	−0.05 (−0.34 to 0.23)	0.23 (−0.12 to 0.58)
rs25531[a]	C	*SLC6A4*	−0.68 (−1.28 to −0.08)[b]	–	−0.03 (−0.37 to 0.30)	–
rs4633	T	*COMT*	−0.03 (−0.77 to 0.71)	−0.22 (−1.18 to 0.74)	0.00 (−0.38 to 0.38)	−0.08 (−0.61 to 0.45)
rs4818	G	*COMT*	−0.50 (−1.16 to 0.16)	−0.30 (−1.33 to 0.74)	−0.10 (−0.45 to 0.26)	−0.09 (−0.67 to 0.48)
rs6971	G	*COMT*	−0.29 (−0.99 to 0.42)	−0.08 (−0.78 to 0.61)	−0.28 (−0.73 to 0.17)	−0.27 (−0.71 to 0.17)
rs6754031	G	*TSPO*	−0.09 (−0.55 to 0.37)	0.42 (−0.19 to 1.04)	−0.01 (−0.27 to 0.26)	0.17 (−0.21 to 0.55)

* All SNPs are included in the model in addition to variables for age, sex, and ancestry principal components. alt., alternate; CI, confidence interval; Ref., reference; SNP, single-nucleotide polymorphism; VARA, Veterans Affairs Rheumatoid Arthritis Registry.
[a] Low minor allele frequency: coefficient compares homozygous major alleles to any other combination of alleles.
[b] $P \leq 0.05$.

a lesser extent than in FORWARD. VARA participants in the highest quartile of the FM GRS had higher RAPID3 scores (β 1.12 [95% CI 0.22–2.02], $P = 0.02$) and higher DAS28 scores (β 0.22 [95% CI 0.01–0.43], $P = 0.04$) than those in other quartiles.

Similarly, the FM GRS was associated with disease activity over all observations in FORWARD and VARA as estimated by PAS-II and RAPID3, respectively (Table 4). Over the course of the study, FORWARD participants in the highest quartile of the FM GRS had a PAS-II score that was 1.02 (95% CI 0.61–1.43, $P < 0.01$) greater (on a 10-point scale) than those in the lowest quartile. VARA participants in the highest quartile had higher RAPID3 scores (β 0.67 [95% CI 0.02–1.33], $P = 0.04$), on a 30-point scale, compared to those in other quartiles. A higher FM GRS was also associated with higher DAS28, with those in the highest quartile having 0.19 (95% CI 0.05–0.33], $P = 0.01$) higher scores (on a 9.4-point scale) compared to those in other quartiles.

The existing pain intensity GRS was strongly associated with higher time-averaged PAS-II, RAPID3, and DAS28 scores, with a similar or greater magnitude of effect as the FM GRS (Table 5). For example, in VARA, participants in the highest quartile of the pain intensity GRS had significantly higher DAS28 scores over all follow-up visits (β 0.38 [95% CI 0.17–0.60], $P < 0.01$) compared to those in the lowest quartile.

GRS and components of disease activity and FM. Although baseline VAS pain and VAS patient global scores were numerically higher among participants in higher quartiles of the FM GRS in both FORWARD and VARA, other measures of baseline disease activity, including CRP levels, swollen joint counts,

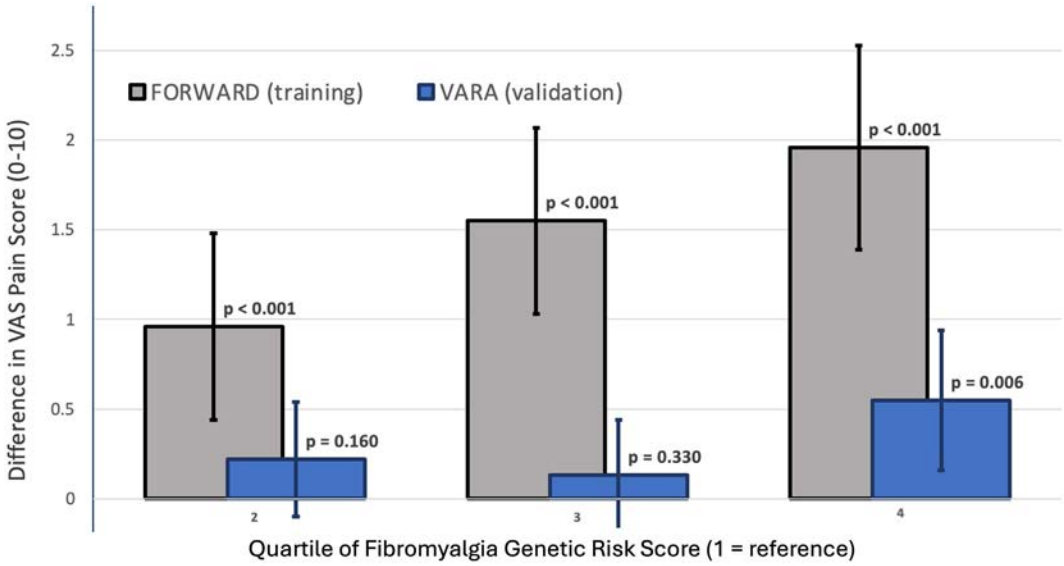

Figure 1. Difference in actual baseline VAS pain scores (mean, 95% confidence interval) by quartile of fibromyalgia genetic risk score (vs quartile 1, lowest) for FORWARD (training; gray) and VARA (validation; blue) data sets at baseline. VARA, Veterans Affairs Rheumatoid Arthritis Registry; VAS, visual analog scale.

and tender joint counts, were similar between quartiles, as shown in Supplemental Table 2.

DISCUSSION

We found that a composite GRS for pain based on 36 FM-related SNPs, generated using data from the FORWARD registry, was associated with baseline pain in an external and demographically distinct cohort (VARA). The external validation supports a modest genetic influence on self-reported pain scores in patients with RA and provides proof-of-concept that such an approach to predicting pain based on genetic variation may be possible. We also found that an existing GRS derived from a GWAS for pain intensity[10] was similarly associated with higher pain scores. The highest quartile of this score was associated with 1-point higher pain scores on a 0 to 10 scale, a difference that may be clinically meaningful. The FM GRS was also significantly associated with RA disease activity scores. This was driven primarily by an association with pain and patient global scores, whereas other components of disease activity (CRP levels, swollen and tender joint counts) were not observed to vary by genetic risk. These data suggest that differences in disease activity between genetic risk quartiles are likely driven by differences in the subjective experience of pain rather than differences in RA-related inflammation.

None of the individual SNPs studied here were significantly associated with pain in both populations studied. The findings from the composite risk score suggest that some weak associations identified here are important but may have too small an effect on their own to reach statistical significance in this sample. Some differences in the estimates across cohorts could also be

Table 3. Longitudinal visual analog pain scales and GRS quartile in multivariable regression with adjustment for age, sex, and genetic ancestry in FORWARD (training) and VARA (validation)*

	Training (FORWARD)[a]		Validation (VARA)[b]	
GRS quartile	β (95% CI)	*P*	β (95% CI)	*P*
Fibromyalgia GRS				
1	Ref	–	Ref	–
2	0.69 (0.29 to 1.10)	0.001	0.09 (−0.14 to 0.32)	0.45
3	1.00 (0.59 to 1.41)	<0.001	0.27 (0.03 to 0.51)	0.03
4	1.28 (0.82 to 1.75)	<0.001	0.36 (0.08 to 0.64)	0.01
Pain intensity GRS				
1	Ref	–	Ref	–
2	0.47 (0.08 to 0.87)	0.02	0.43 (0.19 to 0.68)	<0.001
3	1.12 (0.71 to 1.53)	<0.001	0.76 (0.52 to 1.00)	<0.001
4	1.38 (0.93 to 1.83)	<0.001	1.07 (0.64 to 1.51)	<0.001

* CI, confidence interval; GRS, genetic risk score; Obs, number of observations; Ref, reference; VARA, Veterans Affairs Rheumatoid Arthritis Registry.
[a] N = 759; Obs = 9,814.
[b] N = 2,176; Obs = 26,701.

Table 4. Baseline and longitudinal disease activity scores and FM GRS quartile in FORWARD (training) and VARA (validation)*

	Training (FORWARD)		Validation (VARA)			
	PAS-II		RAPID3		DAS28	
FM GRS quartile	β (95% CI)	*P*	β (95% CI)	*P*	β (95% CI)	*P*
Baseline[a]						
1	Ref	–	Ref	–	Ref	–
2	0.74 (0.32 to 1.16)	0.001	0.55 (−0.18 to 1.29)	0.14	0.11 (−0.06 to 0.28)	0.20
3	0.96 (0.54 to 1.39)	<0.001	0.65 (−0.09 to 1.39)	0.08	0.18 (0.01 to 0.35)	0.04
4	1.39 (0.92 to 1.85)	<0.001	1.12 (0.22 to 2.02)	0.02	0.22 (0.01 to 0.43)	0.04
Longitudinal (all observations)[b]						
1	Ref	–	Ref	–	Ref	–
2	0.55 (0.20 to 0.90)	0.002	0.35 (−0.20 to 0.91)	0.21	0.11 (−0.01 to 0.22)	0.06
3	0.78 (0.42 to 1.14)	<0.001	0.75 (0.19 to 1.31)	0.008	0.19 (0.08 to 0.30)	0.001
4	1.02 (0.61 to 1.43)	<0.001	0.67 (0.02 to 1.33)	0.04	0.19 (0.05 to 0.33)	0.01

* CI, confidence interval; DAS28, Disease Activity Score in 28 joints; FM, fibromyalgia; GRS, genetic risk score; Obs, number of observations; PAS-II, Patient Activity Scale II; RAPID3, Routine Assessment of Patient Index Data 3; Ref, reference; VARA, Veterans Affairs Rheumatoid Arthritis Registry.
[a] PAS-II: N = 710; RAPID3: N = 1,780; DAS28: N = 1,918.
[b] PAS-II: N = 758, Obs = 9,363; RAPID3: N = 2,097, Obs = 21,792; DAS28: N = 2,152, Obs = 26,639.

attributable to differences in population characteristics (eg, sex). Overall, these results highlight the importance of validation of observations across multiple datasets because some seemingly meaningful associations between genetics and phenotype can be observed due to random chance.

Previous investigations have quantified genome-wide associations and have attempted to generate polygenic risk scores for chronic pain syndromes in other populations, but we are not aware of any other studies done in RA.[25] A study by McIntosh et al demonstrated that a polygenic risk score generated from 23andMe data was strongly associated with chronic pain phenotypes in the UK Biobank.[26] Another study from the UK Biobank found that a GRS was modestly predictive of those with chronic back pain.[27] Importantly, here we applied an existing pain intensity score developed in the general population, which performed well and demonstrated a significant and potentially meaningful association (moderate effect sizes) with clinically important RA outcomes. The results from the current study support a genetic basis for chronic pain and emphasize the potential impact that this risk may have on important clinical outcomes, such as disease activity, for patients with RA.

The novel FM GRS in this study modestly predicted self-reported pain in the validation dataset, which suggests that genetic factors related to FM likely have some impact on pain-related outcomes in people with RA. Interestingly, both the novel FM GRS and the existing pain intensity score were predictive of

Table 5. Baseline and longitudinal disease activity scores and existing pain intensity GRS quartile in FORWARD and VARA*

	FORWARD		VARA			
	PAS-II		RAPID3		DAS28	
Pain intensity GRS quartile	β (95% CI)	*P*	β (95% CI)	*P*	β (95% CI)	*P*
Baseline[a]						
1	Ref	–	Ref	–	Ref	–
2	0.42 (−0.01 to 0.87)	0.06	0.71 (−0.08 to 1.49)	0.08	0.05 (−0.13 to 0.23)	0.59
3	1.21 (0.76 to 1.65)	<0.001	1.21 (0.42 to 2.01)	0.003	0.21 (0.03 to 0.39)	0.02
4	1.12 (0.66 to 1.57)	<0.001	1.63 (0.25 to 3.01)	0.02	0.38 (0.06 to 0.71)	0.02
Longitudinal (all observations)[b]						
1	Ref	–	Ref	–	Ref	–
2	0.41 (0.05 to 0.77)	0.02	0.89 (0.32 to 1.45)	0.002	0.17 (0.06 to 0.29)	0.004
3	0.96 (0.60 to 1.33)	<0.001	1.78 (1.21 to 2.35)	<0.001	0.27 (0.15 to 0.40)	<0.001
4	1.21 (0.82 to 1.60)	<0.001	2.66 (1.66 to 3.66)	<0.001	0.38 (0.17 to 0.60)	0.001

* CI, confidence interval; DAS28, Disease Activity Score in 28 joints; GRS, genetic risk score; Obs, number of observations; PAS-II, Patient Activity Scale II; RAPID3, Routine Assessment of Patient Index Data 3; Ref, reference; VARA, Veterans Affairs Rheumatoid Arthritis Registry.
[a] PAS-II: N = 709; RAPID3: N = 1,780; DAS28: N = 1,986.
[b] PAS-II: N = 757, Obs = 9,346; RAPID3: N = 2,097, Obs = 21,792; DAS28: N = 2,152, Obs = 26,639.

pain in RA, but they were poorly correlated. This suggests that many pain-related SNPs work through pathways that are independent from one another. Although the effect size was smaller, the FM GRS, which includes SNPs that are mechanistically involved in FM pain pathways, may be more useful to clinicians aiming to predict RA pain out of proportion to disease activity. However, larger studies may determine which genes are the most important in driving the differences in pain and identify the specific mechanisms involved. Further investigation into these complex genetic pathways could lead to the development of comprehensive risk scores that are closer to the GWAS-based score in predictive value but are as straightforward to interpret as the pathway-specific score.

Other future directions include investigating the association of these SNPs with other outcomes in RA, such as failure to achieve remission and discordance between patient and provider disease activity scores. It would also be of interest to compare the performance of a GRS for pain in patients with RA with and without comorbid FM to determine whether the condition is an effect modifier.

Strengths of this study include the use of two independent and demographically distinct datasets to train and test a novel GRS from FM-related SNPs as well as the incorporation of an existing pain intensity GRS from GWAS data. We created a GRS based on the association between FM genes and VAS pain scores in FORWARD participants and demonstrated that these same genes confer risk in a very distinct and independent patient population (VARA). The observation of more modest effects in the validation cohort is expected and illustrates the common issue of overly optimistic prediction models developed and validated within a single population. Additionally, the evaluation of associations with both individual SNPs and a composite score illustrates that although the individual gene effects are small, they can still be informative when combined. Another strength of this study is the analysis using a previously developed pain intensity GRS from a GWAS of pain. We found similar associations with pain and disease activity scores in these patient populations and validated our approach, demonstrating the potential for stronger risk prediction with more comprehensive scores.

Limitations of this study include the relatively small development data set, as well as limited ancestral and sexual diversity of the studied populations, which limited our ability to perform subgroup analyses within these populations. Additionally, the lack of FM phenotypic data in VARA made it difficult to quantify the degree of effect that can be explained by genotype on that clinical phenotype. We used SNPs identified in a single review article, with the goal of focusing on genetic factors with a known or proposed pathophysiologic connection to amplified pain, and demonstrated associations with FM. Inclusion of additional genetic loci using a more comprehensive, but pathway-agnostic, approach improved performance. Although the current study is a proof-of-concept approach to modeling the genetic risk of pain in patients with RA, more advanced techniques may improve the prediction of similar models and are a logical future goal. In addition to identification of additional relevant genes, methods to integrate other types of data, such as environmental factors, proteomics, and metabolomics, may provide an opportunity to develop clinical tools of sufficient accuracy to guide clinical care.

In conclusion, we report a GRS that can predict self-reported pain in patients with RA. Those with higher genetic risk reported greater pain and greater disease activity components in both the development and the external validation cohorts, though the association was modest. This study provides initial evidence that genetic information could eventually be used to identify patients at higher risk of chronic pain and FM and help guide the approach to management in patients with RA.

AUTHOR CONTRIBUTIONS

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

REFERENCES

1. Hunter TM, Boytsov NN, Zhang X, et al. Prevalence of rheumatoid arthritis in the United States adult population in healthcare claims databases, 2004-2014. Rheumatol Int 2017;37(9):1551–1557.
2. Walsh DA, McWilliams DF. Mechanisms, impact and management of pain in rheumatoid arthritis. Nat Rev Rheumatol 2014;10(10):581–592.
3. Walsh DA, McWilliams DF. Pain in rheumatoid arthritis. Curr Pain Headache Rep 2012;16(6):509–517.
4. Zhao SS, Duffield SJ, Goodson NJ. The prevalence and impact of comorbid fibromyalgia in inflammatory arthritis. Best Pract Res Clin Rheumatol 2019;33(3):101423.
5. Gist AC, Guymer EK, Eades LE, et al. Fibromyalgia remains a significant burden in rheumatoid arthritis patients in Australia. Int J Rheum Dis 2018;21(3):639–646.
6. Lage-Hansen PR, Chrysidis S, Lage-Hansen M, et al. Concomitant fibromyalgia in rheumatoid arthritis is associated with the more frequent use of biological therapy: a cross-sectional study. Scand J Rheumatol 2016;45(1):45–48.
7. Wolfe F, Häuser W, Hassett AL, et al. The development of fibromyalgia--I: examination of rates and predictors in patients with rheumatoid arthritis (RA). Pain 2011;152(2):291–299.
8. Janssen LP, Medeiros LF, Souza ADE, et al. Fibromyalgia: a review of related polymorphisms and clinical relevance. An Acad Bras Cienc 2021;93(suppl 4):e20210618.
9. Meeus M, Vervisch S, De Clerck LS, et al. Central sensitization in patients with rheumatoid arthritis: a systematic literature review. Semin Arthritis Rheum 2012;41(4):556–567.
10. Toikumo S, Vickers-Smith R, Jinwala Z, et al; Million Veteran Program. A multi-ancestry genetic study of pain intensity in 598,339 veterans. Nat Med 2024;30(4):1075–1084.

11. Wolfe F, Michaud K. The National Data Bank for rheumatic diseases: a multi-registry rheumatic disease data bank. Rheumatology (Oxford) 2011;50(1):16–24.
12. Mikuls TR, Reimold A, Kerr GS, et al. Insights and implications of the VA Rheumatoid Arthritis Registry. Fed Pract 2015;32(5):24–29.
13. Infinium Global Screening Array v2.0 Locus Report. Illumina; 2018.
14. Li YR, Li J, Zhao SD, et al. Meta-analysis of shared genetic architecture across ten pediatric autoimmune diseases. Nat Med 2015; 21(9):1018–1027.
15. Das S, Forer L, Schönherr S, et al. Next-generation genotype imputation service and methods. Nat Genet 2016;48(10):1284–1287.
16. Ge T, Chen CY, Ni Y, et al. Polygenic prediction via Bayesian regression and continuous shrinkage priors. Nat Commun 2019;10(1):1776.
17. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81(3):559–575.
18. Chang CC, Chow CC, Tellier LC, et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 2015;4(1):7.
19. Wolfe F, Michaud K, Pincus T. A composite disease activity scale for clinical practice, observational studies, and clinical trials: the patient activity scale (PAS/PAS-II). J Rheumatol 2005;32(12):2410–2415.
20. Pincus T, Yazici Y, Bergman MJ. RAPID3, an index to assess and monitor patients with rheumatoid arthritis, without formal joint counts: similar results to DAS28 and CDAI in clinical trials and clinical care. Rheum Dis Clin North Am 2009;35(4):773–778, viii.
21. van Riel PLCM. The development of the disease activity score (DAS) and the disease activity score using 28 joint counts (DAS28). Clin Exp Rheumatol 2014;32(5 suppl 85):S-65–S-74.
22. Wheeler AM, Baker JF, Riley T, et al. Development and internal validation of a clinical and genetic risk score for rheumatoid arthritis-associated interstitial lung disease. Rheumatology (Oxford) 2025; 64(1):268–275.
23. Zhang HG, McDermott G, Seyok T, et al. Identifying shared genetic architecture between rheumatoid arthritis and other conditions: a phenome-wide association study with genetic risk scores. EBioMedicine 2023;92:104581.
24. Wolfe F, Clauw DJ, Fitzcharles MA, et al. 2016 Revisions to the 2010/2011 fibromyalgia diagnostic criteria. Semin Arthritis Rheum. 2016;46(3):319-329. doi:10.1016/j.semarthrit.2016.08.012
25. Johnston KJA, Adams MJ, Nicholl BI, et al. Genome-wide association study of multisite chronic pain in UK Biobank. PLoS Genet 2019; 15(6):e1008164.
26. McIntosh AM, Hall LS, Zeng Y, et al. Genetic and environmental risk for chronic pain and the contribution of risk variants for major depressive disorder: a family-based mixed-model analysis. PLoS Med 2016; 13(8):e1002090.
27. Tsepilov YA, Elgaeva EE, Nostaeva AV, et al. Development and replication of a genome-wide polygenic risk score for chronic back pain. J Pers Med 2023;13(6):977.

Arthritis Care & Research
Vol. 77, No. 12, December 2025, pp 1418–1425
DOI 10.1002/acr.25584

American College of Rheumatology
Empowering Rheumatology Professionals

Time Trends and Predictors of Gout Remission Over 6 Years

Adwoa Dansoa Tabi-Amponsah,[1] Sarah Stewart,[2] Greg Gamble,[1] Lisa K. Stamp,[3] William J. Taylor,[4] and Nicola Dalbeth[1]

Objective. This study aims to describe the trends in remission rates over 6 years of follow-up among people with gout taking urate-lowering therapy (ULT) and to identify variables that predict remission.

Methods. A post hoc analysis was conducted using data from the Cardiovascular Safety of Febuxostat or Allopurinol in Patients with Gout (CARES) trial, which enrolled people with gout and cardiovascular disease randomized to febuxostat or allopurinol. Gout remission over 6 years of follow-up was measured in participants with at least 1 year of follow-up data using the simplified gout remission definition, requiring the fulfillment of three domains: (1) no gout flares during the past year, (2) at least two serum urate measurements <0.36 mmol/L during the past year, and (3) no tophus. Logistic regression was used to identify baseline predictors of remission.

Results. Achievement of remission increased from 37.4% of participants (1,593/4,259) at year 1 to 63.1% (322/510) at year 6. Over the 6 years, 59.4% of participants achieved remission at least once. More participants receiving febuxostat achieved remission during the first 2 years, primarily because of a higher number achieving the serum urate remission domain. In multivariable analysis, baseline age, race, greater disease severity, presence of comorbidities, and febuxostat treatment were variables significantly associated with remission.

Conclusion. On ULT, fulfillment of remission increases over time and remission can be achieved in most patients. Baseline predictors, including demographics, comorbidities, and disease severity, may be useful to identify people with gout who need more proactive management to achieve remission.

INTRODUCTION

Gout is a chronic rheumatic disease of monosodium urate crystal deposition in the joints, tendons, and other tissues, resulting from sustained elevation of serum urate levels. Gout is characterized by recurrent episodes of acute inflammatory arthritis, known as gout flares.[1] If untreated, gout may progress to the development of subcutaneous tophi and chronic synovitis with joint damage.[1]

For chronic rheumatic diseases, remission is defined as "either a complete absence of disease activity or a level of disease activity so low that it is not troublesome to the patient and portends a later good prognosis."[2] In current gout management and research, the primary strategy is "treat to serum urate-target of <0.36 mmol/L" using urate-lowering therapy (ULT).[3] In 2016, a composite measure of disease activity, "remission," was introduced by rheumatologists and gout researchers with the development of preliminary gout remission criteria, which includes the following: no gout flares and serum urate level of <0.36 mmol/L (6 mg/dL), with no values ≥0.36 mmol/L (6 mg/dL) and absence of tophi; pain caused by gout <2 on a 10-point scale, with no values ≥2; and patient global assessment of gout disease activity <2 on a 10-point scale, with no values ≥2.[4] The 2016 preliminary criteria required that the serum urate, pain, and patient global assessment domains must be measured at least twice over 12 months.[4] Subsequently, a simplified gout remission criteria without patient-reported outcomes has been developed and compared with the 2016 preliminary criteria and has shown good face validity, concurrent validity, responsiveness, and discrimination.[5–7]

To date, analyses of gout remission have been conducted on studies that were 1 to 2 years in length.[5–7] However, the longitudinal trends in gout remission status over more extended periods are unknown. Furthermore, baseline variables that influence

[1]Adwoa Dansoa Tabi-Amponsah, BBiomedSc (Hons), Greg Gamble, MSc, Nicola Dalbeth, MBChB, MD, FRACP: University of Auckland, Auckland, New Zealand; [2]Sarah Stewart, PhD: University of Auckland and Auckland University of Technology, Auckland, New Zealand; [3]Lisa K. Stamp, MBChB, PhD, FRACP: University of Otago, Christchurch, Christchurch, New Zealand; [4]William J. Taylor, MBChB, PhD, FRACP: University of Otago, Wellington, Wellington, New Zealand.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25584).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25584.

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

Submitted for publication December 16, 2024; accepted in revised form June 11, 2025.

SIGNIFICANCE & INNOVATIONS

- Gout remission was analyzed using the simplified gout remission criteria, a definition for clinical remission in gout recently endorsed by the Gout, Hyperuricemia, and Crystal-Associated Disease Network.
- The fulfillment of gout remission over extended periods is currently underexplored in the literature. This is the first study to evaluate gout remission over a 6-year clinical trial, focusing on trends in rates of remission over time and identifying baseline variables associated with achieving remission.
- In this study, there was an increase in the rates of remission over time and baseline age, race, greater disease severity, presence of comorbidities, and febuxostat treatment were variables significantly associated with remission.

fulfillment of gout remission over more extended periods are unknown. Using this new simplified definition of gout remission, this post hoc analysis of the Cardiovascular Safety of Febuxostat or Allopurinol in Patients With Gout and Cardiovascular Comorbidities' (CARES) trial[8] aimed to describe the time trends in remission rates over 6 years of follow-up among people with gout receiving ULT and to identify baseline variables that predict remission.

PATIENTS AND METHODS

A post hoc analysis was conducted using data from the CARES trial.[8] The CARES trial dataset was accessed through the Vivli Platform (request number 8874). This trial has previously been described[8]; in brief, all participants had gout according to the 1977 American Rheumatism Association preliminary gout classification criteria[9] and a history of major cardiovascular disease. Participants were assigned 1:1 to receive once-daily febuxostat or allopurinol, which were titrated to achieve serum urate concentration <0.36 mmol/L (6 mg/dL). Participants were observed prospectively from baseline through to dropout, death, or the end of the 6-year study, whichever occurred first.[8]

Gout remission definition and timepoints. In this analysis, gout remission was assessed using the simplified definition: (a) absence of gout flares over 12 months (gout flares domain), (b) serum urate <0.36 mmol/L at least twice over 12 months at equal distances apart (serum urate domain), and (c) absence of tophus (tophus domain). Study visits were scheduled every 6 months. Remission at year 1 was measured using clinical assessments at month 6 and month 12, year 2 was measured using clinical assessments at month 18 and month 24, year 3 was measured at month 30 and month 36, year 4 at month 42 and month 48, year 5 at month 54 and month 60, and year 6 at month 66 and month 72. Remission was determined for participants with serum urate, tophi, and gout flare assessments from these visits (Supplementary Figure 1).

Statistical analysis. In the CARES trial, there were 6,190 participants enrolled at baseline who received trial medication. There was significant loss to follow-up (45% of participants missed at least one study visit), and 56.6% discontinued the study drug. In this analysis, only participants with at least 1 year of follow-up data were included, and those who discontinued the study drug over the study period were still included in the analysis. Missing data were handled using available case analysis, in which remission was determined only for participants with available data at the analyzed time points without imputation.

Baseline demographics and clinical variables were summarized using standard descriptive statistics, including mean, SD, count, and percent, as appropriate. Remission status and fulfillment of the individual remission domains were described using count and percent.

Logistic regression was used to evaluate the association between intervention group and the fulfillment of individual domains within the simplified criteria, as well as to assess the association between intervention group and remission status, for each year. To account for clustering of patients within years, a generalized linear mixed model with nested random effects was applied to estimate intervention effects on overall remission status. Random intercepts were included for patients and for years nested within patients, with a compound symmetry covariance structure assumed for the repeated measurements.

The 'survival' package in R was used to estimate the median time to first remission for both interventions. Right censoring was applied accordingly, accounting for participants who were lost to follow-up or did not reach remission over the 6-year period. Loss to follow-up was treated similarly to not reaching remission because it was approximately uniform across treatment groups and did not differ systematically. As such, it was assumed to be a random event rather than a competing risk for the event (remission). Cox regression was used to examine the association between the intervention groups and time to remission. The proportional hazards assumption was assessed using Schoenfeld residuals, which indicated that the proportional hazards assumption was met.

Baseline factors associated with the fulfillment of remission at least once over 6 years (selected based on expert clinical knowledge) were assessed using univariable and multivariable logistic regression analysis. The number of variables to be used in the final multivariable model was reduced by selecting from candidate baseline variables with associations of $P < 0.15$ in univariable analyses. From these candidate variables, backward elimination and a modern selection strategy, lasso logistic regression, were used to identify predictor(s) for achieving gout remission.[10–12]

Lasso regression includes an L1 penalty, which encourages sparsity by shrinking some coefficients to exactly zero. This makes lasso particularly useful for variable selection, as it automatically removes irrelevant or less important variables from the model.[10] From these approaches, a final multivariable logistic regression model was produced incorporating variables chosen based on biologic plausibility, parsimony, and goodness of fit. Statistical analysis was performed using R software version 4.3.1 and GraphPad Prism software version 9.3.1.; $P < 0.05$ was used to denote statistical significance.

RESULTS

Baseline clinical characteristics. The baseline clinical characteristics of those used in this analysis and the overall number of participants randomized are reported in Supplementary Table 1. Clinical characteristics were similar across both groups. Among the 4,301 included in the analysis, 2,158 were randomized to receive allopurinol and 2,143 to receive febuxostat. Overall, 21% of the participants had tophi at baseline, 61% had a serum urate concentration <0.54 mmol/L, and 11% had not experienced a gout flare in the past year or more. Of the participants analyzed, 1,693/4,301 discontinued the study drug and this was similar between the intervention groups: 832 (39%) in the allopurinol group and 861 (40%) in the febuxostat group.

Fulfillment of remission over 6 years. Achievement of remission increased over the 6 years. At year 1, 37.4% of participants (1,593/4,259) were in remission; at year 2, 47.8% (1,472/3,081); at year 3, 53.1% (1,182/2,225); at year 4, 55.3% (863/1,560); at year 5, 59.0% (589/999) and at year 6, 63.1% (322/510) (Figure 1).

Participants in the febuxostat group were significantly more likely to fulfill remission at year 1 and year 2. From year 3 through to year 6, there was no difference in the proportion of participants who achieved remission between the intervention groups (Table 1). Overall, when accounting for the clustering of patients within years, those in the febuxostat group had higher odds of getting in remission over the 6 years compared with those in the allopurinol group (odds ratio [OR] 1.29; 95% confidence interval [CI] 1.11–1.49; $P < 0.001$) (Table 1). Additionally, the proportion of missing data was comparable between the intervention groups over the 6 years, indicating that missingness was not systematically related to the intervention and did not bias the comparison of group performance (Supplementary Table 2).

Over the 6 years, 59.4% (2,554/4,301) of all enrolled participants with sufficient data achieved remission at least once. Among participants who achieved remission, 67% had remained on the study drug over the duration of the study.

The median time to the first episode of remission was 2 years for both intervention groups. Those in the febuxostat group had a

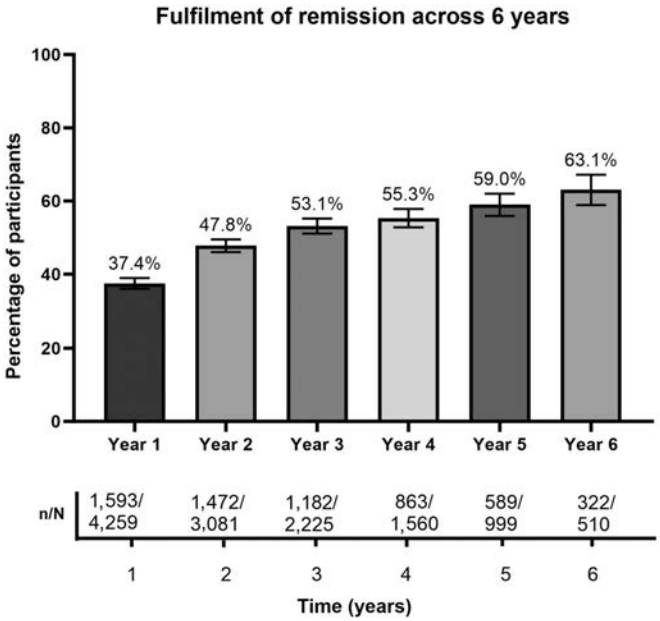

Figure 1. Fulfillment of the simplified remission criteria over 6 years.

10% higher hazard of achieving remission at any given time, with a hazard ratio of 1.10 (95% CI, 1.02–1.19; $P = 0.02$). (Figure 2).

Most participants in remission at year 1 maintained their remission status in subsequent years, ranging from 73.2% to 76.0% of participants (Figure 3). Among those who fulfilled remission at year 6, half (50.3%) had also fulfilled remission at year 1 (Figure 3).

Fulfillment of individual remission domains over 6 years. Fulfillment of the individual domains increased consistently over the 6 years. Data for individual domains are shown in Supplementary Figure 2.

Gout flares domain. At year 1, the gout flares domain was fulfilled by 71.2% of participants (3,033/4,259), rising to 85.8% at year 2 (2,645/3,081), 90.0% at year 3 (2,002/2,225) and year 4 (1,404/1,560), 92.3% at year 5 (922/999), and 92.7% in year 6 (473/510).

Serum urate domain. At year 1, the serum urate domain was fulfilled by 61.0% of participants (2,596/4,259), rising to 62.7% at year 2 (1,931/3,081), 65.2% at year 3 (1,450/2,225), 66.5% at year 4 (1,037/1,560), 67.8% at year 5 (677/999), and 69.8% at year 6 (356/510). In those not reaching the remission urate target over the 6 years, the mean ± SD serum urate concentration was 0.43 ± 0.05 mmol/L.

Tophus domain. At year 1, the tophus domain was fulfilled by 85.2% at year 1 (3,629/4,259), rising to 86.4% at year 2 (2,663/3,081), 88.3% at year 3 (1,965/2,225), 90.4% at year 4 (1,411/1,560), 93.1% at year 5 (930/999), and 95.5% at year 6 (487/510).

Fulfillment of domains by intervention group. There was no difference in fulfillment of the gout flares domain between the

Table 1. Fulfillment of remission by intervention*

	Febuxostat group, n/N (%; 95% CI)	Allopurinol group, n/N (%; 95% CI)	OR (95% CI)[a]	*P* value
Year 1	828/2,118 (39.1; 37.0–41.2)	765/2,141 (35.7; 33.7–37.8)	1.15 (1.02–1.31)	0.02
Year 2	778/1,551 (50.2; 47.7–52.7)	694/1,530 (45.4; 42.9–47.9)	1.21 (1.05–1.40)	0.008
Year 3	612/1,125 (54.4; 51.5–57.3)	570/1,100 (51.8; 48.9–54.8)	1.11 (0.94–1.31)	0.22
Year 4	446/789 (56.5; 53.1–60.0)	417/771 (54.1; 50.6–57.6)	1.10 (0.90–1.35)	0.33
Year 5	310/503 (61.6; 57.3–65.8)	279/496 (56.3; 51.9–60.6)	1.25 (0.97–1.61)	0.08
Year 6	170/263 (64.6; 58.7–70.2)	152/247 (61.5; 55.3–67.4)	1.14 (0.80–1.64)	0.47
Overall[b]	–	–	1.29 (1.12–1.49)	0.006

* CI, confidence interval; OR, odds ratio.
[a] Participants assigned to allopurinol were used as the reference.
[b] Adjusted for the clustering of patients within years.

intervention groups. However, at years 1 and 2, participants in the febuxostat group were significantly more likely to fulfill the serum urate domain. Also, at year 6, participants in the febuxostat group were significantly more likely to fulfill the tophus domain (Supplementary Tables 3–5).

Baseline predictors of at least one episode of remission. In the univariable analyses, many baseline factors were significantly associated with remission (Supplementary Table 6). Race, disease severity, and the presence of certain comorbidities were particularly associated with lower odds of achieving remission. The multivariable model is shown in Table 2; 17.1% of the variance in remission status was explained by this model (χ^2 = 576.75, df = 20, $P < 0.001$; Nagelkerke R^2 = 0.171; C-index = 0.709).

In the multivariable model, baseline age remained an independent predictor of remission. Older participants (aged ≥65 years) had significantly higher odds of achieving remission (OR 1.30; 95% CI 1.14–1.49) at least once over the 6 years. Race was also another demographic factor that remained an independent predictor. Compared with White participants, participants who were Black and from "other" races had significantly lower odds of achieving remission: the OR for Black participants was 0.65 (95% CI 0.54–0.78) and 0.59 (95% CI 0.48–0.72) for "other" participants.

Disease duration remained an independent predictor. Compared with participants with disease duration <5 years, the OR for participants with disease duration 5 to 10 years was 0.80 (95% CI 0.66–0.97) and for those with disease duration >10 years was 0.78 (95% CI 0.67–0.91). The presence of tophi, experience of multiple gout flares in the past year or more, and higher serum

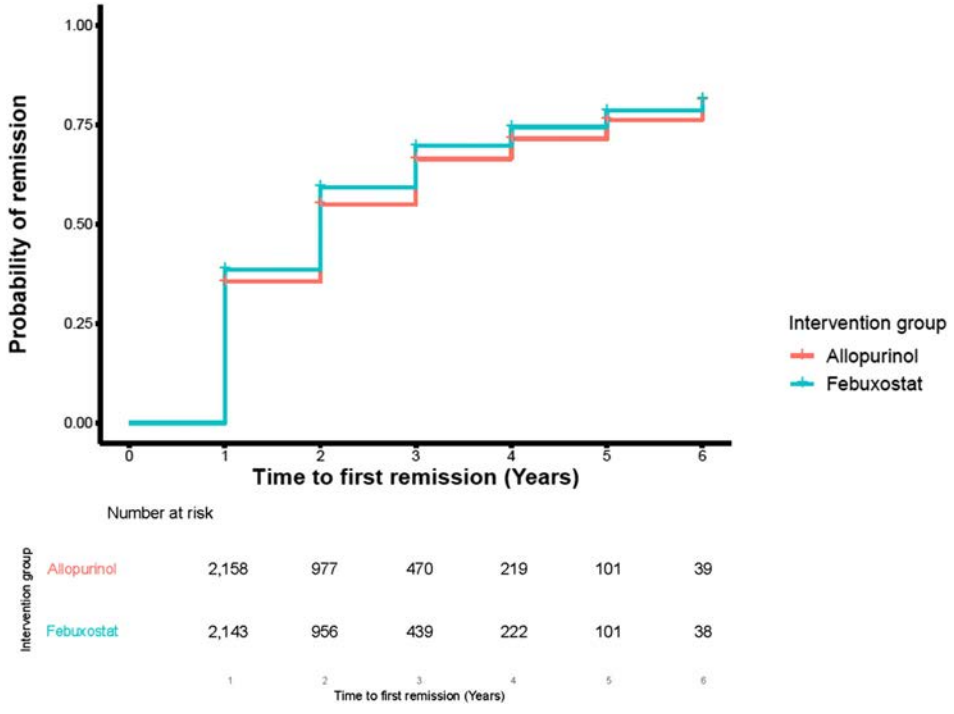

Figure 2. Survival curve for the time to the first episode of remission between the allopurinol and febuxostat intervention groups. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25584/abstract.

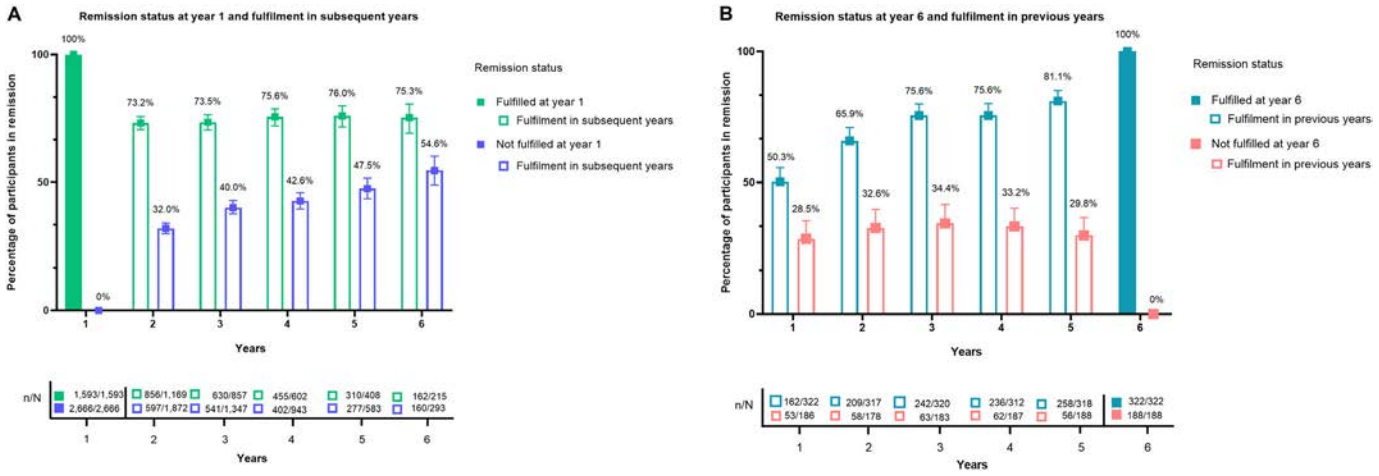

Figure 3. (A) Remission status at year 1 and fulfillment in subsequent years. (B) Remission status at year 6 and fulfillment in previous years.

urate concentrations all remained independent predictors negatively associated with experiencing at least one episode of remission. Neutrophil-lymphocyte ratio (NLR) was also an independent predictor; a higher NLR was associated with lower odds of experiencing remission, with an OR of 0.95 (95% CI 0.95–0.99).

The presence of comorbidities, including a history of diabetes, cerebrovascular disease, congestive heart failure, or peripheral vascular disease, were each independent predictors associated with lower odds of remission. With a history of diabetes, the OR for an episode of remission was 0.80 (95% CI 0.69–0.91); for cerebrovascular disease, the OR was 0.82 (95% CI 0.70–0.96); for congestive heart failure, it was 0.79 (95% CI 0.63–0.96); and for peripheral vascular disease, it was 0.78 (95% CI 0.63–0.96). In the multivariable model, febuxostat remained associated with greater odds of remission, with an OR of 1.23 (95% CI 1.08–1.40; $P = 0.002$).

DISCUSSION

A key goal of gout management should be the achievement of remission and then maintenance of remission in subsequent years. In this study, we described remission status over 6 years in people with gout and cardiovascular disease, using the simplified gout remission criteria. We also identified baseline factors that influence the fulfillment of gout remission over this period.

Our analysis has shown that gout remission is difficult to achieve in the first year of ULT. However, over time there was an increase in the rates of people achieving remission, from 37.4% at year 1 to 63.1% at year 6. This trend in the rates of remission over time is consistent with previous analyses. In our previous analysis of the simplified gout remission definition in a population of people with erosive gout receiving oral ULT over 2 years, rates of remission increased from 27% at year 1 to 44% at year 2.[7] Similarly, in our analysis comparing nurse-led care with usual gout care, remission increased from 18% at year 1 to 43% at year 2.[5] In an observational treat-to-target patient cohort with 5 years of follow-up, Uhlig et al[13] reported that rates of remission increased from 7.7% at year 1 to 45.4 % at year 2 and 58.6% at year 5.

Of the individual domains in this clinical trial, the serum urate domain appeared to be the most difficult to achieve as it had the lowest rates of fulfillment over time, with 69.8% of participants fulfilling it in year 6. This may be because of the maximum dose of ULT given to participants, which was 80 mg daily febuxostat and 600 mg daily allopurinol in participants with normal kidney function. Participants did not get further dose titration once these maximum doses were reached, as opposed to what may happen in clinical practice. This may also be because of the high rates of study drug discontinuation in this study. Low rates of continuous ULT are common in clinical practice and have been reported in previous studies, including a previous meta-analysis that calculated the overall pooled proportion of patients receiving regular uninterrupted ULT across studies from Europe, North America, Oceania, and Asia to be 26% (95% CI 18%–35%).[14] This highlights a key issue in gout management and an area of focus for the achievement of gout remission.

In contrast, the gout flare domain and tophus domain had higher rates of fulfillment, with 92.7% of participants and 95.5% of participants, respectively, at year 6. This, however, can vary depending on the baseline disease severity, as we observed in our study, and as previously reported by Alvarado-de la Barrera et al[15] in an observational study in which 44% of the participants had five or more tophi at baseline and were categorized as having severe gout. In that study, none were able to fulfill the tophus domain over the 5-year study period. The number and size of the tophi was the key reason for why some participants were unable to achieve gout remission, defined using the 2016 preliminary gout remission criteria. Collectively, these data show that once tophi are present, it takes a long time to achieve gout remission and highlights that in aiming for gout remission, people with gout need to begin treatment as early as possible (before tophus development) to increase the chances of being able to achieve remission.

Table 2. Reduced model for baseline variables associated with at least one episode of gout remission over 6 years*

Baseline variables	OR (95% CI)	*P* value
Age		
<65 y	Ref	–
≥65 y	1.30 (1.14–1.49)	<0.001
Sex		
Male	1.08 (0.89–1.31)	0.43
Female	Ref	–
Race		
Black	0.65 (0.54–0.78)	<0.001
White	Ref	–
Other	0.59 (0.48–0.72)	<0.001
Disease duration		
<5 y	Ref	–
5 to <10 y	0.80 (0.66–0.97)	0.02
≥10 y	0.78 (0.67–0.91)	0.002
Presence of tophus		
0	Ref	–
1 to <5	0.31 (0.26–0.36)	<0.001
≥5	0.10 (0.07–0.15)	<0.001
Serum urate		
<0.54 mmol/L	Ref	–
0.54–0.59 mmol/L	0.83 (0.70–0.98)	0.03
0.60–0.65 mmol/L	0.73 (0.59–0.91)	0.005
>0.65 mmol/L	0.60 (0.46–0.77)	<0.001
Number of gout flares in the last year		
0	Ref	–
1–3	0.64 (0.50–0.80)	<0.001
4–6	0.53 (0.41–0.70)	<0.001
>6	0.51 (0.38–0.68)	<0.001
NLR	0.95 (0.90–0.99)	0.03
Diabetes		
Yes	0.80 (0.69–0.91)	<0.001
No	Ref	–
Cerebrovascular disease		
Yes	0.82 (0.70–0.96)	0.014
No	Ref	–
Congestive heart failure		
Yes	0.79 (0.63–0.96)	0.007
No	Ref	–
Peripheral vascular disease		
Yes	0.78 (0.63–0.96)	0.02
No	Ref	–
Intervention group		
Allopurinol	Ref	–
Febuxostat	1.23 (1.08–1.40)	0.002

* χ^2 test = 576.75, df = 20, $P < 0.001$; Nagelkerke $R^2 = 0.171$; and C-index = 0.709. CI, confidence interval; NLR, neutrophil-lymphocyte ratio; OR, odds ratio.

In this analysis, we also observed that, of those fulfilling remission at year 1, 73.2% to 76.0% maintained their remission status in subsequent years. Similarly, of those participants who were in remission in year 6, 50% were also in remission in year 1. This implies that most people who are in remission early in treatment can maintain that status over a long period. Also, for those not in remission at year 1, 32.0% to 54.6% had fulfilled remission in subsequent years, which provides encouragement for the continuation of ULT in aiming for gout remission.

In this study, over the six years, >50% of participants were able to experience at least one episode of remission. In the first 2 years, those in the febuxostat group were significantly more likely to achieve remission than those in the allopurinol group. This is primarily because of the fulfillment of the serum urate domain, which was seen more in the febuxostat group during years 1 and 2. It may be that the intensive serum urate lowering afforded by febuxostat is beneficial for earlier achievement of remission. The intensive urate lowering may have also contributed to the significantly higher odds of people in the febuxostat group fulfilling the tophus domain in year 6.

In this study, multiple baseline variables were independently associated with remission. In the multivariable analysis, febuxostat treatment remained significantly associated with increased odds of remission. Of the demographic variables, older age

(≥65 years) was associated with increased odds of remission. Non-White races (Black and "other") had lower odds of remission, indicating that non-White populations are less likely to experience remission over time. Inequities in gout management and outcomes across different racial groups are well documented,[16–18] and our results suggest that, in treatment aimed at gout remission, there needs to be clinical support tailored to the diverse backgrounds and needs of specific communities. Baseline variables associated with greater disease burden, such as longer disease duration, presence of tophus, higher serum urate concentration, and higher frequency of gout flares, were all associated with lower odds of gout remission. Participants with disease duration >5 years, presence of multiple tophi, serum urate concentration >0.54 mmol/L, and multiple flares in preceding years may require intensive treatment strategies to experience remission and additionally may require long periods of time to be able to achieve remission.[15] Similarly, participants with comorbidities such as diabetes, cerebrovascular disease, congestive heart failure, and peripheral vascular disease may also require additional support to reach gout remission as these were other variables associated with lower odds of remission in the multivariable analysis.

This study had many strengths, including using a dataset that involved a large population of people with gout in a randomized controlled trial of ULT that had frequent visits assessing serum urate, gout flare, and tophus burden. Furthermore, the data were collected over multiple years, thus enabling the longitudinal analysis of gout remission over 6 years. A potential limitation of our analysis is the risk of overadjustment and the misinterpretation of confounder and modifier coefficients, as discussed by Westreich and Greenland.[19] Although we carefully selected the variables and interpreted the results within the broader context, including the univariable analyses provided in the supplementary materials, we acknowledge that residual confounding may still be a concern. A key limitation of this study was the high rate of loss to follow-up, missing data, and study drug discontinuation, which may have influenced remission estimates. Our use of available case analysis, although a practical approach, may introduce bias through overestimation or underestimation of remission rates. Additionally, this approach may affect the validity and generalizability of our findings as participants with available data may not be fully representative of the entire study population. Also, the presence of cardiovascular disease was a key inclusion criterion for the CARES study and as such the results may not be generalizable to other populations of people with gout.

In conclusion, >50% of participants achieved remission at least once over 6 years. More participants receiving febuxostat achieved remission, reflected by faster initial urate lowering. We identified baseline predictors that highlight people with gout who may need more proactive and intensive management to achieve remission.

ACKNOWLEDGMENTS

This publication is based on research using data from the data contributor, Takeda, that has been made available through Vivli, Inc. Vivli has not contributed to or approved, and is not in any way responsible for, the contents of this publication. Open access publishing facilitated by The University of Auckland, as part of the Wiley - The University of Auckland agreement via the Council of Australian University Librarians.

AUTHOR CONTRIBUTIONS

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

REFERENCES

1. Dalbeth N, Choi HK, Joosten LAB, et al. Gout. Nat Rev Dis Primers 2019;5(1):69.
2. Felson D. Defining remission in rheumatoid arthritis. Ann Rheum Dis 2012;71(0 2)(suppl 2):i86–i88.
3. Richette P, Dalbeth N. Treat-to-target or treat-to-dissolve strategy to improve gout treatment. Nat Rev Rheumatol 2024;20(7):393–394.
4. de Lautour H, Taylor WJ, Adebajo A, et al. Development of preliminary remission criteria for gout using Delphi and 1000Minds consensus exercises. Arthritis Care Res (Hoboken) 2016;68(5):667–672.
5. Tabi-Amponsah AD, Doherty M, Sarmanova A, et al. Performance of two gout remission definitions in a two-year randomized controlled trial of nurse-led versus usual gout care. Semin Arthritis Rheum 2024;69:152555.
6. Tabi-Amponsah AD, Stamp LK, Horne A, et al. Analysis of gout remission definitions in a randomised controlled trial of colchicine prophylaxis for people with gout initiating allopurinol. J Rheumatol 2024; 51(11):1135–1140.
7. Tabi-Amponsah AD, Stewart S, Gamble G, et al. Baseline dual-energy computed tomography urate volume predicts fulfillment of gout remission after two years of urate-lowering therapy. Arthritis Care Res (Hoboken) 2024;76(12):1657–1665.
8. White WB, Saag KG, Becker MA, et al; CARES Investigators. Cardiovascular safety of febuxostat or allopurinol in patients with gout. N Engl J Med 2018;378(13):1200–1210.
9. Wallace SL, Robinson H, Masi AT, et al. Preliminary criteria for the classification of the acute arthritis of primary gout. Arthritis Rheum 1977;20(3):895–900.
10. Sauerbrei W, Perperoglou A, Schmid M, et al; for TG2 of the STRATOS initiative. State of the art in selection of variables and functional forms in multivariable analysis-outstanding issues. Diagn Progn Res 2020;4(1):3.
11. Chowdhury MZI, Turin TC. Variable selection strategies and its importance in clinical prediction modelling. Fam Med Community Health 2020;8(1):e000262.
12. Heinze G, Wallisch C, Dunkler D. Variable selection - a review and recommendations for the practicing statistician. Biom J 2018;60(3): 431–449.

13. Uhlig T, Stjärne J, Karoliussen LF, et al. Remission in gout is possible: 5-year follow-up in the NOR-Gout study. Semin Arthritis Rheum 2025; 72:152698.
14. Son CN, Stewart S, Su I, et al. Global patterns of treat-to-serum urate target care for gout: systematic review and meta-analysis. Semin Arthritis Rheum 2021;51(4):677–684.
15. Alvarado-de la Barrera C, López-López CO, Álvarez-Hernández E, et al. Are target urate and remission possible in severe gout? A five-year cohort study. J Rheumatol 2020;47(1):132–139.
16. Te Karu L, Dalbeth N, Stamp LK. Inequities in people with gout: a focus on Māori (indigenous people) of Aotearoa New Zealand. Ther Adv Musculoskelet Dis 2021;13:1759720X2 11028007.
17. Yokose C, McCormick N, Lu N, et al. Nationwide racial/ethnic disparities in US emergency department visits and hospitalizations for gout. Rheumatology (Oxford) 2023;62(6):2247–2251.
18. McCormick N, Choi HK. Racial disparities in the modern gout epidemic. J Rheumatol 2022;49(5):443–446.
19. Westreich D, Greenland S. The table 2 fallacy: presenting and interpreting confounder and modifier coefficients. Am J Epidemiol 2013; 177(4):292–298.

Arthritis Care & Research
Vol. 77, No. 12, December 2025, pp 1426–1435
DOI 10.1002/acr.25594

Proportion of Acceptable Symptom State Nearly Tripled With Improvements in Patient-Reported Outcomes for All Symptom State Subgroups: A Registry Study of More Than 15,000 Patients With Osteoarthritis in Digital Education and Exercise Therapy

Ali Kiadaliri,[1] L. Stefan Lohmander,[2] Amanda E. Nelson,[3] and Leif E. Dahlberg[4]

Objective. This study investigated trajectories of patient acceptable symptom state (PASS) among participants of digital education and exercise therapy for knee and hip osteoarthritis.

Methods. A longitudinal observational study among individuals aged at least 40 years who participated in the digital program. Participants completed PASS (yes/no) at enrollment and at least one follow-up during 1 year after enrollment (N = 15,253). Group-based trajectory modeling was used to identify groups with distinct PASS trajectories. We used multinomial logistic regression and linear random intercept models to explore predictors and compare changes in patient-reported outcome measures (PROMs) across trajectory subgroups.

Results. The proportion of participants reporting acceptable symptom state rose from 17.4% (95% confidence interval [CI] 16.8%–18.1%) at enrollment to 42.4% (95% CI 41.6%–43.1%) and 48.9% (95% CI 47.5%–50.2%) at 3- and 12-month follow-ups, respectively. We identified four PASS trajectories: (1) "persistently not achieving PASS" (PNAP) (45.1%), (2) "early sustained PASS" (ESP) (34.8%), (3) "gradually increasing satisfaction" (GIS) (10.8%), and (4) "early PASS, later unacceptable PASS" (9.3%). Among baseline variables, female sex, older age, nonmetropolitan residence, lower education, knee osteoarthritis, fear of movement, no walking difficulties, no wish for surgery, and better PROMs were generally associated with higher odds of following trajectories other than the PNAP. All trajectories experienced improvements in PROMs, with generally larger improvements in the ESP and GIS groups than the other two groups.

Conclusion. The percentage of participants achieving PASS almost tripled at 12 months. Improvements in PROMs across all PASS trajectories highlights the importance of distinction between "feeling better" and "feeling good."

INTRODUCTION

Clinical guidelines consistently advocate for education, exercise, and weight management as first-line treatments for all persons with osteoarthritis (OA),[1] even though these treatments are underused.[2] Although traditionally first-line treatments are delivered face-to-face, there has been a considerable increase in digital delivery in recent years. Regardless of the mode of delivery, education and exercise therapy reduce pain and improve function and health-related quality of life.[3–5]

In addition to patient-reported outcome measures (PROMs), an emerging concept in evaluating the treatment outcomes and success from the patient's perspective is patient acceptable symptom state (PASS).[6] PASS is a single dichotomized question (yes/no) outcome tool to evaluate a patient's satisfaction with their current state of symptoms and response to treatment.[7] It should be noted that even though in the original PASS question introduced by Tubach et al,[6] "current PASS" was evaluated with no specification of the time spent in the state, other anchor questions have been developed incorporating the time spent in the state

Supported by the Foundation for People with Movement Disability in Skåne and the Greta and Johan Kock Foundation. Dr Nelson's work was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases, NIH (grant K24-AR-081368).

[1]Ali Kiadaliri, PhD: Lund University, Lund, and Karolinska Institute, Stockholm, Sweden; [2]L. Stefan Lohmander, MD, PhD: Lund University, Lund, Sweden; [3]Amanda E. Nelson, PhD: University of North Carolina, Chapel Hill; [4]Leif E. Dahlberg, MD, PhD: Lund University, Lund, and Arthro Therapeutics, Malmö, Sweden.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25594) .

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25594.

Address correspondence via email to Ali Kiadaliri, PhD, at ali.kiadaliri@med.lu.se.

Submitted for publication November 7, 2024; accepted in revised form June 23, 2025.

SIGNIFICANCE & INNOVATIONS

- We used registry data from participants of a digital education and exercise therapy with explored trajectories of patient acceptable symptom state (PASS) from before to 1 year after enrollment in the program.
- The percentage of individuals reporting acceptable symptom state nearly tripled at 1 year after enrollment.
- There were four distinct PASS trajectories with several factors including female sex, older age, lower education, fear of movement, no walking difficulties, no wish for surgery, and better patient-reported outcome measures generally associated with higher odds of following trajectories other than "persistently not achieving PASS."
- All PASS trajectories experienced improvements in patient-reported outcome measures, highlighting the importance of distinguishing between "feeling better" and "feeling good."

including "the next few months" (also known as "future PASS") and "the rest of your life" (known as "lifelong PASS").[8] Moreover, anchor questions with multiple response choices (eg, Likert-style questions) have also been developed.[9,10] The PASS anchor question can be domain-specific (eg, pain, symptoms, or function) or generic, covering multiple domains.[8,11] It has also been suggested that PASS can serve as a clinical benchmark to facilitate comparison across different PROMs,[12] even though considerable variability in PASS estimates (eg, by patient characteristics and settings, types of intervention, type of anchors, wording of anchor question, and the response choices) may limit cross study comparability.[11,13]

Although PASS itself provides valuable insights about patients' current state and achieving PASS is a patient-centered outcome,[14] it has mainly been used as an anchor to determine the cutoff points for interpreting other PROMs in OA (ie, determining a clinically meaningful threshold for a PROM to classify patients to those who consider themselves "well" and satisfied with treatment and those who do not).[12,15,16] In this regard, whereas minimal important change (MIC) reflects the improvement ("feeling better"), PASS reflects the satisfaction with health state ("feeling good").[17] However, evidence on longitudinal changes in the responses to the PASS question itself following participation in the OA first-line treatments including digital treatment is nonexistent. Studies that explore factors associated with achieving PASS among patients with OA receiving first-line treatments are limited.[11,18] Understanding these predictors is critical for tailoring interventions to improve patient satisfaction and outcomes.

This study aimed to monitor longitudinal changes in the responses to a PASS anchor question and to identify subgroups with distinct trajectories of PASS among individuals with knee or hip OA participating in a digital education and exercise program in Sweden. We also assessed the associations between pretreatment variables and assignment to these PASS trajectory subgroups. Finally, we compared the longitudinal changes in PROMs across identified PASS trajectory subgroups.

PATIENTS AND METHODS

Study design and setting. We conducted a retrospective observational longitudinal study using data from the Joint Academy registry, a digital platform for delivering education and exercise therapy for hip and knee OA in Sweden.[19] The digital program focuses on exercise, physical activity, and education and is delivered through a smartphone application. Participants in the digital program must have had either (1) a prior radiographic and/or clinical diagnosis of hip or knee OA from a physiotherapist or physician or (2) clinical OA confirmed through a telephone consultation or a physical visit with an orthopedic surgeon or physiotherapist. The self-management digital program includes explanatory video lectures on OA and physical activity; individualized exercises adjusted with participants' progression in the program, which is monitored by a physiotherapist through a dedicated digital interface; and goal setting guided by the their personal physiotherapist.[20] Participants receive regular supervision from their physiotherapist via telephone or video calls, with the option of asynchronous chat communications during participation in the program. The study was approved by the Swedish Ethical Review Authority (Dnr: 2021–01713, 2021–06-16) and performed in accordance with the Declaration of Helsinki. Digital informed consent was obtained from participants at enrollment.

Participants. We included all consecutive participants aged ≥40 years who enrolled in the program in Sweden between January 1, 2019, and September 30, 2021, and provided their digital consent for research at enrollment (n = 16,811). We excluded those who did not respond to any follow-up (n = 1,430) and those with no responses within 4 weeks from the assumed follow-up date (n = 128). Given the similarities in the baseline characteristics of individuals with knee and hip OA as well as comparable trajectories of adherence to the digital treatment and changes in pain among them,[21,22] we combined them in the present study and adjusted for the index joint (ie, knee vs hip OA) in our analyses.

PASS. PASS was evaluated by a "yes" or "no" response to the following question: "Considering your knee/hip function, do you feel that your current state is satisfactory? You should take all activities during your daily life, sport and recreational activities, your level of pain and other symptoms, and your knee/hip-related quality of life into account."[23]

PROMs. Participants in the digital program responded to several PROMs covering three core outcome domains in knee and hip OA (ie, pain, physical function, and health-related quality of life).[24] These PROMs include the 12-item short forms of the Knee Injury and OA Outcome Score (KOOS-12)[25] and Hip disability and OA Outcome Score (HOOS-12)[26] (both 0–100, worst to best), which assess pain, physical function, and joint-related quality of life; EuroQol 5-Dimension 5-Level Questionnaire (EQ-5D-5L) (the Swedish value set[27] ranging from –0.31 to 1, worst to best), which assesses health-related quality of life; and the daily activity impairment (0–10, best to worst) measured using the Work Productivity and Activity Impairment questionnaire. We calculated the KOOS-12/HOOS-12 summary score as the average of the KOOS-12/HOOS-12 pain, function, and quality of life scale scores. Performance-based function was measured using the 30-second chair stand test (30-s CST).[28] All these measures are reliable, valid, and commonly used among persons with OA.[25,26,28,29] The responses to PASS, PROMs, and the 30-s CST were self-reported through the app at enrollment and at 3-month intervals up to 1 year following enrollment.

Covariates. We included the following variables self-reported through the app at enrollment as covariates: sex, age, body mass index (BMI), place of residence (living in the three metropolitan municipalities of Stockholm, Gothenburg, or Malmö: yes/no), education (less than high school, high school, and college/university degree), physical activity ≥150 min/week (yes/no), index joint (knee or hip), readiness to do exercise ("How ready are you to start doing exercise for your knee/hip on a daily basis?" 0–10, not at all ready to extremely ready), wish for surgery (yes/no), fear of movement (yes/no), walking difficulties (yes/no), and self-reported doctor-diagnosed coexisting conditions (diabetes, lung diseases, balance troubles, rheumatoid arthritis, and cardiovascular diseases). Physical activity was measured using the responses to two validated questions designed by The Swedish National Board of Health and Welfare (see the Supplemental Materials for detailed information on these questions).[30]

Statistical analysis. Patient characteristics at enrollment were reported as mean and SD and numbers (proportions). We computed the standardized difference to compare baseline characteristics of included and excluded participants and applied a threshold of 0.1 to define a meaningful difference.[31] The differences were addressed using inverse probability of inclusion weighting in analyses.

We used Poisson regression with a robust error variance to estimate prevalence ratios quantifying the strength of the associations between covariates and PASS status at enrollment.[32] A Sankey diagram was used to visualize the participant PASS status during the follow-up. To estimate the changes in proportion (with 95% confidence intervals [CIs]) of persons with PASS "Yes" from enrollment up to 1 year after enrollment, we used multilevel mixed-effects Poisson regression with robust error variance.

To identify subgroups with similar trajectories of PASS status from enrollment to 1 year after enrollment, we used group-based trajectory modeling (GBTM) using Stata's "traj" command.[33] Because PASS was a binary outcome, we used logistic distribution in the estimation. The basic GBTM assumes missing at random, which is violated if, for example, persons unsatisfied with their symptoms were more likely to stop the treatment. To account for this, we used the dropout extension suggested by Haviland et al allowing the trajectory group-specific probability of dropout based on the previous value of the dependent variable.[34] We estimated models with two to five groups with all possible combinations of trajectory shapes (from linear to cubic). We then sorted these models based on the absolute Bayesian information criterion (BIC) from the smallest to largest. The final trajectory model had the smallest absolute BIC while fulfilling all the following criteria: (1) average posterior probability of assignment >0.7 for each group, (2) the odds of correct classification less than five for each group, 3) no more than five percent of participants assigned to each group; and (4) the relative entropy, a measure of the degree of separation between groups, >0.7.[33,35] Days from enrollment in each response occasion was used as our time variable. After selecting the final model, each participant was assigned to the trajectory group with the highest posterior probability of membership. Subsequently, we used multinomial logistic regression to explore the associations between the PROMs and covariates (described above) at enrollment and identified PASS trajectories.

To compare the changes in PROMs across the identified PASS trajectories from enrollment up to 1 year after enrollment, we used a linear random intercept model with robust SEs. Separate models adjusted for all covariates were estimated for each PROM. The results are reported as predicted marginal means with 95% CI using the "margin" command in Stata. The KOOS-12/HOOS-12 scores at enrollment were missing for 1,422 (8.5%) participants (included and excluded). We used multiple imputation by chained equations to fill these missing values at enrollment (10 imputations) using PROMs other than KOOS-12/HOOS-12 and covariates, all measured at enrollment, as input for the imputation model. In the main analyses, we included the KOOS-12/HOOS-12 summary score as a predictor/covariate, in a sensitivity analysis we included KOOS-12/HOOS-12 domain-specific scores. All statistical analyses were implemented in Stata v.18.

RESULTS

We included 15,253 persons with mean ± SD age of 64.3 ± 8.8 years, and 75.3% were female (Table 1). Although the proportion of PASS responders were similar among those included and excluded from the study, there were meaningful differences in

Table 1. Baseline characteristics of the participants included and excluded from the study*

	Included (n = 15,253)	Excluded (n = 1,558)	Absolute standardized difference	
			Before weighting	After weighting
Female, n (%)	11,478 (75.3)	1,126 (72.3)	0.068	0.011
Age, mean (SD), years	64.3 (8.8)	65.9 (9.8)	0.170	0.046
Living in metropolitan cities, n (%)	2,949 (19.3)	336 (21.6)	0.055	0.002
Education, n (%)				
Less than high school	1,327 (8.7)	154 (9.9)	0.041	0.017
High school	5,521 (36.2)	542 (34.8)	0.029	0.023
College/university	8,405 (55.1)	862 (55.3)	0.004	0.013
Body mass index, mean (SD), kg/m^2	27.1 (4.7)	27.2 (5.0)	0.025	0.009
Physically active ≥150 min/week, n (%)	10,885 (71.4)	1,015 (65.2)	0.134	0.005
Diabetes, n (%)	849 (5.6)	140 (9.0)	0.132	0.007
Lung diseases, n (%)	1,589 (10.4)	191 (12.3)	0.058	0.002
Balance troubles, n (%)	506 (3.3)	83 (5.3)	0.099	0.010
Rheumatoid arthritis, n (%)	689 (4.5)	104 (6.7)	0.094	0.000
Cardiovascular diseases, n (%)	1,145 (7.5)	174 (11.2)	0.126	0.013
Knee as the index joint, n (%)	9,120 (59.8)	831 (53.3)	0.130	0.019
Walking difficulties, n (%)	9,976 (65.4)	1,036 (66.5)	0.023	0.004
Fear of movement, n (%)	2,312 (15.2)	222 (14.3)	0.026	0.009
Wish for surgery, n (%)	2,309 (15.1)	304 (19.5)	0.117	0.014
Readiness for exercise (0–10), mean (SD)	9.3 (1.3)	8.8 (1.7)	0.317	0.005
30-s CST, mean (SD)	12.7 (4.3)	12.2 (4.5)	0.111	0.009
KOOS-12/HOOS-12 pain (0–100), mean (SD)	52.9 (16.1)	52.4 (17.2)	0.033	0.013
KOOS-12/HOOS-12 function (0–100), mean (SD)	62.4 (18.7)	61.6 (19.6)	0.044	0.009
KOOS-12/HOOS-12 QoL (0–100), mean (SD)	46.0 (16.7)	45.9 (17.9)	0.007	0.004
KOOS-12/HOOS-12 summary score (0–100), mean (SD)	53.8 (15.1)	53.3 (16.2)	0.032	0.010
Activity impairment (0–10), mean (SD)	3.9 (2.4)	4.1 (2.5)	0.073	0.007
EQ-5D-5L index (−0.31 to 1), mean (SD)	0.85 (0.17)	0.82 (0.20)	0.149	0.006
PASS, n (%)	2,639 (17.3)	282 (18.1)	0.021	0.009

* 30-s CST, 30-second chair stand test; EQ-5D-5L, EuroQol 5-Dimension 5-Level Questionnaire; HOOS-12, Hip disability and Osteoarthritis Outcome Score; KOOS-12, Knee Injury and Osteoarthritis Outcome Score; PASS, patient acceptable symptom state; QoL, quality of life.

several characteristics at enrollment that disappeared after using inverse probability of inclusion weighting. Among those included in the study, 2,639 (17.3%) responded "Yes" to the PASS question at enrollment. A higher proportion of achieving PASS (ie, responding "Yes" to the PASS question) at enrollment was reported for participants with older age, low education, normal weight, coexisting cardiovascular diseases, hip OA, no walking difficulties, no fear of movement, no wish for surgery, lower readiness for exercise, and better PROMs (Table 2 and Supplementary Table S1).

In terms of the number of responses to the PASS question, 7,872 (51.6%) participants provided four to five responses to the PASS question over the study period with a slightly lower number of responses among those with PASS "Yes" than "No" status at enrollment (Supplementary Table S2). The flow of the participant responses to the PASS question from enrollment to 1 year after enrollment is displayed in the Sankey diagram (Figure 1). There were 104 unique patterns of responses to the PASS question, with the most common pattern (12.7% of all responses) being a "No" response to PASS at enrollment and the 3-month follow-up with missing responses at the remaining follow-ups (Supplementary Table S3). Among participants, 5,593 (39.3%) never achieved PASS (ie, PASS "No" status), 1,749 (11.5%) always achieved PASS (ie, PASS "Yes" status), and 7,511 (49.2%) changed their PASS. The proportion of individuals with a PASS "Yes" response rose from 17.4% (95% CI 16.8–18.1) at enrollment to 42.4% (95% CI 41.6–43.1), 46.8% (95% CI 45.9–47.7), 50.0% (95% CI 49.0–51.1) and 48.9% (95% CI 47.5–50.2) at 3, 6, 9, and 12 month follow-ups, respectively.

GBTM suggested that a four-trajectory group model best fit the data. Although several models with five groups had smaller BIC, they did not fulfill at least two other criteria (Supplementary Table S4). The largest trajectory group (45.1% of participants) included individuals who followed a "persistently not achieving PASS" trajectory (Figure 2). Approximately 34.7% of participants followed an early PASS achievement that sustained over follow-up ("early sustained PASS"), whereas 10.8% of participants followed a trajectory characterized by "gradually increasing satisfaction." The smallest group included 9.3% of participants who achieved PASS early with later unacceptable PASS. The baseline characteristics of participants assigned to these trajectory groups are reported in Supplementary Table S5. Among covariates measured at enrollment, female sex, older age, living outside metropolitan cities, lower education, knee OA, fear of movement, no walking difficulties, no wish for surgery, and better PROMs were generally associated with higher odds of not being in the "persistently not achieving PASS" group (Table 3). The results were similar when the KOOS-12/HOOS-12 domain scale scores were

Table 2. Associations between participants' characteristics and PASS at enrollment*

	PASS: no (n = 12,614)	PASS: yes (n = 2,639)	Prevalence ratio (95% CI)
Female, n (%)	9,482 (75.2)	1,996 (75.6)	1.02 (0.94–1.11)
Age, mean (SD), years	63.9 (8.8)	66.1 (8.6)	1.02 (1.02–1.03)
Living in metropolitan cities, n (%)	2,430 (19.3)	519 (19.7)	1.02 (0.94–1.12)
Education, n (%)			
Less than high school	1,035 (8.2)	292 (11.1)	1.00 (ref)
High school	4,511 (35.8)	1,010 (38.3)	0.83 (0.74–0.93)
College/university	7,068 (56.0)	1,337 (50.7)	0.72 (0.65–0.81)
Body mass index, n (%)			
<25 kg/m^2	4,483 (35.5)	1,063 (40.3)	1.00 (ref)
25–29.9 kg/m^2	5,162 (40.9)	1,065 (40.4)	0.89 (0.83–0.96)
≥30 kg/m^2	2,969 (23.5)	511 (19.4)	0.76 (0.69–0.84)
Physically active ≥150 min/week, n (%)	8,986 (71.2)	1,899 (72.0)	1.03 (0.96–1.12)
Diabetes, n (%)	687 (5.5)	162 (6.1)	1.11 (0.96–1.28)
Lung diseases, n (%)	1,331 (10.6)	258 (9.8)	0.93 (0.83–1.05)
Balance troubles, n (%)	435 (3.5)	71 (2.7)	0.81 (0.65–1.01)
Rheumatoid arthritis, n (%)	571 (4.5)	118 (4.5)	0.99 (0.83–1.17)
Cardiovascular diseases, n (%)	921 (7.3)	224 (8.5)	1.14 (1.01–1.29)
Knee as the index joint, n (%)	7,603 (60.3)	1,517 (57.5)	0.91 (0.85–0.98)
Walking difficulties, n (%)	8,961 (71.0)	1,015 (38.5)	0.33 (0.31–0.35)
Fear of movement, n (%)	2,038 (16.2)	274 (10.4)	0.65 (0.58–0.73)
Wish for surgery, n (%)	2,177 (17.3)	132 (5.0)	0.30 (0.25–0.35)
Readiness for exercise (0–10), mean (SD)	9.3 (1.3)	9.0 (1.5)	0.89 (0.87–0.91)
30-s CST, mean (SD)	12.6 (4.3)	13.4 (4.4)	1.03 (1.03–1.04)
KOOS-12/HOOS-12 summary score (0–100), mean (SD)	51.6 (14.2)	64.4 (14.1)	1.65 (1.61–1.69)[a]
Activity impairment (0–10), mean (SD)	4.2 (2.3)	2.5 (2.0)	0.74 (0.73–0.76)
EQ-5D-5L index (−0.31 to 1), mean (SD)	0.83 (0.17)	0.92 (0.12)	1.69 (1.56–1.82)[b]

* 30-s CST, 30-second chair stand test; CI, confidence interval; EQ-5D-5L, EuroQol 5-Dimension 5-Level Questionnaire; HOOS-12, Hip disability and Osteoarthritis Outcome Score; KOOS-12, Knee Injury and Osteoarthritis Outcome Score; PASS, patient acceptable symptom state.
[a] Per 10-unit difference.
[b] Per 0.1-unit difference.

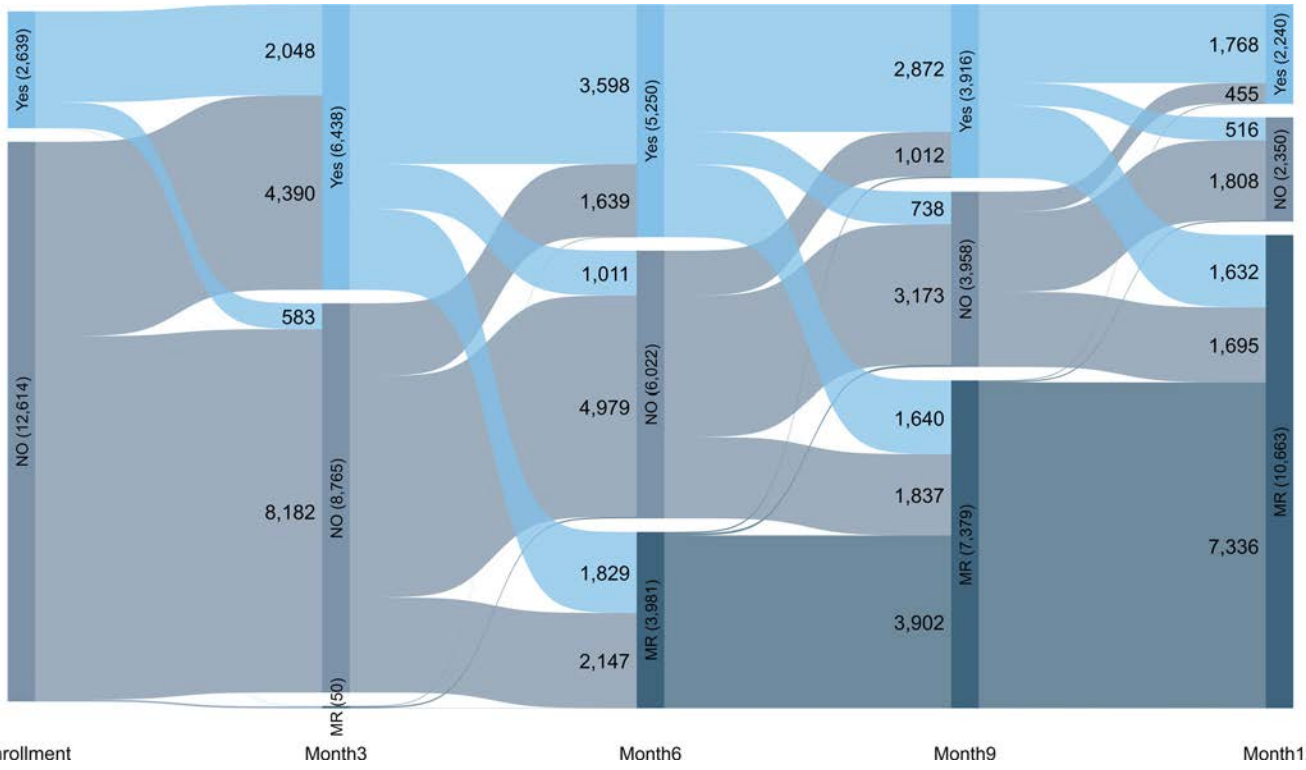

Figure 1. Flow of response to the question on patient acceptable symptom state. MR, missing response; No, responded "no"; Yes, responded "yes."

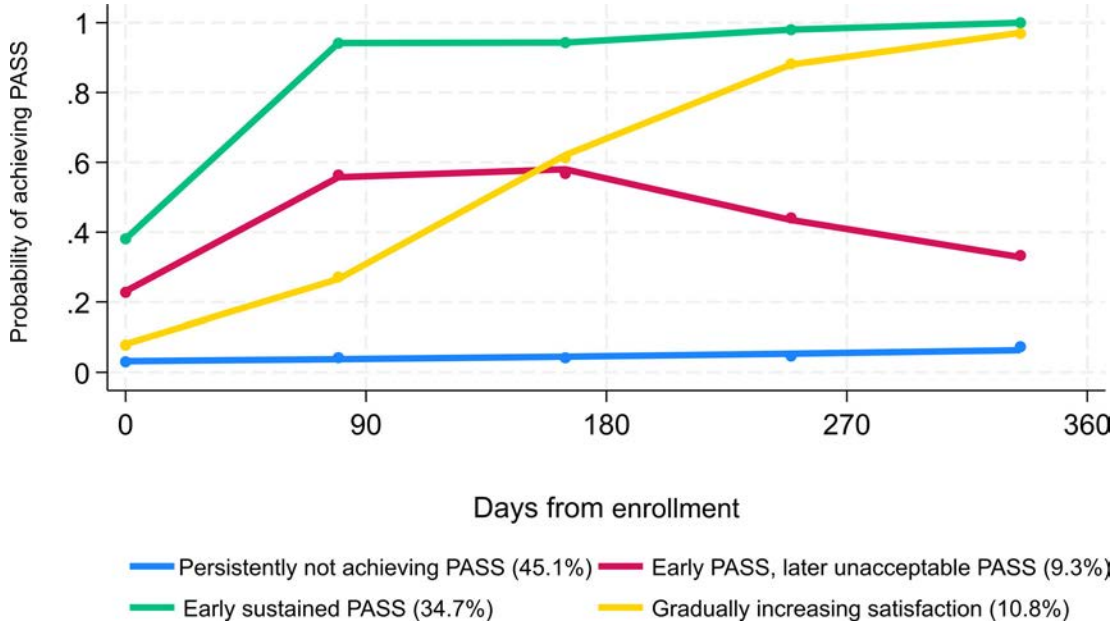

Figure 2. Subgroups with distinct trajectory of PASS from enrollment up to 1-year after enrollment. PASS< patient acceptable symptom state. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25594/abstract.

used (Supplementary Table S6). Although all trajectory groups experienced improvements in PROMs, the magnitude of these improvements were generally larger for those in "early sustained PASS" and "gradually increasing satisfaction" groups compared with the other two groups (Figure 3, Supplementary Figure S1 and Supplementary Table S7).

Table 3. ORs (95% CIs) for patient acceptable symptom state trajectory groups from multinomial logit model ("persistently not achieving patient acceptable symptom state" group as the reference)*

	EPLUP, OR (95% CI)	ESP, OR (95% CI)	GIS, OR (95% CI)
Female	1.33 (1.15–1.53)	1.32 (1.20–1.45)	1.35 (1.18–1.55)
Age	1.02 (1.01–1.03)	1.02 (1.01–1.02)	1.01 (1.00–1.01)
Living in metropolitan cities	0.82 (0.71–0.96)	0.77 (0.70–0.85)	0.76 (0.66–0.88)
Education			
Less than high school	1.00	1.00	1.00
High school	0.74 (0.60–0.92)	0.76 (0.65–0.89)	0.86 (0.70–1.07)
College/university	0.57 (0.46–0.70)	0.51 (0.44–0.59)	0.73 (0.60–0.90)
Body mass index			
<25 kg/m^2	1.00	1.00	1.00
25–29.9 kg/m^2	1.07 (0.93–1.23)	1.00 (0.92–1.10)	0.97 (0.86–1.11)
≥30 kg/m^2	1.10 (0.93–1.30)	0.98 (0.87–1.10)	0.87 (0.74–1.02)
Physically active ≥150 min/week	0.95 (0.83–1.08)	1.05 (0.96–1.16)	1.00 (0.88–1.11)
Diabetes	1.07 (0.83–1.38)	1.05 (0.87–1.26)	1.47 (1.17–1.84)
Lung diseases	0.99 (0.82–1.20)	1.06 (0.93–1.22)	1.06 (0.88–1.27)
Balance troubles	1.16 (0.85–1.58)	0.90 (0.71–1.14)	1.00 (0.73–1.36)
Rheumatoid arthritis	1.15 (0.87–1.51)	1.15 (0.95–1.40)	0.91 (0.68–1.20)
Cardiovascular diseases	1.07 (0.86–1.34)	1.07 (0.92–1.25)	0.83 (0.66–1.05)
Knee as the index joint	1.16 (1.02–1.31)	1.13 (1.04–1.22)	1.17 (1.05–1.42)
Walking difficulties	0.65 (0.56–0.75)	0.53 (0.48–0.58)	0.77 (0.67–0.88)
Fear of movement	1.09 (0.92–1.30)	1.04 (0.92–1.17)	1.22 (1.05–1.42)
Wish for surgery	0.50 (0.41–0.61)	0.38 (0.33–0.44)	0.52 (0.44–0.62)
Readiness for exercise (0–10)	0.99 (0.95–1.04)	0.97 (0.94–1.00)	1.02 (0.98–1.07)
30-s CST	1.00 (0.99–1.02)	1.00 (0.99–1.01)	1.00 (0.99–1.01)
KOOS-12/HOOS-12 summary score (0–100)[a]	1.12 (1.05–1.19)	1.44 (1.38–1.51)	1.15 (1.09–1.23)
Activity impairment (0–10)	0.88 (0.85–0.91)	0.91 (0.89–0.93)	0.97 (0.94–1.01)
EQ-5D-5L index (−0.31 to 1)[b]	1.04 (0.98–1.09)	1.05 (1.00–1.09)	1.07 (1.02–1.12)

* 30-s CST, 30-second chair stand test; CI, confidence interval; EPLUP, early patient acceptable symptom state, later unacceptable patient acceptable symptom state; EQ-5D-5L, EuroQol 5-Dimension 5-Level Questionnaire; ESP, early sustained patient acceptable symptom state; GIS, gradually increasing satisfaction; HOOS-12, Hip disability and Osteoarthritis Outcome Score; KOOS-12, Knee Injury and Osteoarthritis Outcome Score; OR, odds ratio.
[a] Per 10-unit difference.
[b] Per 0.1-unit difference.

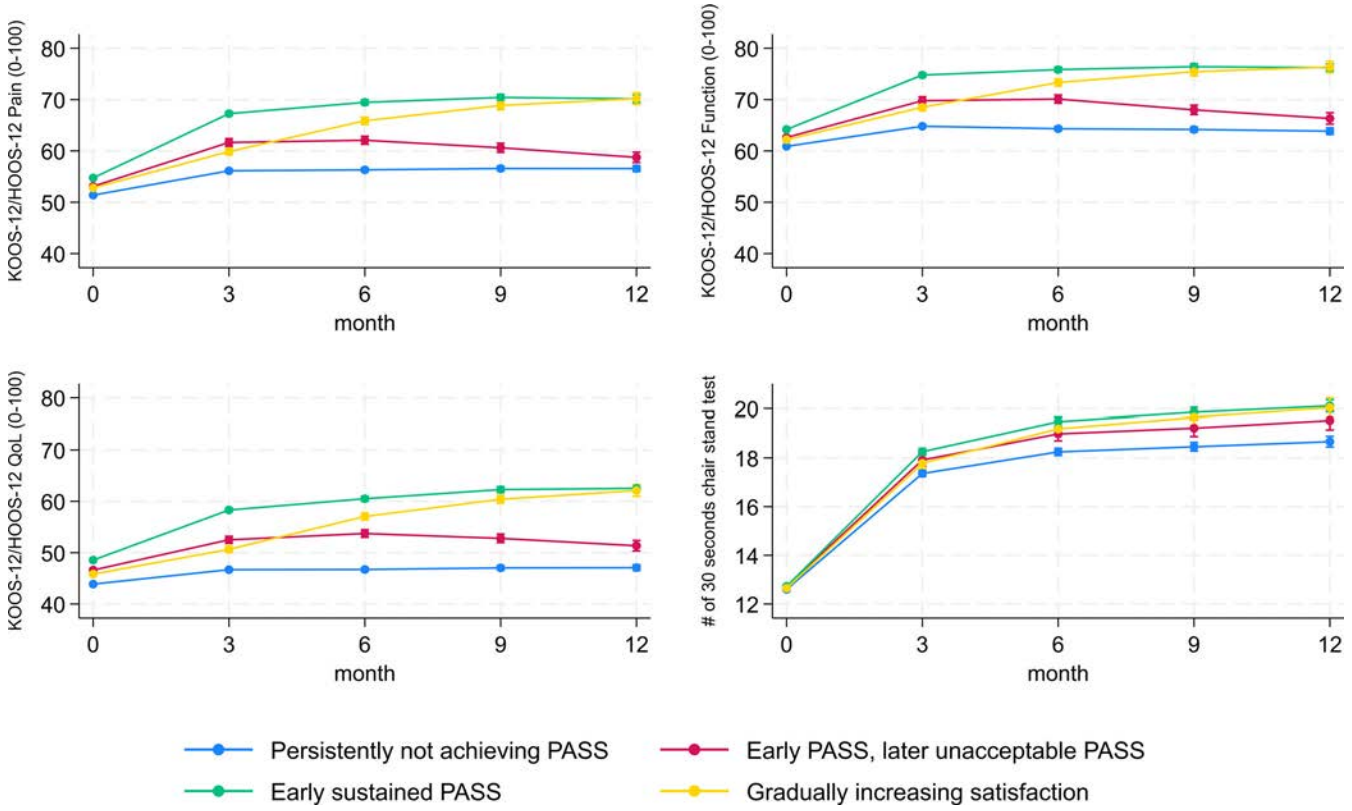

Figure 3. Predicted mean (95% confidence interval) in patient-reported outcome measures across PASS trajectory groups. Estimates obtained from random intercept model adjusted for covariates and the outcome of interest, all measured at enrollment. HOOS-12, Hip disability and Osteoarthritis Outcome Score; KOOS-12, Knee Injury and Osteoarthritis Outcome Score; PASS, patient acceptable symptom state; QoL, quality of life. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25594/abstract.

DISCUSSION

Our results showed that among participants in a digital education and exercise therapy program, the proportion of individuals reporting acceptable symptom state rose substantially from approximately 17% at enrollment to approximately 49% at 1 year after enrollment in the program. We identified four groups with distinct PASS trajectories over 1 year of participation in the program, with the largest group (45.1% of participants) persistently not achieving PASS, followed by a group (34.7% of participants) experiencing an early sustained PASS trajectory. Our results suggested that younger male patients with hip OA, poorer PROMs, walking difficulties, and a wish for surgery at enrollment were less likely to achieve PASS throughout participation in the program. As expected, larger improvements in PROMs were reported among those who experienced more favorable changes in their PASS.

Previous studies reported improvements in different PROMs, including pain, function, and quality of life, as well as a reduced desire for surgery following education and exercise therapy for knee and hip OA.[3–5,36] Consistent with these reports, our results suggest that the percentage of participants who considered themselves to have an acceptable symptom state almost tripled from before to 1 year after participation in a digital education and exercise therapy program. These improvements in PASS are of clinical importance and, similar to improvements in other PROMs, they may reduce the wish for surgery and delay surgery in those with OA.[36] As the most common use of the PASS question is to establish cutoff points for other PROMs, it has generally been collected at follow-up time points with little information on changes in the responses to the PASS question before and after an intervention. An exception is a previous study among a cohort of patients with knee OA receiving first-line treatments in the Netherlands, where the proportion of those achieving PASS rose from 49% at baseline to 56% at 3 months following treatment.[11] The between-study differences in the proportions and the magnitude of changes are likely because of differences in patient characteristics (eg, younger age, more male patients and higher BMI compared with our study), intervention (face-to-face vs digital treatment), and most importantly the difference in the PASS question, where we investigated the "current PASS" whereas the previous study investigated the "lifelong PASS" (ie, "Think about all consequences of the knee OA in the last week. If you were to remain for the rest of your life as you were during the last week, would the current state be acceptable or unacceptable for you?"[11]).

We observed important individual variations in the responses to the PASS question over follow-up. This is consistent with previous studies suggesting large individual variability in the response

to exercise including digital exercise therapy.[20,37–39] In the present study, despite considerable rises in PASS proportions, approximately 45% of the participants were persistently not achieving PASS throughout the 1-year observation period. This suggests that digital education and exercise therapy may not be sufficient for all participants, and that additional modifications in the program, adjunct, and/or alternative treatments may be needed for this subgroup. In this regard, identifying patients who are less likely to benefit from the intervention can facilitate early modifications to treatment plans. In the present study, several characteristics, including male sex, younger age, living in metropolitan cities, higher education, hip OA, no fear of movement, walking difficulties, wish for surgery, and worse PROMs before intervention, were associated with following the "persistently not achieving PASS" trajectory. Consistent with our findings, previous studies mainly conducted in patients receiving surgical treatments for hip or knee OA have reported that better pretreatment PROMs were associated with a higher probability of achieving PASS (with PASS being defined as cutoffs in another PROM in these studies) after surgery.[18,40–44]

However, the findings on other factors are mixed. For example, although male sex, White race, no smoking, lower BMI, fewer comorbidities, and better preoperative PROMs were associated with higher probability of achieving PASS defined by the cutoff for HOOS–Joint Replacement (JR) among patients who underwent total hip arthroplasty,[40] no associations among age, sex, race, ethnicity, BMI, smoking status, and comorbidities were observed with achieving PASS defined by the cutoff for KOOS-JR among patients who underwent total knee arthroplasty.[44] Interestingly, Hadad et al reported mixed associations between the area deprivation index (a measure of socioeconomic disadvantage) and PASS achievement, depending on the PROM used to define the PASS, in a cohort of patients who underwent total knee arthroplasty.[41] The lower probability of achieving PASS throughout the participation into the program among participants with walking difficulties, wish for surgery, and worse PROMs at enrollment may suggest that initial symptom severity and functional impairment are critical factors influencing the outcomes of exercise therapy. This is consistent with recent findings suggesting that those with a relatively short symptom duration[39] or less disease severity[20] may benefit more from exercise therapy. This highlights the importance of providing the OA first-line treatments, including exercise therapy, as early as possible in OA management, as those with milder symptoms may be more responsive to these treatments. It should be noted that these findings do not mean that those with more disease severity should be excluded from first-line treatments, but some modifications might be needed for this subgroup.

Given that both PASS and PROMs capture patients' perceptions about their symptoms, larger improvements in PROMs among participants achieving PASS were anticipated. The estimated improvements in KOOS-12/HOOS-12 subscales, 30-s CST, and the Swedish EQ-5D-5L score in all PASS trajectory subgroups were clinically important except for the quality of life subscales and EQ-5D-5L in the "persistently not achieving PASS" subgroup.[16,45–47] A previous systematic review reported a minimal clinically important difference (MCID) of 6.5 for the KOOS quality of life subscale (which is the same as the KOOS-12 quality of life subscale) among nonsurgical patients with OA.[48] Accordingly, the differences between the "persistently not achieving PASS" and "early PASS, later unacceptable PASS" groups and the other two groups were clinically important at all follow-ups. Moreover, the differences in the Swedish EQ-5D-5L scores between the "persistently not achieving PASS" group and two groups with favorable changes in PASS (ie "early sustained PASS" and "gradually increasing satisfaction") were larger than the MCID.[47] A lack of MIC and MCID values for activity impairment in patients with OA hindered clinical interpretations of observed improvements and differences for this outcome.

The observed improvements in PROMs among those assigned to the "persistently not achieving PASS" trajectory reinforce the importance of the distinction between "feeling better" and "feeling good" when assessing patient outcomes.[49,50] In other words, experiencing a clinically relevant change in PROMs ("feeling better") does not guarantee that patients are satisfied with their current state ("feeling good"). Hence, it is important to complement the changes in PROMs with PASS status to provide a more holistic assessment of treatment outcomes in the OA context. We encourage future studies to include PASS as an outcome measure and evaluate its changes from baseline when assessing the effectiveness of different OA treatments.

To our knowledge, this is the first study to explore individual variations in the trajectories of the responses to a PASS anchor question among individuals with hip or knee OA. The large sample size and the use of the responses to the direct PASS question (instead of PROM cutoff point estimates) are the main strengths of the present study. Several limitations of the present study should be acknowledged. A single binary question might not be sufficient to adequately capture patients' satisfaction with their symptom state. We have used a multidomain PASS anchor question with no specification of the time spent in the state ("current PASS") in the present study, and the results may not be generalizable to other types of PASS anchor questions. The lack of a control group limits our ability to attribute observed changes in the PASS status and PROMs solely to the digital intervention, as other factors (eg, concurrent additional treatments and contextual factors) could have influenced these outcomes. We relied on self-reported data that are prone to response biases. Voluntary self-selection into the digital intervention limits the generalizability of the findings to similar populations who have access to and are comfortable with digital technologies. Some possibly important predictors of PASS status, such as symptom duration, OA structural severity (eg, the Kellgren-Lawrence grade), smoking status, and country of birth were not available in this dataset.

Our study documented that the percentage of individuals reporting an acceptable symptom state almost tripled, rising from 17% to 49% from before to 1 year after participation in a digital education and exercise therapy program, among >15,000 people with knee or hip OA. Despite large individual variabilities in the PASS trajectories, all trajectories experienced improvements in PROMs over 1 year of follow-up, with larger improvements observed among those with better PASS trajectories. We identified several factors associated with PASS trajectories that may be used to modify or better target the digital treatment to achieve PASS. Our findings highlighted the distinction between "feeling better" and "feeling good" and the need for using measures of both concepts to provide a more comprehensive understanding of OA therapy outcomes.

AUTHOR CONTRIBUTIONS

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

REFERENCES

1. Conley B, Bunzli S, Bullen J, et al. Core recommendations for osteoarthritis care: a systematic review of clinical practice guidelines. Arthritis Care Res (Hoboken) 2023;75(9):1897–1907.
2. Mazzei DR, Whittaker JL, Kania-Richmond A, et al. Do people with knee osteoarthritis use guideline-consistent treatments after an orthopaedic surgeon recommends nonsurgical care? A cross-sectional survey with long-term follow-up. Osteoarthr Cartil Open 2022;4(2):100256.
3. Sasaki R, Honda Y, Oga S, et al. Effect of exercise and/or educational interventions on physical activity and pain in patients with hip/knee osteoarthritis: a systematic review with meta-analysis. PLoS One 2022;17(11):e0275591.
4. Yang Y, Li S, Cai Y, et al. Effectiveness of telehealth-based exercise interventions on pain, physical function and quality of life in patients with knee osteoarthritis: a meta-analysis. J Clin Nurs 2023;32(11-12):2505–2520.
5. Chen T, Or CK, Chen J. Effects of technology-supported exercise programs on the knee pain, physical function, and quality of life of individuals with knee osteoarthritis and/or chronic knee pain: a systematic review and meta-analysis of randomized controlled trials. J Am Med Inform Assoc 2021;28(2):414–423.
6. Tubach F, Ravaud P, Baron G, et al. Evaluation of clinically relevant states in patient reported outcomes in knee and hip osteoarthritis: the patient acceptable symptom state. Ann Rheum Dis 2005;64(1):34–37.
7. Conti F, Ceccarelli F, Massaro L, et al. Evaluation of the patient acceptable symptom state (PASS) in Italian patients affected by systemic lupus erythematosus: association with disease activity indices. PLoS One 2013;8(9):e73517.
8. Katchamart W, Narongroeknawin P, Suppa-Udom B, et al. Factors associated with and cutoff points for patient acceptable symptom state (PASS) in rheumatoid arthritis. Clin Rheumatol 2020;39(3):779–786.
9. Shaffrey I, Nguyen J, Conti M, et al. Defining the patient acceptable symptom state (PASS) for PROMIS after total ankle replacement. J Bone Joint Surg Am 2024;106(15):1404–1410.
10. Escobar A, Gonzalez M, Quintana JM, et al. Patient acceptable symptom state and OMERACT-OARSI set of responder criteria in joint replacement. Identification of cut-off values. Osteoarthritis Cartilage 2012;20(2):87–92.
11. Mahler EAM, Boers N, Bijlsma JWJ, et al. Patient acceptable symptom state in knee osteoarthritis patients succeeds across different patient-reported outcome measures assessing physical function, but fails across other dimensions and rheumatic diseases. J Rheumatol 2018;45(1):122–127.
12. Georgopoulos V, Smith S, McWilliams DF, et al. Harmonising knee pain patient-reported outcomes: a systematic literature review and meta-analysis of patient acceptable symptom state (PASS) and individual participant data (IPD). Osteoarthritis Cartilage 2023;31(1):83–95.
13. Paulsen A, Djuv A, Dalen I. Clinical cut-offs for hip- and knee arthroplasty outcome - minimal clinically important improvement (MCII) and patient acceptable symptom state (PASS) of patient-reported outcome measures (PROM). Qual Life Res 2025;34(4):1147–1158.
14. Martinez-Harms R, Barnett C, Bril V. Time to achieve a patient acceptable symptom state in myasthenia gravis. Front Neurol 2023;14:1187189.
15. Kiadaliri A, Cronström A, Dahlberg LE, et al. Patient acceptable symptom state and treatment failure threshold values for Work Productivity and Activity Impairment and EQ-5D-5L in osteoarthritis. Qual Life Res 2024;33(5):1257–1266.
16. Cronström A, Ingelsrud LH, Nero H, et al. Interpretation threshold values for patient-reported outcomes in patients participating in a digitally delivered first-line treatment program for hip or knee osteoarthritis. Osteoarthr Cartil Open 2023;5(3):100375.
17. Kvien TK, Heiberg T, Hagen KB. Minimal clinically important improvement/difference (MCII/MCID) and patient acceptable symptom state (PASS): what do these concepts mean? Ann Rheum Dis 2007;66(Suppl 3) (suppl 3):iii40–iii41.
18. Boffa A, Andriolo L, Franceschini M, et al. Minimal clinically important difference and patient acceptable symptom state in patients with knee osteoarthritis treated with PRP injection. Orthop J Sports Med 2021;9(10):23259671211026242.
19. Dahlberg LE, Grahn D, Dahlberg JE, et al. A web-based platform for patients with osteoarthritis of the hip and knee: a pilot study. JMIR Res Protoc 2016;5(2):e115.
20. Kiadaliri A, Hörder H, Lohmander LS, et al. Weekly pain trajectories among people with knee or hip osteoarthritis participating in a digitally delivered first-line exercise and education treatment. Pain Med 2024;25(4):291–299.
21. Kiadaliri A, Lohmander LS, Dahlberg LE. Two-thirds maintain high adherence to digital education and exercise therapy with comparable outcomes across adherence clusters: a registry study including data from over 14 000 patients in Sweden. J Orthop Sports Phys Ther 2025;55(1):56–67.
22. Roos EM, Grønne DT, Thorlund JB, et al. Knee and hip osteoarthritis are more alike than different in baseline characteristics and outcomes: a longitudinal study of 32,599 patients participating in supervised education and exercise therapy. Osteoarthritis Cartilage 2022;30(5):681–688.
23. Ingelsrud LH, Granan LP, Terwee CB, et al. Proportion of patients reporting acceptable symptoms or treatment failure and their associated KOOS values at 6 to 24 months after anterior cruciate ligament

reconstruction: a study from the Norwegian Knee Ligament Registry. Am J Sports Med 2015;43(8):1902–1907.

24. Bellamy N, Kirwan J, Boers M, et al. Recommendations for a core set of outcome measures for future phase III clinical trials in knee, hip, and hand osteoarthritis. Consensus development at OMERACT III. J Rheumatol 1997;24(4):799–802.
25. Gandek B, Roos EM, Franklin PD, et al. A 12-item short form of the Knee injury and Osteoarthritis Outcome Score (KOOS-12): tests of reliability, validity and responsiveness. Osteoarthritis Cartilage 2019; 27(5):762–770.
26. Gandek B, Roos EM, Franklin PD, et al. A 12-item short form of the Hip disability and Osteoarthritis Outcome Score (HOOS-12): tests of reliability, validity and responsiveness. Osteoarthritis Cartilage 2019; 27(5):754–761.
27. Sun S, Chuang LH, Sahlén KG, et al. Estimating a social value set for EQ-5D-5L in Sweden. Health Qual Life Outcomes 2022;20(1):167.
28. Dobson F, Hinman RS, Roos EM, et al. OARSI recommended performance-based tests to assess physical function in people diagnosed with hip or knee osteoarthritis. Osteoarthritis Cartilage 2013; 21(8):1042–1052.
29. Kiadaliri A. Psychometric properties of the EQ-5D-5L in patients with knee or hip osteoarthritis: confirmatory factor analysis and item response theory. Osteoarthritis Cartilage 2024;32:S196–S197.
30. Olsson SJ, Ekblom Ö, Andersson E, et al. Categorical answer modes provide superior validity to open answers when asking for level of physical activity: a cross-sectional study. Scand J Public Health 2016;44(1):70–76.
31. Austin PC. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. Stat Med 2009;28(25):3083–3107.
32. Barros AJ, Hirakata VN. Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence ratio. BMC Med Res Methodol 2003;3(1):21.
33. Nagin DS, Jones BL, Elmer J. Recent advances in group-based trajectory modeling for clinical research. Annu Rev Clin Psychol 2024; 20(1):285–305.
34. Haviland AM, Jones BL, Nagin DS. Group-based trajectory modeling extended to account for nonrandom participant attrition. Sociol Methods Res 2011;40(2):367–390.
35. van der Nest G, Lima Passos V, Candel MJJM, et al. An overview of mixture modelling for latent evolutions in longitudinal data: modelling approaches, fit statistics and software. Adv Life Course Res 2020; 43:100323.
36. Dell'Isola A, Nero H, Dahlberg LE, et al. Within-person change in patient-reported outcomes and their association with the wish to undergo joint surgery during a digital first-line intervention for osteoarthritis. Osteoarthritis Cartilage 2023;31(9):1257–1264.
37. Holden MA, Hattle M, Runhaar J, et al; STEER OA Patient Advisory Group; OA Trial Bank Exercise Collaborative. Moderators of the effect of therapeutic exercise for knee and hip osteoarthritis: a systematic review and individual participant data meta-analysis. Lancet Rheumatol 2023;5(7):e386–e400.
38. Kiadaliri A, Lohmander LS, Ignjatovic MM, et al. Digital self-management of hip and knee osteoarthritis and trajectories of work and activity impairments. BMC Musculoskelet Disord 2023;24(1):207.
39. van Middelkoop M, Schiphof D, Hattle M, et al. People with short symptom duration of knee osteoarthritis benefit more from exercise therapy than people with longer symptom duration: an individual participant data meta-analysis from the OA trial bank. Osteoarthritis Cartilage 2024;32(12):1620–1627.
40. Goh GS, Tarabichi S, Baker CM, et al. Should we aim to help patients "feel better" or "feel good" after total hip arthroplasty? Determining factors affecting the achievement of the minimal clinically important difference and patient acceptable symptom state. J Arthroplasty 2023;38(2):293–299.
41. Hadad MJ, Pasqualini I, Klika AK, et al. High area deprivation index is associated with not achieving the patient-acceptable symptom state after TKA. Clin Orthop Relat Res 2024;482(8):1428–1438.
42. Connelly JW, Galea VP, Rojanasopondist P, et al. Which preoperative factors are associated with not attaining acceptable levels of pain and function after TKA? Findings from an International Multicenter Study. Clin Orthop Relat Res 2020;478(5):1019–1028.
43. Perrot S, Bertin P. "Feeling better" or "feeling well" in usual care of hip and knee osteoarthritis pain: determination of cutoff points for patient acceptable symptom state (PASS) and minimal clinically important improvement (MCII) at rest and on movement in a national multicenter cohort study of 2414 patients with painful osteoarthritis. Pain 2013; 154(2):248–256.
44. Goh GS, Baker CM, Tarabichi S, et al. The paradox of patient-reported outcome measures: should we prioritize "feeling better" or "feeling good" after total knee arthroplasty? J Arthroplasty 2022; 37(9):1751–1758.
45. Mostafaee N, Rashidi F, Negahban H, et al. Responsiveness and minimal important changes of the OARSI core set of performance-based measures in patients with knee osteoarthritis following physiotherapy intervention. Physiother Theory Pract 2024;40(5):1028–1039.
46. Wright AA, Cook CE, Baxter GD, et al. A comparison of 3 methodological approaches to defining major clinically important improvement of 4 performance measures in patients with hip osteoarthritis. J Orthop Sports Phys Ther 2011;41(5):319–327.
47. Kiadaliri A, Lohmander S, Dahlberg LE. Meaningful changes for EQ-5D-5L and 30 second chair stand test after digital education and exercise therapy for knee or hip osteoarthritis. Osteoarthritis Cartilage 2025;33:S259–S260.
48. Silva MDC, Perriman DM, Fearon AM, et al. Minimal important change and difference for knee osteoarthritis outcome measurement tools after non-surgical interventions: a systematic review. BMJ Open 2023;13(5):e063026.
49. Tubach F, Dougados M, Falissard B, et al. Feeling good rather than feeling better matters more to patients. Arthritis Rheum 2006;55(4): 526–530.
50. Roos EM, Boyle E, Frobell RB, et al. It is good to feel better, but better to feel good: whether a patient finds treatment 'successful' or not depends on the questions researchers ask. Br J Sports Med 2019; 53(23):1474–1478.

Arthritis Care & Research
Vol. 77, No. 12, December 2025, pp 1436–1442
DOI 10.1002/acr.25601

American College of Rheumatology
Empowering Rheumatology Professionals

Prevalence of Pincer Morphology in Early Adolescents From the General Population: A Population-Based Study

Delong Chen, MD,[1,2] Fleur Boel, MSc,[1,2] Suzanne de Vos-Jakobs, MD,[3] Pim van Klij, MD, PhD,[4,5] Michiel M. A. van Buuren, MD,[2] Sita M. A. Bierma-Zeinstra, PhD,[2] and Rintje Agricola, MD, PhD[2]

Objective. Pincer morphology can lead to femoroacetabular impingement syndrome (FAIS) and may be a modifiable risk factor for hip osteoarthritis (OA). Currently, no studies investigate the prevalence of pincer morphology in early adolescence, which is the period when this bony shape likely develops. The purpose of this study was to estimate the prevalence and birth-assigned sex distribution of pincer morphology in early adolescents from the general population in the Netherlands.

Methods. This study was embedded in the Generation R Study, a population-based prospective cohort in Rotterdam, the Netherlands. Around the age of 13 years, participants underwent high-resolution dual-energy x-ray absorptiometry of their full-body and right hip. The lateral center edge angle (LCEA) was automatically determined based on landmarks outlining the hip contour, and pincer morphology was defined as a LCEA ≥ 40°. The overall and birth-assigned sex-specific prevalence was presented as a percentage with 95% confidence interval (CI).

Results. A total of 3,986 adolescents (median age 13.5 years [2.5th–97.5th percentile: 13.2–14.6]; 46.8% male) were included. The overall prevalence of pincer morphology was 3.1% (95% CI 2.6–3.6). The prevalence in male and female adolescents was 3.0% (95% CI 2.2–3.7) and 3.3% (95% CI 2.5–4.0), respectively.

Conclusion. Among early adolescents from the general population in the Netherlands, the estimated prevalence of pincer morphology was 3.1%. Male and female adolescents had a similar prevalence of pincer morphology. These findings could inform the timing of prevention strategies for pincer morphology and potentially reduce the risk of FAIS and hip OA.

INTRODUCTION

Osteoarthritis (OA) is a major cause of disability worldwide,[1,2] and femoroacetabular impingement syndrome (FAIS) is an important risk factor of hip OA.[3] FAIS is a motion-related disorder that affects the hip joint and results from abnormal contact between the acetabulum and the femoral head.[4] Depending on the anatomic morphology, FAIS can be distinguished based on the presence of cam and/or pincer morphology. Cam morphology involves an extra bone formation at the anterolateral femoral head-neck junction, whereas pincer morphology exhibits either focal or global overcoverage of the femoral head by the acetabulum. Pincer morphology is not only associated with FAIS but also may lead to labral tears and circumferential acetabular cartilage damage.[5,6] Although the relationship between pincer morphology and hip OA is controversial,[7–10] a recent study from the Worldwide Collaboration on OsteoArthritis prediCtion for the Hip (World COACH) consortium reported that hips with pincer morphology have a 1.59 (95% confidence interval [CI] 1.16–2.20) times higher odds of developing radiographic hip OA within eight years.[11]

Until now, the exact age at which pincer morphology starts to develop had not yet been determined. A retrospective study based on abdominal computed tomography scans from 225 patients aged 2 to 19 years without hip complaints reported that pincer morphology first developed at the age of 12 years.[12]

Dr Chen's work was supported by the China Scholarship Council PhD Fellowship for his PhD study at Erasmus Medical Center, Rotterdam, the Netherlands (grant 202108430032) and this publication is part of the Flagship Healthy Joints which is (partly) financed by Convergence Health and Technology.

[1]The Generation R Study Group, Erasmus Medical Center, University Medical Center Rotterdam, Rotterdam, the Netherlands; [2]Department of Orthopaedics and Sports Medicine, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands; [3]Department of Orthopaedics and Sports Medicine, Erasmus MC - Sophia Children's Hospital, University Medical Center Rotterdam, Rotterdam, the Netherlands; [4]Department of Orthopedic Surgery and Sports Medicine, Academic Center for Evidence Based Medicine, Amsterdam IOC Center ACHSS, Amsterdam UMC, Location University of Amsterdam, Amsterdam, the Netherlands; [5]Amsterdam Movement Sciences, Sports, Amsterdam, the Netherlands.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25601.

Address correspondence via email to Rintje Agricola, MD, PhD, at r.agricola@erasmusmc.nl.

Submitted for publication January 7, 2025; accepted in revised form June 11, 2025.

SIGNIFICANCE & INNOVATIONS

- Based on a cross-sectional analysis of 3,986 adolescents aged 13 years in a population-based cohort (Generation R), we found that the overall prevalence of pincer morphology was 3.1% (95% confidence interval [CI] 2.6–3.6), with similar prevalence in male (3.0% [95% CI 2.2–3.7)] and female adolescents (3.3% [95% CI 2.5–4.0)].
- Given the relatively low prevalence in this population of early adolescents, pincer morphology might also develop further during skeletal maturation of the hip like cam morphology.
- Our findings could inform the timing of prevention strategies for pincer morphology, potentially reducing the risk of femoroacetabular impingement syndrome and hip osteoarthritis.

To the best of our knowledge, no large population-based epidemiologic studies have examined the prevalence of pincer morphology in the early adolescent population. Interestingly, there are more studies available on the development of cam morphology, which show that cam morphology is an adaptive response to vigorous hip loading and develops gradually over time during adolescence.[13,14] However, it is unknown whether this concept holds for pincer morphology, as well. Understanding the prevalence of pincer morphology in early adolescence could be critical for informing the timing of prevention strategies that may help slow its progression to hip OA. We therefore aimed to estimate the overall and birth-assigned sex-specific prevalence of pincer morphology in a general population of 3,986 early adolescents.

METHODS

Study design and participants. The data used in this study were derived from the Generation R Study, a population-based prospective cohort aiming to investigate the growth, development, and health of children from fetal life onwards in Rotterdam, the Netherlands. A total of 9,778 pregnant women whose delivery was expected between April 2002 and January 2006 were enrolled in the study, and their children became part of the Generation R cohort. Detailed designs and methods on the Generation R cohort were previously described elsewhere.[15] Around the age of 5 (Focus 5), 9 (Focus 9), and 13 years (Focus 13), all participating children and their parents were invited to visit the Generation R research center in the Erasmus MC Sophia Children's Hospital. For the current study, we included all participants who had both full-body and right hip high-resolution dual-energy x-ray absorptiometry (DXA) scans available at the Focus 13 visit. Exclusion criteria were 1) participants with an incomplete acetabulum depicted on the DXA image; 2) the presence of movement artifacts; and/or 3) the presence of an artifact in the region of interest. This study was approved by the Medical Ethical Committee of Erasmus Medical Center (MEC-2015-749), and written informed consents were obtained from the participants and their parents. We followed the Strengthening the Reporting of Observational Studies in Epidemiology reporting guidelines.[16]

DXA scan. All participants underwent DXA scans of their full-body and right hip by well-trained investigators using the General Electric (GE)-Lunar IntelligentDXA (iDXA) densitometer (GE Healthcare, Madison, WI, USA). The full-body and right hip DXA scans were done sequentially with the participants in the same position. Before scanning, all participants were required to remove heavy clothing, shoes, and any metal accessories. Afterward, they were placed in a supine position with the hands flat at their side. The legs were slightly separated and rotated internally, with the big toes touching. The feet were fastened in this position with a Velcro strip to avoid movement.

Pincer morphology. Pincer morphology was determined by the lateral center edge angle (LCEA) on the DXA images. The LCEA defines the bony coverage of the acetabulum by the femoral head. The presence of pincer morphology was defined as a LCEA $\geq 40°$.[10,12,17] The LCEA was automatically calculated based on landmarks outlining the hip contour with in-house developed software (Figure 1).[18] All landmarks were placed automatically using the BoneFinder software (www.bone-finder.com; The University of Manchester, UK).[19] A visual inspection was conducted to ensure correct placement of landmarks and manual adjustments (DC and FB) were made when necessary. The LCEA was determined using the following steps: first, the center of the femoral head was automatically computed based on the best-fitting circle around the femoral head. Next, the LCEA was formed by the line from the center of the femoral head to the most lateral bony edge of the acetabulum and the line from the center of the femoral head perpendicular to the horizontal reference line of the pelvis as determined on the full-body DXA image (Figure 1).[18] Details of this method have been published previously and show an intermethod reliability between automatic measurement and manual measurement of 0.95 (95% CI 0.87–0.98).[18]

The datasets used in the current study are available to researchers after reasonable request to the management team of the Generation R Study. More detailed information is available on the following website (https://generationr.nl/researchers/collaboration/).

Statistical analyses. We compared the characteristics of included and excluded adolescents using Mann–Whitney U tests for nonnormally distributed continuous data, Student's t-tests for normally distributed continuous data, and chi-square tests for categorical data. We estimated the prevalence of pincer morphology and presented it as a percentage with 95% CI stratified by sex. The 95% CI was computed assuming a binomial distribution. The chi-square test was used to examine the difference in birth-assigned sex-specific prevalence. The overall and birth-

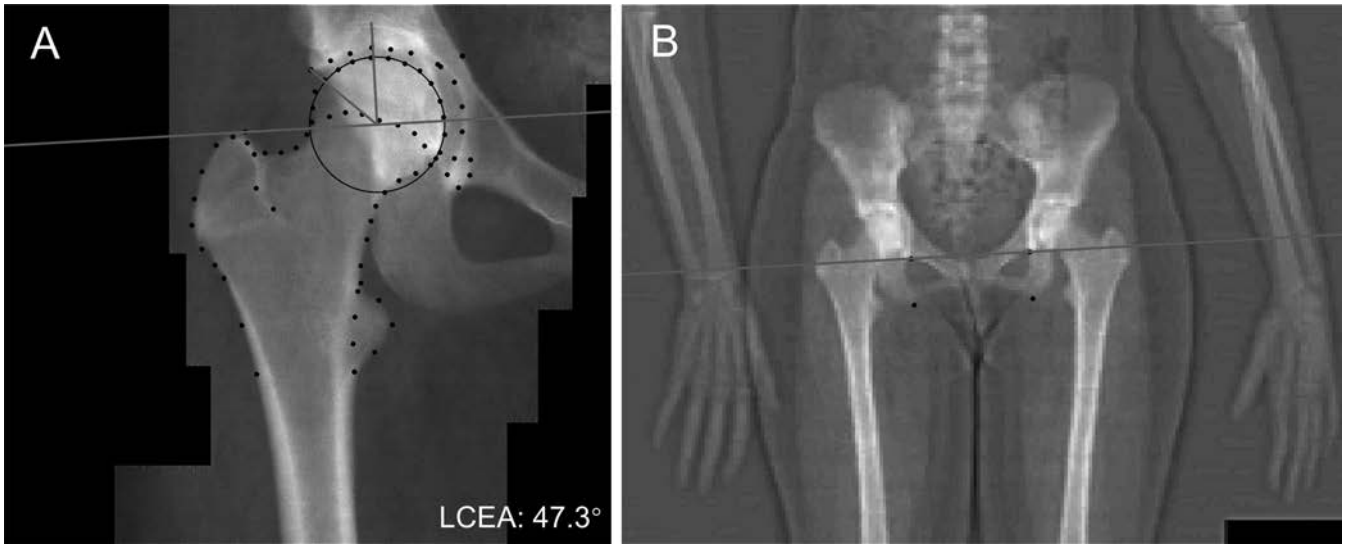

Figure 1. The measurement of LCEA. On the hip DXA image, we outlined the shape of the proximal femur and acetabulum with 80 landmarks. The center of the femoral head was automatically determined by the best-fitting circle, which was based on the landmarks outlining the femoral head. The LCEA of 47.3 degrees in this sample (A) is formed by the intersection of two lines: one from the center of the femoral head to the most lateral bony edge of the acetabulum and the other from the center of the femoral head perpendicular to the horizontal reference line of the pelvis as determined on the full-body DXA image (B). DXA, dual-energy x-ray absorptiometry; LCEA, lateral center edge angle.

assigned sex-specific distribution of LCEA was assessed using histograms. To address potential selection bias on our findings, we calculated the weighted prevalence using inverse probability weighting (IPW) adjusting for birth-assigned sex, body mass index (BMI), and height.[20] We performed statistical analyses using R Statistical software (v4.2.1; R Core Team 2022). The 95% CIs were computed using the binom-package,[21] histograms were generated using the ggplot2-package,[22] and IPW was computed using the ipw-package.[23] A P value < 0.05 was considered as statistically significant.

RESULTS

Characteristics of study participants. A total of 4,929 participants attended the Focus 13 visit, of whom 3,986

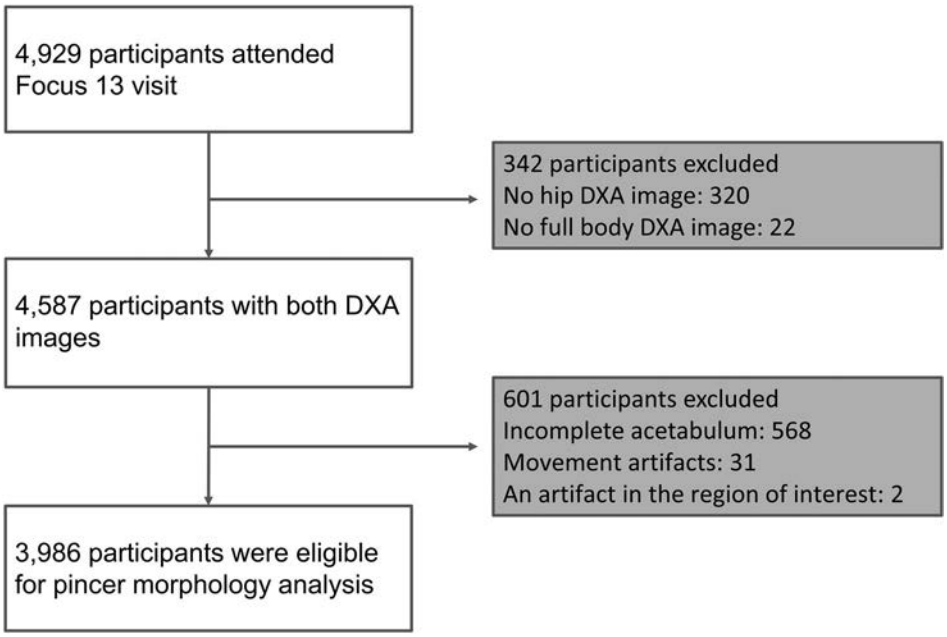

Figure 2. Flowchart of participants. DXA, dual-energy x-ray absorptiometry.

Table 1. Demographic characteristics of included and excluded adolescents*

Characteristic	Included adolescents (N = 3,986)	Excluded adolescents (N = 943)	*P* value
Age, median (2.5th–97.5th percentile) y	13.5 (13.2–14.6)	13.6 (13.2–14.7)	0.567
Male, n (%)	1,864 (46.8)	556 (59.0)	<0.001
Female, n (%)	2,122 (53.2)	387 (41.0)	
BMI, median (2.5th–97.5th percentile)	19.2 (15.2–29.3)	18.9 (15.1–28.3)	0.003
Height, mean ± SD, cm	164.0 (7.9)	165.9 (8.2)	<0.001

* BMI, body mass index.

participants (81%) had sufficient quality right hip and full-body DXA images available and were analyzed in this study (Figure 2). Table 1 shows the characteristics of the included and excluded early adolescents. The median age of included adolescents was 13.5 years (2.5th–97.5th percentile: 13.2–14.6), and 1,864 (46.8 %) were male adolescents. The median BMI of included adolescents was 19.2 (2.5th–97.5th percentile: 15.2–29.3), and the mean height was 164.0 cm (SD 7.9). The excluded adolescents were more frequently male, slightly taller, and had slightly lower BMI than the included adolescents (Table 1).

Prevalence of pincer morphology. The overall and birth-assigned sex-specific LCEA distributions are presented in Figure 3. The overall prevalence of pincer morphology was 3.1% (95% CI 2.6–3.6) in early adolescents. In the sex subgroup analysis, the prevalence of pincer morphology was similar in male (3.0% [95% CI 2.2–3.7]) and female participants (3.3% [95% CI 2.5–4.0]) (Table 2). The weighted prevalence was similar to the unweighted prevalence in the total group and birth-assigned sex subgroup, as shown in Table 2.

DISCUSSION

This study included 3,986 early adolescents from the general population of Rotterdam, the Netherlands, and used high-resolution DXA scans to determine the prevalence of pincer morphology. We found that the overall prevalence of pincer morphology was 3.1% (95% CI 2.6–3.6) and that it was similar in male (3.0% [95% CI 2.2–3.7]) and female participants (3.3% [95% CI 2.5–4.0]).

The previously reported prevalence of pincer morphology varies widely in literature because of different study populations and inconsistent definitions. To date, there is a paucity of literature on the prevalence of pincer morphology in early adolescents. A population-based study of 2,081 participants (mean age, 18.6 years) in Norway found that 24% of them had a pincer morphology as defined by the presence of a posterior wall sign,

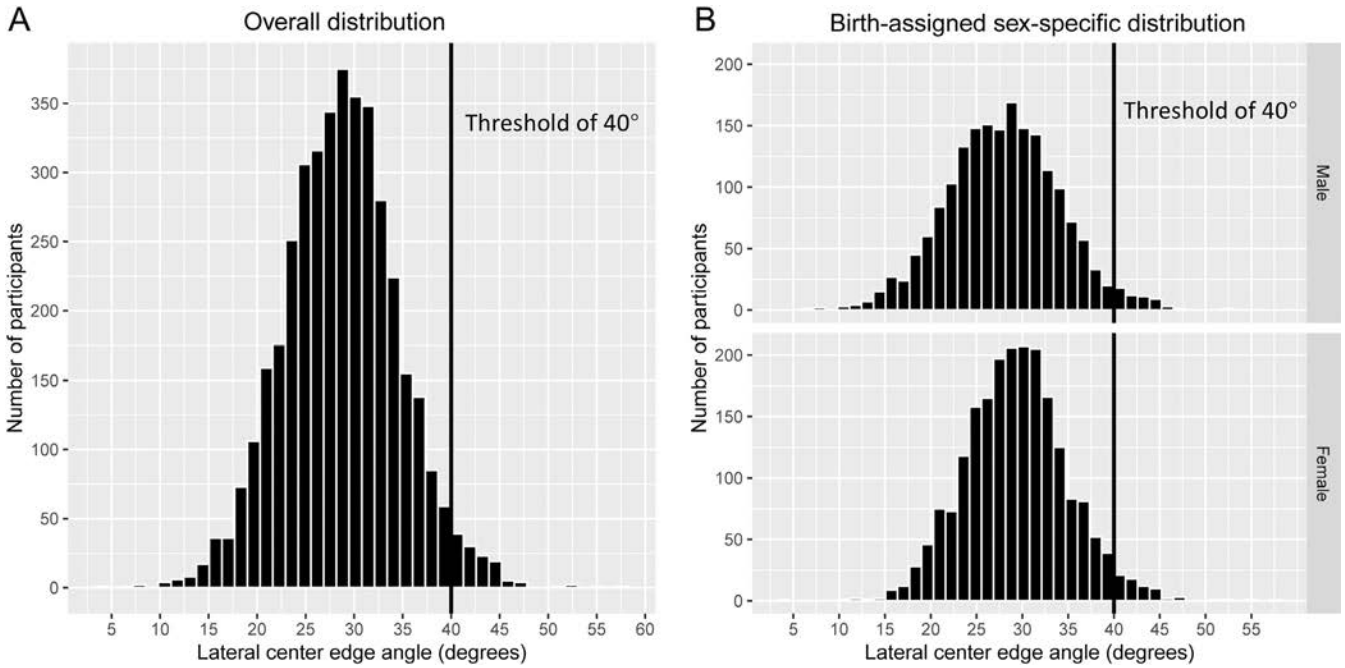

Figure 3. The overall (A) and birth-assigned sex-specific (B) distribution of the lateral center edge angle among the included early adolescents. The black solid line indicates the 40° threshold of the lateral center edge angle.

Table 2. The prevalence of pincer morphology among early adolescents stratified by birth-assigned sex*

Characteristic	Cases, n/N	LCEA ≥40° Prevalence, % (95% CI)	χ^2	*P* value
Unweighted prevalence				
Overall	124/3,986	3.1 (2.6–3.6)		
Sex				
Male	55/1,864	3.0 (2.2–3.7)	0.298	0.585
Female	69/2,122	3.3 (2.5–4.0)		
Weighted prevalence				
Overall	–	3.1 (2.6–3.7)		
Sex				
Male	–	3.0 (2.2–3.8)	0.203	0.653
Female	–	3.3 (2.5–4.0)		

* CI, confidence interval; LCEA, lateral center edge angle.

crossover sign, or excessive acetabular coverage on the anteroposterior (AP) pelvic radiograph.[24] In a cross-sectional study of 6,807 individuals with a mean age of 62.7 years from the UK Biobank, pincer morphology was defined as a LCEA ≥ 45° and found in 8.5% of participants using DXA scans of the left hip.[25] Our results showed the prevalence of pincer morphology was 3.1% in early adolescents, lower than most of the previously reported prevalences in adults. Several factors could explain this result. First, the coverage of the acetabulum may still increase during skeletal maturation, which could lead to the development of pincer morphology. Some prospective studies have investigated the development of cam morphology and revealed that it gradually increases in size during adolescence.[13,14] Interestingly, the low prevalence of pincer morphology in our study suggests that it might also gradually develop during skeletal growth similar to cam morphology. Prospective studies are needed to confirm this observation. Secondly, we only used the LCEA, one of the most commonly used objective measures, to quantify pincer morphology. Other definitions, such as the crossover sign, posterior wall sign, and coxa profunda, have also been reported to determine pincer morphology. However, previous studies on these measures reported poor reliability and specificity.[26–28] The definition of pincer morphology lacks a validated measure and accompanying thresholds, such as for the LCEA. In adults, the LCEA threshold to define pincer morphology varies across studies. Some large prospective studies used a threshold of 33.7° or 40°[10,29] to investigate the association of pincer morphology and hip OA, whereas other studies used a threshold of 45°.[11,25] The recent Lisbon Agreement provided guidance and criteria for defining pincer morphology, including global overcoverage identified by the presence of protrusio acetabuli or a Wiberg center edge angle (WCEA) ≥40° (or ≥35° with an acetabular index <0°).[30] We used the same threshold value but chose to use the LCEA instead of the WCEA in this study because pincer impingement usually occurs between the femoral head-neck junction and the most lateral bony edge of the acetabulum.

Research on sex differences in the prevalence of pincer morphology has yielded inconclusive results. A population-based study of 2,596 participants (mean age, 63 years) in the United States found that 7% of male participants and 10% of female participants had pincer morphology as defined by a LCEA ≥ 40° or acetabular protrusio in the AP pelvic radiograph.[31] Faber et al included 6,807 participants (mean age 62.7 years) from UK Biobank in a cross-sectional study in which pincer morphology was defined as LCEA ≥45° and present in 8.9% of male participants and 8.1% of female participants on the DXA images of left hips.[25] Our results are in line with these findings and also indicate the similar prevalence of pincer morphology between male and female adolescents at a mean age of 13 years. However, results from other studies are contrary to our findings. In a Norwegian population-based cohort study, Laborie et al studied the AP pelvic radiograph of 2,081 participants aged 17.2 to 20 years. They found that 34.3% of male participants had at least one sign of pincer morphology (posterior wall sign, crossover sign, or excessive acetabular coverage) compared with 16.6% of female participants.[24] The reason for these inconsistent results, which complicate direct comparison, are the heterogeneous populations, diverse imaging modalities (radiographs vs DXA scans), and variety in definitions for pincer morphology (LCEA thresholds or specific radiographic signs). Although our study found the similar prevalence of pincer morphology between sexes in early adolescence, follow-up of the current cohort can reveal if these sex-specific differences become apparent during or after skeletal maturation.

Selection bias is a critical consideration in this study, as 19% (943 of 4,929) of adolescents who visited the Generation R research center were excluded because of missing DXA scans or LCEA measurements. To address this selection bias, we applied IPW using a logistic regression model adjusting for sex, BMI, and height—factors that showed statistically significant differences between included and excluded participants (Table 1). The similarity between weighted and unweighted prevalence estimates suggests that the selection bias had a limited impact on our findings. However, generalizability to the broader Dutch population requires caution because our study did not weight results to the entire Dutch adolescent population.

The major strength of this study is the population-based design with a large sample size in an early adolescent population. Moreover, automated measurement is fast and highly reproducible for LCEA calculation, which helps eliminate observer bias that is inherently present with manual measurement. Our study has some limitations that should be addressed. First, our study only included the right hips for analyses, so the prevalence of pincer morphology was based on hip level rather than person level. Moreover, analyzing only in right hips underestimates the prevalence of pincer morphology at the person level because participants may have unilateral pincer morphology in their left hips. Second, DXA is a less common imaging modality for determining

pincer morphology. However, it provides sufficient resolution to identify hip morphologies with less radiation burden than radiographs and has been validated against AP pelvic radiographs in adults (intraclass correlation coefficient for LCEA: 0.93 [95% CI 0.91–0.94]).[32] Unlike three-dimensional imaging modalities or additional lateral radiographs, the anterior center edge angle could not be obtained from the AP DXA hip images, so this may lead to an underestimation of the prevalence of pincer morphology.[33] Third, care should be taken when generalizing our results to late adolescents because of the potential for further development of pincer morphology when the study population becomes older.

In conclusion, our study provides a large-scale objective evaluation of pincer morphology among early adolescents from the general population in the Netherlands. Pincer morphology was present in 3.1% of early adolescents and is similarly prevalent in male and female participants. Our study provides valuable data for this age group and serves as a reference for investigating the development of pincer morphology in men and women throughout adolescence. Given its higher prevalence in adults, pincer morphology might also develop further during skeletal maturation of the hip like cam morphology. Our findings could inform the timing of prevention strategies for pincer morphology, potentially reducing the risk of FAIS and hip OA.

ACKNOWLEDGMENTS

The Generation R Study is conducted by the Erasmus MC, University Medical Center Rotterdam in close collaboration with the Erasmus University Rotterdam and the city of Rotterdam. We gratefully acknowledge the contribution of children and parents. The general design of Generation R Study is made possible by long-term financial support from Erasmus MC, University Medical Center Rotterdam, the Netherlands, Organization for Health Research and Development and the Dutch Ministry of Health, Welfare and Sport.

AUTHOR CONTRIBUTIONS

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

REFERENCES

1. Collaborators GBDO; GBD 2021 Osteoarthritis Collaborators. Global, regional, and national burden of osteoarthritis, 1990 to 2020 and projections to 2050: a systematic analysis for the Global Burden of Disease Study 2021. Lancet Rheumatol 2023;5(9):e508–e522.
2. Glyn-Jones S, Palmer AJ, Agricola R, et al. Osteoarthritis. Lancet 2015;386(9991):376–387.
3. Agricola R, van Buuren MMA, Kemp JL, et al. Femoroacetabular impingement syndrome in middle-aged individuals is strongly associated with the development of hip osteoarthritis within 10-year follow-up: a prospective cohort study (CHECK). Br J Sports Med 2024; 58(18):1061–1067.
4. Griffin DR, Dickenson EJ, O'Donnell J, et al. The Warwick Agreement on femoroacetabular impingement syndrome (FAI syndrome): an international consensus statement. Br J Sports Med 2016;50(19): 1169–1176.
5. Beck M, Kalhor M, Leunig M, et al. Hip morphology influences the pattern of damage to the acetabular cartilage: femoroacetabular impingement as a cause of early osteoarthritis of the hip. J Bone Joint Surg Br 2005;87(7):1012–1018.
6. Pascual-Garrido C, Li DJ, Grammatopoulos G, et al; ANCHOR Group. The pattern of acetabular cartilage wear is hip morphology-dependent and patient demographic-dependent. Clin Orthop Relat Res 2019; 477(5):1021–1033.
7. Nelson AE, Stiller JL, Shi XA, et al. Measures of hip morphology are related to development of worsening radiographic hip osteoarthritis over 6 to 13 year follow-up: the Johnston County Osteoarthritis Project. Osteoarthritis Cartilage 2016;24(3):443–450.
8. Nicholls AS, Kiran A, Pollard TC, et al. The association between hip morphology parameters and nineteen-year risk of end-stage osteoarthritis of the hip: a nested case-control study. Arthritis Rheum 2011; 63(11):3392–3400.
9. Saberi Hosnijeh F, Zuiderwijk ME, Versteeg M, et al. Cam deformity and acetabular dysplasia as risk factors for hip osteoarthritis. Arthritis Rheumatol 2017;69(1):86–93.
10. Agricola R, Heijboer MP, Roze RH, et al. Pincer deformity does not lead to osteoarthritis of the hip whereas acetabular dysplasia does: acetabular coverage and development of osteoarthritis in a nationwide prospective cohort study (CHECK). Osteoarthritis Cartilage 2013;21(10):1514–1521.
11. Riedstra NS, Van Buuren MM, Boel F, et al. Pincer morphology is a risk factor for developing radiographic hip osteoarthritis; data from The World Coach Consortium. Osteoarthritis Cartilage 2023;31: S249–S250.
12. Monazzam S, Bomar JD, Dwek JR, et al. Development and prevalence of femoroacetabular impingement-associated morphology in a paediatric and adolescent population: a CT study of 225 patients. Bone Joint J 2013;95-B(5):598–604.
13. Agricola R, Heijboer MP, Ginai AZ, et al. A cam deformity is gradually acquired during skeletal maturation in adolescent and young male soccer players: a prospective study with minimum 2-year follow-up. Am J Sports Med 2014;42(4):798–806.
14. van Klij P, Heijboer MP, Ginai AZ, et al. Cam morphology in young male football players mostly develops before proximal femoral growth plate closure: a prospective study with 5-yearfollow-up. Br J Sports Med 2019;53(9):532–538.
15. Kooijman MN, Kruithof CJ, van Duijn CM, et al. The Generation R Study: design and cohort update 2017. Eur J Epidemiol 2016; 31(12):1243–1264.
16. von Elm E, Altman DG, Egger M, et al; STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Ann Intern Med 2007;147(8):573–577.
17. Riedstra NS, Boel F, van Buuren M, et al. Pincer morphology is not associated with hip osteoarthritis unless hip pain is present: follow-up data from a prospective cohort study. Arthritis Care Res (Hoboken) 2024;76(5):644–651.
18. Boel F, de Vos-Jakobs S, Riedstra NS, et al. Automated radiographic hip morphology measurements: an open-access method. Osteoarthr Imaging 2024;4(2):100181.

19. Lindner C, Thiagarajah S, Wilkinson JM, et al; arcOGEN Consortium. Fully automatic segmentation of the proximal femur using random forest regression voting. IEEE Trans Med Imaging 2013;32(8):1462–1472.
20. Sayon-Orea C, Moreno-Iribas C, Delfrade J, et al. Inverse probability weighting and multiple imputation for evaluating selection bias in the estimation of childhood obesity prevalence using data from electronic health records. BMC Med Inform Decis Mak 2020;20(1):9.
21. Dorai-Raj S. binom: Binomial Confidence Intervals for Several Parameterizations. 2022.
22. Wickham H. ggplot2: Elegant Graphics for Data Analysis. New York, NY: Springer-Verlag; 2016.
23. van der Wal WM, Geskus RB. {ipw}: An {R} package for inverse probability weighting. J Stat Softw 2011;43(13):1–23.
24. Laborie LB, Lehmann TG, Engesæter IO, et al. Prevalence of radiographic findings thought to be associated with femoroacetabular impingement in a population-based cohort of 2081 healthy young adults. Radiology 2011;260(2):494–502.
25. Faber BG, Ebsim R, Saunders FR, et al. Cam morphology but neither acetabular dysplasia nor pincer morphology is associated with osteophytosis throughout the hip: findings from a cross-sectional study in UK Biobank. Osteoarthritis Cartilage 2021;29(11):1521–1529.
26. Larson CM, Moreau-Gaudry A, Kelly BT, et al. Are normal hips being labeled as pathologic? A CT-based method for defining normal acetabular coverage. Clin Orthop Relat Res 2015;473(4):1247–1254.
27. Zaltz I, Kelly BT, Hetsroni I, et al. The crossover sign overestimates acetabular retroversion. Clin Orthop Relat Res 2013;471(8):2463–2470.
28. Nepple JJ, Lehmann CL, Ross JR, et al. Coxa profunda is not a useful radiographic parameter for diagnosing pincer-type femoroacetabular impingement. J Bone Joint Surg Am 2013;95(5):417–423.
29. Thomas GE, Palmer AJ, Batra RN, et al. Subclinical deformities of the hip are significant predictors of radiographic osteoarthritis and joint replacement in women. A 20 year longitudinal cohort study. Osteoarthritis Cartilage 2014;22(10):1504–1510.
30. Mascarenhas VV, Castro MO, Afonso PD, et al. The Lisbon Agreement on femoroacetabular impingement imaging-part 2: general issues, parameters, and reporting. Eur Radiol 2021;31(7):4634–4651.
31. Raveendran R, Stiller JL, Alvarez C, et al. Population-based prevalence of multiple radiographically-defined hip morphologies: the Johnston County Osteoarthritis Project. Osteoarthritis Cartilage 2018;26(1):54–61.
32. Boel F, Wortel J, van Buuren MMA, et al. DXA images vs. pelvic radiographs: reliability of hip morphology measurements. Osteoarthritis Cartilage 2025;33(2):283–292.
33. Herfkens J, van Buuren MMA, Riedstra NS, et al. Adding false-profile radiographs improves detection of developmental dysplasia of the hip, data from the CHECK cohort. J Hip Preserv Surg 2022;9(1):3–9.

Arthritis Care & Research
Vol. 77, No. 12, December 2025, pp 1443–1452
DOI 10.1002/acr.25581

American College of Rheumatology
Empowering Rheumatology Professionals

Effect of Nonsteroidal Anti-Inflammatory Drugs on Sacroiliac Joint Inflammation, as Seen on Magnetic Resonance Imaging, in Axial Spondyloarthritis

Gareth T. Jones,[1] Alexander N. Bennett,[2] Raj Sengupta,[3] Pedro M. Machado,[4] Helena Marzo-Ortega,[5] Lorna Aucott,[1] Margaret A. Hall-Craggs,[6] Timothy J. P. Bray,[6] Alan Bainbridge,[7] Ruaridh M. Gollifer,[6] and Gary J. Macfarlane[1]

Objective. Imaging evidence of active sacroiliitis is important for diagnosis, classification, and monitoring of axial spondyloarthritis (axSpA). However, there is no consistent guidance on whether patients should temporarily stop non-steroidal anti-inflammatory drugs (NSAIDs) before magnetic resonance imaging (MRI). The aim of this study was to determine whether NSAIDs lead to an underestimation of active sacroiliitis, as observed using MRI.

Methods. Adults with axSpA were recruited from rheumatology clinics and undertook NSAID washout for one to two weeks before a sacroiliac joint MRI scan. Images were read by two independent readers and adjudicated by a third if required. Those who had a positive result for active sacroiliitis, as per internationally recognized criteria, underwent a second scan six weeks after recommencing daily NSAIDs. We determined the proportion of participants who had a negative scanning result while taking NSAIDs after a previous positive result when NSAID-free. Images were also scored using semiquantitative methods comprising lesion size and intensity, and a subset of participants underwent quantitative MRI (qMRI) to provide an objective evaluation of any inflammatory changes.

Results. From 34 centers across the United Kingdom, 311 participants (median age 42 years; 62% male) were recruited; 286 (92%) completed the NSAID washout and underwent the first MRI scan. From 146 participants with active sacroiliitis, follow-up scans (while taking NSAIDs) were obtained from 124 (85%), at which point 25 participants had a negative result (20.2%; 95% confidence interval 13.5%–28.3%). Semiquantitative and qMRI methods supported these findings.

Conclusion. One-fifth of patients showed full resolution of active sacroiliitis lesions when NSAIDs were present. In clinical practice, if patients with axSpA are willing to attempt a one- to two-week NSAID washout before MRI, this should be considered.

INTRODUCTION

Axial spondyloarthritis (axSpA) is a chronic inflammatory disease predominantly affecting the spine and sacroiliac joints. Two disease phenotypes, radiographic axSpA (r-axSpA) and non-radiographic axSpA, are identified based on the presence or absence of structural changes of sacroiliitis on plain radiography. Disease classification has evolved over recent decades, and the

Presented in part at the Annual Scientific Meeting of the British Society for Rheumatology, Manchester (UK), 2023.

Versus Arthritis had no role in study design, other than as part of peer review during the grant application process. Thereafter, they had no role in data collection, data analysis, data interpretation, or writing of the report. The views expressed herein are those of the authors and do not necessarily represent those of the NHS, the NIHR, or the UK Department of Health.

Supported by Versus Arthritis (award 21022). Drs Machado, Hall-Craggs, Bray, and Gollifer's work was supported by the NIHR University College London Hospitals Biomedical Research Centre. Dr Marzo-Ortega's work was supported by the NIHR Leeds Biomedical Research Centre.

[1]Gareth T. Jones, PhD, Lorna Aucott, PhD, Gary J. Macfarlane, MD (Hons): University of Aberdeen, Aberdeen, United Kingdom; [2]Alexander N. Bennett, PhD: Defence Medical Rehabilitation Centre, Stanford Hall, United Kingdom, and Imperial College London, London, United Kingdom; [3]Raj Sengupta, MBBS: Royal National Hospital for Rheumatic Diseases and University of Bath, Bath, United Kingdom; [4]Pedro M. Machado, PhD: University College London, University College London Hospitals NHS Foundation Trust, and London North West University Healthcare NHS Trust, London, United Kingdom; [5]Helena Marzo-Ortega, PhD: Leeds Teaching Hospitals NHS Trust and University of Leeds, Leeds, United Kingdom; [6]Margaret A. Hall-Craggs, PhD, Timothy J. P. Bray, PhD, Ruaridh M. Gollifer, PhD: University College London and University College London Hospitals NHS Foundation Trust, London, United Kingdom; [7]Alan Bainbridge, PhD: University College London Hospitals NHS Foundation Trust, London, United Kingdom.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25581).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25581.

Address correspondence via email to Gareth T. Jones, PhD, at gareth.jones@abdn.ac.uk.

Submitted for publication November 6, 2024; accepted in revised form May 20, 2025.

SIGNIFICANCE & INNOVATIONS

- In axial spondyloarthritis, there are no internationally recognized guidelines for whether patients should discontinue anti-inflammatory medication before sacroiliac joint magnetic resonance imaging (MRI). Practice differs, and evidence is lacking.
- This study demonstrates that a sizable proportion of patients exhibit significant improvements in sacroiliitis in the presence of nonsteroidal anti-inflammatory drugs (NSAIDs). Improvements in lesion severity were observed in >40% of patients, with one in five showing full resolution.
- Despite concerns about increases in pain during an NSAID washout period, this study also provides evidence that a brief NSAID washout can be tolerated by almost all patients who attempt it.
- Future imaging guidelines should incorporate this evidence and recommend that if a patient is willing to attempt NSAID washout before MRI, they should be encouraged and supported to do this, including provision of other non–anti-inflammatory medication if required.

now widely accepted Assessment of SpondyloArthritis international Society (ASAS) classification criteria allow for magnetic resonance imaging (MRI) and various clinical features to phenotype patients further.[1]

MRI is a standard imaging modality in axSpA, and a key factor aiding diagnosis and classification is the presence of bone marrow edema (BME) lesions suggestive of active sacroiliitis. Indeed, in the United Kingdom, the National Institute for Health and Care Excellence guidance for the diagnosis and management of axSpA recommends that diagnosis can be supported using MRI,[2] and acquisition and interpretation of MRI findings is facilitated by a joint consensus from radiologists and rheumatologists under the auspices of the British Society for Spondyloarthritis.[3]

Nonsteroidal anti-inflammatory drugs (NSAIDs) are a common first-line therapy in axSpA, with continuous treatment recommended for persistent symptoms.[4] Because NSAIDs are readily available without prescription, many patients are, or have been, exposed to these drugs when they first present to rheumatology. A common clinical consideration during axSpA diagnostic workup is whether NSAIDs should be withdrawn at the time of imaging; it is plausible that their presence could limit the ability to observe inflammation. Currently, no guidelines exist for whether NSAIDs should be discontinued before MRI, and evidence is still lacking. A small UK study examined patients who met the modified New York criteria for ankylosing spondylitis[5] and who were eligible for tumor necrosis factor (TNF) inhibition.[6] Among nine patients with follow-up MRI, 8 of 22 inflammatory sacroiliac joints lesions had resolved following six weeks of NSAIDs (etoricoxib), although the study was too small to be confident about any conclusions. Another small study of 20 Belgian patients with newly diagnosed axSpA found that none had a normal sacroiliac joint MRI finding after six weeks of optimal-dose NSAIDs, although the intensity of lesions was decreased.[7] Yet real-world evidence shows that NSAIDs can lead to near resolution of severe BME lesions in presumed reactive sacroiliitis over an eight-week period.[8]

We hypothesized that among patients with axSpA, the use of NSAIDs would mask underlying active sacroiliitis when observed using MRI. Not only is this important in terms of diagnosis, but it also has potential implications for treatment (correct diagnosis is required to ensure appropriate management) and research (to determine eligibility for some clinical trials).

The primary objective of the current study was to determine the proportion and characteristics of patients with axSpA with active sacroiliitis on MRI in the absence of NSAIDs, which resolved after NSAIDs were reintroduced. Although the main objective considered sacroiliitis as present or absent, in reality, it is a continuum, and it is possible, therefore, that NSAID use decreases the appearance of active sacroiliitis but insufficiently to alter classification. Thus, the secondary objective was to determine the extent that NSAIDs affected the severity of active sacroiliitis using a semiquantitative scoring method. Furthermore, a subset of patients underwent quantitative MRI (qMRI) to provide an objective evaluation of any changes in inflammation with NSAIDs.

PATIENTS AND METHODS

The DyNAMISM study (Do Non-steroidal Anti-inflammatory Drugs Mask Inflammation in Spondyloarthritis on MRI?) was a non-randomized crossover study that recruited participants from NHS rheumatology clinics across the United Kingdom. Eligible participants were adults with established axSpA or suspected axSpA in whom a sacroiliac joint MRI scan was clinically indicated. In addition, they needed to be taking, or be about to commence, daily NSAIDs. To optimize the chances of observing active BME lesions, participants were required to have had a previous positive MRI result demonstrating active sacroiliitis and/or be male and HLA–B27 positive.[9] Potential participants were excluded if they had received biologic therapy within the past six months, were currently receiving other anti-inflammatory medication (eg, intramuscular or intravenous glucocorticoids), had any NSAID contraindications (eg, active peptic ulceration or severe hepatic or renal dysfunction) or contraindications to MRI (eg, having a metallic or conducting foreign body or severe claustrophobia), or were deemed potentially unable to understand the consent process.

Potential participants were invited to attend a baseline screening visit to confirm their eligibility. Clinical data were extracted from participants' medical notes, including whether they fulfilled the modified New York classification criteria for ankylosing spondylitis[5] and/or ASAS classification for axSpA[1]; presence or absence of spondyloarthritis features and

extramusculoskeletal manifestations; and targeted medical history, including C-reactive protein (CRP) level, year of symptom onset, and year first seen by a rheumatologist (a proxy for year of diagnosis). Participants were asked to complete questionnaires such as the Bath Ankylosing Spondylitis Disease Activity Index (BASDAI),[10] which, along with patient global disease activity and CRP, allowed computation of the Ankylosing Spondylitis Disease Activity Score (ASDAS).[11] Other measures included the Bath Ankylosing Spondylitis Patient Global Score,[12] the Bath Ankylosing Spondylitis Functional Index,[13] general health (Patient-Reported Outcomes Measurement Information System Scale v1.2 Global Health),[14,15] disease-specific quality of life measures (Ankylosing Spondylitis Quality of Life questionnaire),[16] and the modified American College of Rheumatology preliminary diagnostic criteria for fibromyalgia.[17]

Participants were asked to discontinue all NSAID medication(s) for one week before undergoing a sacroiliac joint MRI scan (scan 1) using a standardized protocol including coronal-oblique T1-weighted and STIR sequences with 3-mm slices.[18] A one-week washout period (equivalent to more than five half-lives of the most commonly used NSAIDs) was thought to be sufficient based on drug pharmacokinetics (ie, unlike bisphosphonates, NSAIDs do not specifically target bone tissue). However, after approximately 20% of participants had been recruited, a protocol amendment was implemented to increase the washout period to two weeks because fewer participants than expected had a positive result at scan 1 and there was concern that the one-week washout was not long enough.

All scans were anonymized and randomly allocated to two experienced readers (ANB, PMM, or HM-O). Agreement between readers was examined using the kappa statistic, but in the event of any disagreement, a third reader adjudicated the result. Magnetic resonance images were accessed via a dedicated electronic platform created by the study coordinating center. Scans had a unique study identifier, and thus readers were blinded to participants' clinical characteristics and scan date. On the first reading, active sacroiliitis was categorized as present (positive scan) or absent (negative scan) as per internationally recognized ASAS definitions, revised by Lambert et al.[19] The amount of inflammatory signal required to define a positive scan result followed the specifications given by Rudwaleit et al[20]: if there was a single lesion suggesting active inflammation, it must be present on at least two consecutive MRI slices; if there was more than one lesion, then one slice may be sufficient.

Those with a positive baseline MRI finding (scan 1) were invited for a second MRI scan (scan 2) six weeks after restarting daily NSAIDs. By default, participants restarted whatever drug and dose they were on previously; changes were permitted but were not dictated by study protocol. For each participant, both scans were performed using the same MRI scanner and were again read independently by two readers, with adjudication if necessary. After the initial six-week period (during which, by definition, all scans had to be scan 1), readers were also blinded to time point.

After study completion, all positive scans for scan 1 as well as all scans (positive or negative) for scan 2 were assessed for the degree of active sacroiliitis using the Leeds MRI scoring method, a semiquantitative approach that scores each joint quadrant using a combination of lesion size and intensity from 0 (no active sacroiliitis) to 3 (severe sacroiliitis), giving a theoretical range of 0 to 24, and where a BME lesion grading of ≥2 is considered clinically significant.[21,22] This secondary outcome was scored by one single reader, blind to scan chronology, and after first establishing that agreement between readers was excellent: the observed pairwise mean differences in scores were 0.15 (SD 2.33), −0.77 (SD 2.66), and −0.92 (SD 2.02) on the 0 to 24 scale (see Supplementary Figure 1).

Chemical shift-encoded MRI (also known as Dixon MRI) was performed in a subset of patients to allow quantitative assessment of changes in fat content in the bone marrow, as fat content is known to reduce in the presence of BME.[23,24] Fat fraction (FF) maps were generated as described by Bainbridge et al[25] and detailed in Supplementary Table 1. If specialist Dixon proton density FF packages such as mDixon Quant (Philips), Dixon FQ in the Liver Lab package (Siemens), or IDEAL IQ (GE) were available, these were used for the study. Alternatively, sites used a base-level Dixon option such as mDixon (Philips), DIXON (Siemens), or LAVA FLEX (GE). The FF maps were analyzed using the semiautomated BEACH software tool[24] by two readers (independent of the first two readers and adjudicator, with no access to their scores) who had been trained in the use of the tool. The BEACH tool was used to semiautomatically generate standardized regions of interest in the subchondral marrow of the sacroiliac joints; these regions were then analyzed to generate a set of histographic parameters for each patient at the 10th, 25th, 50th, 75th, and 90th percentiles (denoted FF_{10}, FF_{25}, FF_{median}, FF_{75}, and FF_{90}, respectively). Over the pairs of scans, positive changes in FF parameters were taken to indicate improving inflammation and/or increasing fat metaplasia due to NSAIDs, whereas negative changes were taken to indicate worsening inflammation. As described previously,[24] FF parameters targeting the lower end of the distribution (FF_{10} and FF_{25}) may be more sensitive to edema or active inflammation, whereas those targeting the upper end of the distribution (FF_{75} and FF_{90}) may be more sensitive to fat metaplasia, meaning that the contributions of these processes can be distinguished.

The primary analysis would determine the proportion of participants who changed from positive MRI results at scan 1 (not taking NSAIDs) to negative MRI results at scan 2 (taking NSAIDs). Jarrett et al reported MRI findings on 11 patients with r-axSpA with inflammation in the sacroiliac joints.[6] Of the nine with follow-up MRI, three (33%) showed complete resolution with NSAIDs. The current study was powered on the more conservative estimate that 20% of participants would change from a positive MRI

finding to a negative MRI finding after NSAIDs were reintroduced. Thus, 246 patients would be required to obtain a 95% confidence interval (CI) of ±5% around a prevalence estimate of 20%.

The proportion of participants who changed from positive to negative MRI findings was estimated along with an exact 95% CI. Factors associated with resolution of active sacroiliitis in scan 2 were examined using Poisson regression. Results are presented as risk ratios, with 95% CIs computed with robust SEs.

Leeds scoring was summarized using simple descriptive statistics, and agreement between MRI readers was examined using pairwise Bland–Altman difference plots.[26] The change in qMRI parameters with NSAIDs was estimated using mixed-effect linear regression modeling with maximum likelihood estimation; the fixed effects were the dependent (selected histographic parameter) and independent (treatment) variables, whereas the random effects were the patient and participant recruitment center. Z scores calculated from the regression were then used to determine one-tailed *P* values (because of the specific directional hypothesis that NSAIDs would decrease inflammation). Interobserver agreement of qMRI parameters was evaluated using Bland–Altman limits of agreement analysis.

Data may be made available on reasonable request. The study was approved by West of Scotland Research Ethics Committee 3 (reference 17/WS/0041). All participants provided informed written consent.

RESULTS

Between June 2017 and February 2020, 311 participants were recruited from 34 centers across England and Scotland. Participants had a median age of 42 years (interquartile range [IQR] 32–52), 62% were male, and the majority (87%) were of White racial origin. All had a clinical diagnosis of axSpA; 247 (79%) met the ASAS classification criteria for axSpA, of whom 107 (43%) had r-axSpA. Median ASDAS and BASDAI scores were 2.8 (IQR 2.0–3.4) and 4.5 (IQR 2.8–6.2), respectively. Other baseline characteristics are shown in Table 1.

At baseline, most participants (83%) were taking naproxen, etoricoxib, or ibuprofen. A full list of NSAIDs is shown in Supplementary Table 2. Of the 311 participants, 286 (92%) completed the washout period and underwent scan 1. The reasons for failing to undergo scan 1 are shown in Table 2. The median washout duration was 17 (IQR 14–26) days. Fewer than half of participants reported a disease flare in the week before the scan (45%), although 69% experienced a worsening in disease activity and/or spinal pain. However, even in this group, the magnitude of worsening was modest; the median changes in the BASDAI and spinal pain scores were 0.8 (IQR 0.3–1.6) and 1 (IQR 0–2), respectively. The distribution of change in the BASDAI and spinal pain scores is shown in Supplementary Figure 2. Due to a high amount of missing data for CRP, it was only possible to calculate change in ASDAS for 148 participants. Mean change was 0.05 (95% CI –0.02 to 0.11), and only three participants reported a clinical important worsening over the washout period.[27]

Interreader agreement was good (pairwise comparisons ranged from 77% to 88%; κ = 0.53, 0.60, and 0.76), and fewer

Table 1. Characteristics of study participants*

	Participants (N = 311)[a]	
	n (%)	Median (IQR)
Demographic characteristics		
Sex		
Female	117 (37.6)	–
Male	194 (62.4)	–
Age, y	–	42 (32–52)
Race and ethnicity		
White	271 (87.1)	–
Non-White[b]	40 (4.2)	–
Clinical characteristics		
Symptom duration, y	–	9 (4–20)
Time since diagnosis, y	–	1 (0–7)
CRP ≥4 mg/L		
No	168 (54.0)	–
Yes	143 (46.0)	–
CRP,[c] mg/L	–	6 (2.5–11)
Disease classification		
Radiographic axSpA	107 (34.4)	–
Non-radiographic axSpA	140 (45.0)	–
Did not meet ASAS criteria	64 (20.6)	–
HLA–B27		
Positive	171 (55.0)	–
Negative	71 (22.8)	–
Unknown	69 (22.2)	–
Patient-reported characteristics		
Disease activity		
ASDAS[c]	–	2.8 (2.0–3.4)
BASDAI	–	4.5 (2.8–6.2)
Function, BASFI	–	3.1 (1.5–5.6)
Global health		
BAS-G	–	5 (3–6.5)
PROMIS physical[d]	–	15 (14–16)
PROMIS mental[d]	–	13 (11–16)
Quality of life, ASQoL	–	7 (3–12)
Fibromyalgia[e]		
Yes	68 (22.5)	–
No	234 (77.5)	–
Daily NSAIDs, before washout		
Yes	266 (85.5)	–
No	45 (14.5)	–

* ASAS, Assessment of SpondyloArthritis international Society; ASDAS, Ankylosing Spondylitis Disease Activity Score; ASQoL, Ankylosing Spondylitis Quality of Life questionnaire; axSpA, axial spondyloarthritis; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; BAS-G, Bath Ankylosing Spondylitis Patient Global Score; CRP, C-reactive protein; IQR, interquartile range; NSAID, nonsteroidal anti-inflammatory drug; PROMIS, Patient-Reported Outcomes Measurement Information System.
[a] Precise denominator may vary due to missing data.
[b] n = 13 Asian/Asian British (Indian, Pakistani, Bangladeshi); n = 3 Black British/Caribbean/Black African; n = 24 mixed or other.
[c] Absolute CRP values were available only for 192 participants. This also influenced the number of participants for whom it was possible to compute an ASDAS value.
[d] PROMIS global physical health and global mental health scales.
[e] 2016 revisions to the 2010/2011 fibromyalgia diagnostic criteria.[17]

Table 2. Reasons for dropout before scan 1*

	n
Failed washout attempt	
Could not tolerate washout period	2
Adverse event/required NSAID for other condition	4
Voluntary withdrawal from study[a]	1
Did not attend scan appointment	3
Other/unknown	15

* NSAID, nonsteroidal anti-inflammatory drug.
[a] Patient completed washout period and reported feeling very well. He was unwilling to recommence NSAIDs, as was required by study protocol.

than 10% required third-reader adjudication. In total, 146 participants (51%) had a positive scan result for active sacroiliitis at scan 1, and a follow-up scan (while taking NSAIDs) was obtained from 124 (85% of those eligible). The median between-scan interval was 42 (IQR 42–48) days. Twenty-five participants (20.2%) had a negative scan result for active sacroiliitis (95% CI 13.5%–28.3%).

There was no evidence of a sex difference in the risk of scoring negative for sacroiliitis while taking NSAIDs (risk ratio [male vs female] 1.25; 95% CI 0.59–2.68). Indeed, there were no significant associations between any demographic, clinical, or patient-reported characteristics and the risk of sacroiliitis after recommencing NSAIDs (see Table 3).

Semiquantitative scoring (Leeds MRI scoring method). The median total Leeds scores were 4 (IQR 2–8) and 3 (IQR 2–7) for scan 1 and scan 2, respectively. While taking NSAIDs, more than two-fifths of participants experienced a reduction in the Leeds score (44%; 95% CI 35%–54%), with a median change of –2 (IQR –3 to –1).

qMRI analysis. One hundred four patients underwent pre-NSAID Dixon scans for FF measurement, of whom 39 had repeat scans. Changes in FF histographic parameters with NSAIDs are summarized in Table 4 and Figure 1. FF measurements were higher on scans acquired while participants were taking NSAIDs compared to the scans acquired while participants were not taking NSAIDs, with significant increases observed for FF_{75} and FF_{90}. The observed increases in the values of FF_{10}, FF_{25}, and FF_{median} were numerically smaller than those in FF_{75} and FF_{90} and did not reach statistical significance. Bland–Altman limits of agreement analysis showed good interobserver agreement generally, for all histographic parameters (example plots are shown in Supplementary Figure 3).

DISCUSSION

This study demonstrates that among patients with axSpA with evidence of active sacroiliitis after an NSAID washout, one in five exhibit resolution of sacroiliitis six weeks after NSAIDs have been reintroduced. This provides evidence of the effect of NSAIDs on active sacroiliitis and shows that NSAID use before MRI may have an important impact on clinical diagnosis. Further, we show that although patients may experience a deterioration in disease activity and spinal pain during NSAID washout, almost all patients who attempted NSAID washout were able to tolerate it. Thus, stopping NSAIDs for one to two weeks before MRI scanning in clinical practice should be considered.

There are a number of methodologic issues to discuss. Firstly, it would have been possible to answer the research question with a randomized crossover trial. However, this was discounted for reasons of inefficiency and concerns over participant well-being. For example, some participants would have a negative scan result at both time points and would not contribute any data to the main research question despite having undergone the entire study protocol, including two MRI scans and an NSAID washout period. The real challenge of interpretation therefore is whether change in sacroiliitis after NSAID reintroduction might be due to the natural course of MRI inflammation. It would have been interesting to rescan some participants six weeks after a positive scan result for sacroiliitis but without restarting NSAIDs. This was considered and would have provided some data on the natural history of sacroiliitis but was discounted; it would have been necessary to ask participants to remain off their usual medication for four times longer than was necessary to answer the main research question. This was felt not to be justifiable from an ethical perspective.

The pertinent issue, however, is whether the change in sacroiliitis at scan 2 could be explained by natural variation in sacroiliitis. The results might be explained by regression to the mean if study participants were all recruited at a time of high disease activity (assuming corresponding active inflammatory lesions). However, there was no significant difference in the proportion of participants who had a negative scan result at scan 2 between those who did versus did not report a flare at the time of scan 1. Current evidence on the natural variation of sacroiliitis over the short term remains scarce. A recent study showed that short-lived fluctuations within a few days in MRI-determined BME were more common with longer-acting TNF inhibitors and corresponded with a subjective loss of clinical response before the next scheduled dose.[28] However, in a previous study, 29 participants with normal sacroiliac joint radiographs who fulfilled the ASAS criteria for inflammatory back pain underwent four MRI scans over a 12-week period. Results showed that of the 10 who had a positive result for inflammation in the spine and/or sacroiliac joints, all these participants had subsequent positive results at 4, 8, and 12 weeks.[29] Nearly half the patients continued their regular NSAID intake during the study, and none were receiving biologic therapy. Thus, although it is possible that the resolution of inflammatory lesions in the DyNAMISM study (over a six-week period) occurred spontaneously, this earlier study suggests that rapid fluctuation from a positive to a negative result for

Table 3. Factors associated with a positive or negative scan result while taking NSAIDs*

	Scan 2, sacroiliitis, n (%)		
	Positive (n = 99)[a]	Negative (n = 25)[a]	RR[b] (95% CI)
Demographic characteristics			
Sex			
Female	38 (82.6)	8 (17.4)	–
Male	61 (78.2)	17 (21.8)	1.25 (0.59–2.68)
Age			
≤39 y	52 (83.9)	10 (16.1)	–
≥40 y	47 (75.8)	15 (24.2)	1.50 (0.73–3.09)
Race and ethnicity			
White	85 (78.7)	23 (21.3)	–
Non-White	14 (87.5)	2 (12.5)	0.57 (0.15–2.27)
Clinical characteristics			
Symptom duration			
≤8 y	52 (82.5)	11 (17.5)	–
≥9 y	47 (78.3)	13 (21.7)	1.24 (0.60–2.56)
Time since diagnosis			
≤1 y	50 (80.6)	12 (19.4)	–
≥2 y	49 (80.3)	12 (19.7)	1.02 (0.49–2.09)
CRP ≥4 mg/L			
No	48 (78.7)	13 (21.3)	–
Yes	51 (81.0)	12 (19.0)	0.89 (0.44–1.81)
ASAS axSpA criteria			
No	13 (65.0)	7 (35.0)	–
Yes	86 (82.7)	18 (17.3)	0.49 (0.24–1.03)
Radiographic axSpA			
No	71 (82.6)	15 (17.4)	–
Yes	28 (73.7)	10 (26.3)	1.51 (0.74–3.06)
HLA–B27			
Positive	60 (83.3)	12 (16.7)	–
Negative	23 (82.1)	5 (17.9)	1.07 (0.41–2.77)
Unknown	16 (66.7)	8 (33.3)	2.00 (0.93–4.32)
Patient-reported characteristics			
Disease activity			
ASDAS			
≤3.99	34 (79.1)	9 (20.9)	–
≥4.00	36 (83.7)	7 (16.3)	0.78 (0.32–1.91)
BASDAI			
≤4.4	53 (85.5)	9 (14.5)	–
≥4.5	45 (73.8)	16 (26.2)	1.81 (0.86–3.78)
Function, BASFI			
≤2.7	52 (83.9)	10 (16.1)	–
≥2.8	46 (75.4)	15 (24.6)	1.52 (0.74–3.13)
Global health			
BAS-G			
≤4.5	54 (84.4)	10 (15.6)	–
≥5.0	44 (74.6)	15 (25.4)	1.63 (0.79–3.34)
PROMIS physical[c]			
≤15	63 (77.8)	18 (22.2)	–
≥16	35 (83.3)	7 (16.7)	0.75 (0.34–1.66)
PROMIS mental[c]			
≤13	50 (79.4)	13 (20.6)	–
≥14	48 (80.0)	12 (20.0)	0.97 (0.48–1.96)
Quality of life, ASQoL			
≤6	51 (79.7)	13 (20.3)	–
≥7	47 (79.7)	12 (20.3)	1.00 (0.50–2.02)
Fibromyalgia[d]			
No	81 (80.2)	20 (19.8)	–
Yes	17 (77.3)	5 (22.7)	1.15 (0.48–2.73)
Duration of NSAID washout			
NSAID washout			
≤14 days[e]	10 (25.0)	30 (75.0)	–

(Continued)

Table 3. (*Cont'd*)

	Scan 2, sacroiliitis, n (%)		
	Positive (n = 99)[a]	Negative (n = 25)[a]	RR[b] (95% CI)
15–21 days	8 (17.8)	37 (82.2)	0.71 (0.31–1.63)
≥22 days	7 (17.9)	32 (82.1)	0.72 (0.30–1.70)

Note: Asian/Asian British (Indian, Pakistani, Bangladeshi); Black British/Caribbean/Black African; or Other.
* ASAS, Assessment of SpondyloArthritis international Society; ASDAS, Ankylosing Spondylitis Disease Activity Score; ASQoL, Ankylosing Spondylitis Quality of Life questionnaire; axSpA, axial spondyloarthritis; BASDAI, Bath Ankylosing Spondylitis Disease Activity Index; BASFI, Bath Ankylosing Spondylitis Functional Index; BAS-G, Bath Ankylosing Spondylitis Patient Global Score; CI, confidence interval; CRP, C-reactive protein; NSAID, nonsteroidal anti-inflammatory drug; PROMIS, Patient-Reported Outcomes Measurement Information System; RR, risk ratio.
[a] Precise denominator may vary due to missing data.
[b] RR and 95% CI given per category. The reference category is always listed first. Continuous variables have been categorized and split above or below the median value due to violations of linearity assumptions.
[c] PROMIS global physical health and global mental health scales.
[d] 2016 revisions to the 2010/2011 fibromyalgia diagnostic criteria.[17]
[e] Includes 5 participants who had a washout duration of ≤7 days and 35 who had a washout duration of 8–14 days.

sacroiliitis is uncommon. Hence, we believe that the observed changes in the current study can reasonably be attributed to NSAIDs.

It is important to highlight that the DyNAMISM study population is predominantly male. This is due, at least in part, to the fact that in the early part of recruitment, we prioritized men who were HLA–B27 positive to increase study efficiency; this is a subgroup known to have higher likelihood of a positive sacroiliac joint MRI finding.[9] Nevertheless, this criterion was broadened during the study to include anyone (male or female) with a prior positive MRI finding. In terms of external validity, there is no reason to believe that if NSAIDs affect appearance of MRI in men why this would not also be the same in women, an assumption that was borne out by the results.

Unfortunately, 22% of patients were missing data on HLA–B27 status, although this was expected because HLA–B27 was not a protocol-mandated test, and other studies have had similar findings. For example, in the British Society for Rheumatology Biologics Register for Ankylosing Spondylitis, a large registry that recruited from around 80 hospitals across Great Britain, >40% of participants were missing information on HLA–B27.[30] Pertinent to the current study, although knowledge of HLA–B27 status might influence diagnosis and can definitely influence disease classification, it will not necessarily influence the decision of whether a patient is referred for imaging. Further, HLA–B27 status was not a determinant of whether a patient changed status (from positive to negative on MRI).

At recruitment, all participants were taking (or were to commence) daily NSAIDs, and after scan 1, all restarted (or commenced) daily NSAIDs. Scan 2 was undertaken six weeks subsequently. ASAS/EULAR recommendations for the management of axSpA recommend that before escalating to biologic or targeted synthetic disease-modifying antirheumatic drugs, non-pharmacological treatments and at least two NSAIDs should be tried over a four-week period.[4] We would argue that if four weeks is sufficient to determine effectiveness of NSAIDs in a management context, six weeks should be sufficient to observe any effect on active sacroiliitis should such an effect exist. It is possible that the full effect of NSAIDs on active sacroiliitis may take longer, but if this is the case, then the actual effect observed in the current study would be an underestimate.

From a clinical perspective, a potential concern is whether patients can tolerate an NSAID washout period in the first place. The majority of patients in the DyNAMISM study reported an increase in disease activity and/or spinal pain during the washout period, and 45% of participants reported a flare in the week before scan 1. However, increases in disease activity and spinal pain were relatively small; even among those with a deterioration in BASDAI and/or spinal pain scores, the median increase was ≤1 point. Reassuringly, almost all participants (>90%) who

Table 4. Change in qMRI parameters while taking NSAIDs*

FF[a]	Change with NSAID, β coefficient[b]	95% CI[b]	Z score[b]	*P* value
FF_{mean}	0.550	−0.215 to 1.315	1.51	0.080
FF_{median}	0.568	−0.233 to 1.369	1.39	0.082
FF_{10}	0.086	−0.928 to 1.010	0.17	0.434
FF_{25}	0.495	−0.370 to 1.360	1.12	0.131
FF_{75}	0.809	−0.065 to 1.683	1.81	0.035[c]
FF_{90}	0.916	0.056 to 1.776	2.09	0.019[c]

* Comparison of histographic parameters obtained between pre- and posttreatment patients. CI, confidence interval; FF, fat fraction; NSAID, nonsteroidal anti-inflammatory drug; qMRI, quantitative magnetic resonance imaging.
[a] Subscript numbers denote the 10th, 25th, 75th, and 90th percentiles of FF in the defined regions of interest.
[b] Coefficient, 95% CIs, and Z values were derived from the mixed-effect linear regression.
[c] *P* values were calculated separately and refer to a one-tailed z test.

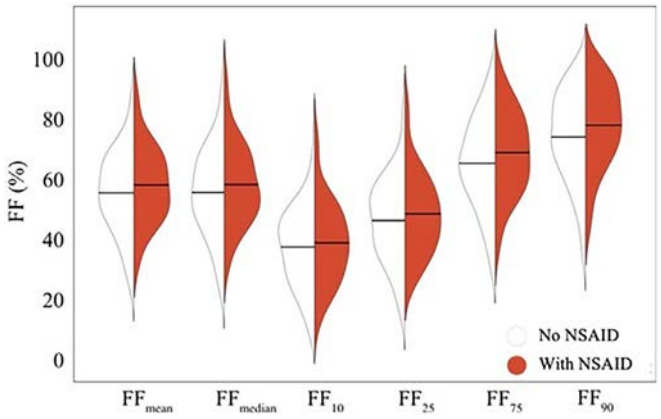

Figure 1. Violin plots demonstrating changes in quantitative magnetic resonance imaging parameters. FF-based histographic parameters are shown from before and after treatment. Note that the variability in the baseline FF values is effectively dealt with in the mixed-effect linear regression because each patient acts as their own control. FF, fat fraction; NSAID, nonsteroidal anti-inflammatory drug.

attempted the NSAID washout were able to achieve it, and in the subgroup with ASDAS measurements both before and after washout, only a small minority (2%) reported a clinically important worsening.

The DyNAMISM study protocol did not stipulate the specific NSAID treatment. Instead, participants took the brand and dose as prescribed by their rheumatologist, reflecting real-world clinical practice. The half-life of NSAIDs vary, from <1 hour[31] to around 48 hours,[32] so it is possible that if a washout period is too short, the drugs might still be exerting an effect, increasing the chance of a negative scan result. In contrast, if the washout period is too long, although increasing the chances of patients having a positive scanning result (if indeed the results are positive), there may be implications for patient comfort and safety. In the current study, very few participants (<5%) had an NSAID washout period of less than one week, and it was not meaningful to examine this group separately. Instead, the study population was divided into washout periods of ≤14 days, 15 to 21 days, and ≥22 days. There was no evidence that a longer washout duration was associated with response (a negative result at scan 2) after NSAIDs were reintroduced. The current study provides no evidence of any benefit of extending the washout beyond 14 days.

We were unable to identify any clinical characteristics that would indicate, in advance, who might be likely to change status (from positive for sacroiliitis to negative). Some small differences were evident, although there were no consistent or significant differences between groups, but the study was not specifically powered to detect such differences. It was interesting that compared to those who did not, participants who met ASAS axSpA classification criteria were half as likely to have a negative scan result in the presence of NSAIDs. Despite this, still around one in five participants changed status after NSAIDs were reintroduced. The high proportion of missing data for HLA–B27 status has already been discussed. However, it is interesting to note that among those for whom data were available, there was no difference in the proportion who scanned positive at scan 2 between those who were B27 positive versus negative (83.3% vs 82.1%).

In the current study, for the primary outcome, all scans were read by two readers (with adjudication by a third, if required). All were rheumatologists with a specialist interest in axSpA and experience in applying the ASAS criteria for active inflammatory lesions (positive MRI finding) of the sacroiliac joint.[20] Interreader agreement was good, with only 12% of all scans requiring adjudication. Reliability in reading the Leeds MRI scoring method was also examined. Here, each quadrant of each sacroiliac joint is scored from 0 to 3, giving a total score of 0 to 24. Although it is harder to gain agreement on a semiquantitative scale, interreader reliability was excellent. The Leeds method was chosen because of its feasibility, reader familiarity, and the fact that it incorporates a subjective assessment of extent and intensity of the BME lesion as a proxy of lesion severity. Indeed, results from the Leeds scoring method showed that 44% of participants exhibited a reduction in BME severity in the presence of NSAIDs, more than twice the proportion who changed status from ASAS positive to ASAS negative. This suggests that NSAIDs have an effect on decreasing the severity of inflammation, although this may not always be sufficient to fully suppress it.

To explore this effect further, a subset of participants underwent qMRI. This novel feature of the study provides quantitative support for the hypothesis that NSAIDs reduce inflammation on MRI. We observed larger increases in the parameters targeting the upper "fatty" end of the distribution than in the parameters targeting the lower "inflamed" end of the distribution, suggesting an emergence or "uncovering" of fat metaplasia in the bone marrow of participants not taking NSAIDs. Further qMRI studies may help to define the response phenotype more precisely. In general, qMRI analyses may provide useful secondary end points in clinical studies, as the subjectivity inherent to conventional MRI reading is minimized. A further benefit is that with emerging techniques, MRI changes due to BME can potentially be disentangled from those due to structural damage, which can otherwise confound interpretation.[33] The use of these qMRI techniques and/or artificial intelligence–based methods could allow a more detailed analysis clarifying the role of NSAIDs in modulating BME and also its impact on disease progression in axSpA.[34]

In conclusion, these findings provide important insights on the impact of NSAIDs on the appearance of MRI-determined BME lesions representative of active sacroiliitis in axSpA. Significant improvements in the severity of the lesions were observed in >40% of participants, with full resolution seen in 20%. Some patients will be unwilling to attempt an NSAID washout, fearful of the increase in pain and disease activity. However, this study also provides reassuring evidence that among patients willing to attempt a washout, almost all can successfully achieve

it. These data have important implications for axSpA diagnosis and classification and, in clinical practice, support consideration of a 14-day NSAID washout before MRI.

ACKNOWLEDGMENTS

We are indebted to the staff and participants at the recruiting centers. We also acknowledge study coordinators Dr Karen Forrest Keenan and Laura Moir, the Centre for Healthcare Randomised Trials (CHaRT) for providing and maintaining the information technology infrastructure, and Dolapo Ayansina for statistical input during the study.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft (A CRediT statement detailing author contributions is shown in Supplementary Table 3a/3b). As corresponding author, Dr Jones confirms that all authors have provided the final approval of the version to be published and takes responsibility for the affirmations regarding article submission (eg, not under consideration by another journal), the integrity of the data presented, and the statements regarding compliance with institutional review board/Declaration of Helsinki requirements.

REFERENCES

1. Rudwaleit M, van der Heijde D, Landewé R, et al. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. Ann Rheum Dis 2009;68(6):777–783.
2. National Institute for Health and Care Excellence. Spondyloarthritis in over 16s: diagnosis and management. 2017. www.nice.org.uk/guidance/ng65
3. Bray TJP, Jones A, Bennett AN, et al; British Society of Spondyloarthritis (BRITSpA). Recommendations for acquisition and interpretation of MRI of the spine and sacroiliac joints in the diagnosis of axial spondyloarthritis in the UK. Rheumatology (Oxford) 2019;58(10):1831–1838.
4. Ramiro S, Nikiphorou E, Sepriano A, et al. ASAS-EULAR recommendations for the management of axial spondyloarthritis: 2022 update. Ann Rheum Dis 2023;82(1):19–34.
5. van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. Arthritis Rheum 1984;27(4):361–368.
6. Jarrett SJ, Sivera F, Cawkwell LS, et al. MRI and clinical findings in patients with ankylosing spondylitis eligible for anti-tumour necrosis factor therapy after a short course of etoricoxib. Ann Rheum Dis 2009;68(9):1466–1469.
7. Varkas G, Jans L, Cypers H, et al. Brief report: six-week treatment of axial spondyloarthritis patients with an optimal dose of nonsteroidal antiinflammatory drugs: early response to treatment in signal intensity on magnetic resonance imaging of the sacroiliac joints. Arthritis Rheumatol 2016;68(3):672–678.
8. Dubash S, Pease C, Aslam A, et al. Acute unilateral sacroiliitis mimicking infection on magnetic resonance imaging with response to nonsteroidal antiinflammatory drugs: a distinct presentation of spondyloarthritis? J Rheumatol 2018;45(12):1708–1710.
9. van Onna M, Jurik AG, van der Heijde D, et al. HLA-B27 and gender independently determine the likelihood of a positive MRI of the sacroiliac joints in patients with early inflammatory back pain: a 2-year MRI follow-up study. Ann Rheum Dis 2011;70(11):1981–1985.
10. Garrett S, Jenkinson T, Kennedy LG, et al. A new approach to defining disease status in ankylosing spondylitis: the Bath Ankylosing Spondylitis Disease Activity Index. J Rheumatol 1994;21(12):2286–2291.
11. Lukas C, Landewé R, Sieper J, et al; Assessment of SpondyloArthritis international Society. Development of an ASAS-endorsed disease activity score (ASDAS) in patients with ankylosing spondylitis. Ann Rheum Dis 2009;68(1):18–24.
12. Jones SD, Steiner A, Garrett SL, et al. The Bath Ankylosing Spondylitis Patient Global Score (BAS-G). Br J Rheumatol 1996;35(1):66–71.
13. Calin A, Garrett S, Whitelock H, et al. A new approach to defining functional ability in ankylosing spondylitis: the development of the Bath Ankylosing Spondylitis Functional Index. J Rheumatol 1994;21(12):2281–2285.
14. The Patient-Reported Outcomes Measurement Information System (PROMIS). Scale v1.2 – Global Health. Accessed November 21, 2016. http://www.healthmeasures.net/administrator/components/com_instruments/uploads/Global%20Health%20Scale%20v1.2%2008.22.2016.pdf
15. Cella D, Yount S, Rothrock N, et al; PROMIS Cooperative Group. The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. Med Care 2007;45(5 suppl 1):S3–S11.
16. Doward LC, Spoorenberg A, Cook SA, et al. Development of the ASQoL: a quality of life instrument specific to ankylosing spondylitis. Ann Rheum Dis 2003;62(1):20–26.
17. Wolfe F, Clauw DJ, Fitzcharles MA, et al. 2016 revisions to the 2010/2011 fibromyalgia diagnostic criteria. Semin Arthritis Rheum 2016;46(3):319–329.
18. Schueller-Weidekamm C, Mascarenhas VV, Sudol-Szopinska I, et al. Imaging and interpretation of axial spondylarthritis: the radiologist's perspective--consensus of the Arthritis Subcommittee of the ESSR. Semin Musculoskelet Radiol 2014;18(3):265–279.
19. Lambert RGW, Bakker PA, van der Heijde D, et al. Defining active sacroiliitis on MRI for classification of axial spondyloarthritis: update by the ASAS MRI working group. Ann Rheum Dis 2016;75(11):1958–1963.
20. Rudwaleit M, Jurik AG, Hermann KGA, 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(10):1520–1527.
21. Marzo-Ortega H, McGonagle D, O'Connor P, et al. Efficacy of etanercept in the treatment of the entheseal pathology in resistant spondylarthropathy: a clinical and magnetic resonance imaging study. Arthritis Rheum 2001;44(9):2112–2117.
22. Marzo-Ortega H, McGonagle D, O'Connor P, et al. Baseline and 1-year magnetic resonance imaging of the sacroiliac joint and lumbar spine in very early inflammatory back pain. Relationship between symptoms, HLA-B27 and disease extent and persistence. Ann Rheum Dis 2009;68(11):1721–1727.
23. Bray TJP, Bainbridge A, Punwani S, et al. Simultaneous quantification of bone edema/adiposity and structure in inflamed bone using chemical shift-encoded MRI in spondyloarthritis. Magn Reson Med 2018;79(2):1031–1042.
24. Bray TJP, Sakai N, Dudek A, et al. Histographic analysis of oedema and fat in inflamed bone marrow based on quantitative MRI. Eur Radiol 2020;30(9):5099–5109.
25. Bainbridge A, Bray TJP, Sengupta R, et al. Practical approaches to bone marrow fat fraction quantification across magnetic

resonance imaging platforms. J Magn Reson Imaging 2020;52(1): 298–306.

26. Altman DG, Bland JM. Measurement in medicine: the analysis of method comparison studies. Statistician 1983;32(3):307–317.
27. Molto A, Gossec L, Meghnathi B, et al; ASAS-FLARE study group. An Assessment in SpondyloArthritis International Society (ASAS)-endorsed definition of clinically important worsening in axial spondyloarthritis based on ASDAS. Ann Rheum Dis 2018;77(1): 124–127.
28. Harrison SR, Ansell R, Mathieson HR, et al. Subjective loss of clinical response to TNFi in axSpA relates to recurrence of MRI bone marrow oedema particularly with long-acting agents. Rheumatology (Oxford) 2022;61(SI):SI86–SI91.
29. Sengupta R, Marzo-Ortega H, McGonagle D, et al; British Society for Spondyloarthritis. Short-term repeat magnetic resonance imaging scans in suspected early axial spondyloarthritis are clinically relevant only in HLA-B27–positive male subjects. J Rheumatol 2018;45(2): 202–205.
30. Jones GT, Dean LE, Pathan E, et al. Real-world evidence of TNF inhibition in axial spondyloarthritis: can we generalise the results from clinical trials? Ann Rheum Dis 2020;79(7):914–919.
31. Awtry EH, Loscalzo J. Aspirin. Circulation 2000;101(10):1206–1218.
32. Verbeeck RK, Richardson CJ, Blocka KL. Clinical pharmacokinetics of piroxicam. J Rheumatol 1986;13(4):789–796.
33. Gollifer RM, Bray TJP, Kruezi A, et al. A multiparametric alternative to short inversion-time inversion recovery for imaging inflammation: T_{2water} and fat fraction measurement using chemical shift-encoded turbo spin-echo MRI. Magn Reson Med 2023;90(3):1010–1024.
34. Hepburn C, Jones A, Bainbridge A, et al. Volume of hyperintense inflammation (V_{HI}): a quantitative imaging biomarker of inflammation load in spondyloarthritis, enabled by human-machine cooperation. PLoS One 2023;18(4):e0284508.

Arthritis Care & Research
Vol. 77, No. 12, December 2025, pp 1453–1458
DOI 10.1002/acr.25593

American College of Rheumatology
Empowering Rheumatology Professionals

BRIEF REPORT

Assessment of Rheumatology Fellows' Skills as Clinical Teachers Through Self-Assessment and Direct Observation

David Leverenz,[1] Marcy B. Bolster,[2] Lisa Criscione-Schreiber,[1] Jon Golenbiewski,[3] Faye Hant,[4] Rumey Ishizawar,[5] Jennifer Schmidt,[4] Amanda Snyder,[5] Rachel Wolfe,[3] and Eli M. Miloslavsky[2]

Objective. The purpose of this study is to assess rheumatology fellows' teaching skills through an observed structured teaching exercise (OSTE), self-assessment, and survey of fellows' teaching experiences.

Methods. Rheumatology fellows from five institutions participated in an in-person OSTE, involving a simulated teaching encounter with a standardized learner. Trained faculty observers rated each OSTE encounter to assess the fellows' proficiency as a clinical teacher in the following domains: learning environment, learner assessment, presenting material, feedback, and overall teaching ability. Before the OSTE, fellows completed a self-assessment of their teaching skills according to those same five domains. In addition, they completed a post-OSTE survey assessing their experience with teaching during rheumatology fellowship training and their experience with the OSTE station itself.

Results. A total of 25 fellows completed the OSTE and self-assessment. According to preceptor ratings on the OSTE, the domain with the highest average proficiency was presenting material (4.16, SD 0.46), and the lowest was learner assessment (3.06, SD 1.56). There was no significant correlation between OSTE ratings and fellow self-assessment in any domain. Of the 23 fellows (92%) who completed the post-OSTE survey, only 57% agreed they had received high-quality feedback on their teaching skills during fellowship training, and 100% agreed they received effective feedback during the OSTE.

Conclusion. Fellows' self-assessed teaching ability does not correlate with direct observation. Interventions, such as this OSTE, are useful for providing high-quality feedback on fellows' teaching skills.

INTRODUCTION

When rheumatology fellows teach during clinical encounters, they strengthen their own knowledge through the active engagement of learners and improve patient care by promoting effective communication between teams.[1] As such, teaching has been identified as part of the core competencies of rheumatology training by the Accreditation Council for Graduate Medical Education (ACGME), in that over half of the "level 5" milestones refer to educating, coaching, and/or role modeling the subcompetency to others. Teaching also facilitates development of skills that are critical for effective patient education in future independent practice. In addition, teaching and mentorship of medical students and residents are critical to recruiting trainees to the specialty, as most trainees have little to no clinical experience in rheumatology.[2] Because fellows have considerable interaction with students and residents in the setting of clinical care, their influence is particularly important in addressing the anticipated rheumatology workforce shortage.

Despite the importance of training rheumatology fellows to be effective teachers, little is known about the current state of rheumatology fellows' teaching skills. Trainee-as-teacher programs have increased in medical schools and residency programs, though teacher training during subspecialty fellowship training remains an underdeveloped area.[3] Additionally, it is uncertain how these programs impact current fellows' teaching

[1]David Leverenz, MD, MEd, Lisa Criscione-Schreiber, MD, MEd: Duke University School of Medicine, Durham, North Carolina; [2]Marcy B. Bolster, MD, Eli M. Miloslavsky, MD: Massachusetts General Hospital, Boston; [3]Jon Golenbiewski, DO, Rachel Wolfe, MD: Wake Forest School of Medicine, Winston-Salem, North Carolina; [4]Faye Hant, MD, Jennifer Schmidt, MD: Medical University of South Carolina, Charleston; [5]Rumey Ishizawar, MD, Amanda Snyder, MD: University of North Carolina at Chapel Hill.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25593).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25593.

Address correspondence via email to David Leverenz, MD, MEd, at David.Leverenz@duke.edu.

Submitted for publication December 19, 2024; accepted in revised form June 23, 2025.

SIGNIFICANCE & INNOVATIONS

- Rheumatology fellows' self-assessment of their teaching ability does not correlate with faculty ratings of their teaching ability in an observed structured teaching exercise (OSTE).
- According to faculty ratings on the OSTE, fellows' highest teaching proficiency is presenting material and the lowest is learner assessment.
- Only 57% of fellows agreed that they had received high-quality feedback on their teaching skills during fellowship training, and 100% agreed that they received effective feedback during the OSTE.

skills or, more importantly, their ability to accurately self-assess those skills given the importance of active reflection on one's own abilities, a crucial component of Kolb's experiential learning cycle (practice, feedback, reflection, practice).[4] As of 2017, only half of rheumatology fellows had access to training as teachers during residency, 60% had an interest in pursuing a teaching-focused career, and nearly all fellows (88%) reported an interest in enhancing their teaching skills.[5]

Defining the current state of fellows' teaching skills is critical for the development and enrichment of effective teacher training curricula during fellowship. To investigate this area of interest, we assessed fellow teaching skills with direct observation and compared those observations with fellows' self-assessment of their teaching skills. We also explored factors that may impact fellow teaching or self-assessment skills.

MATERIALS AND METHODS

The study took place as part of the Carolina Fellows Collaborative–Massachusetts General Hospital (MGH) Winter Symposium in January 2024. Rheumatology fellows from Duke University, MGH, Medical University of South Carolina, the University of North Carolina, and Wake Forest University attended an in-person, two-day collaborative learning symposium. Each participating program is similar in terms of curriculum and fellowship size (four to six ACGME fellows per program). All fellows attending this symposium were required to participate in an observed structured teaching exercise (OSTE) with the option to participate in this study, which was determined exempt by the MGH Institutional Review Board.

Similar to the observed structured clinical examination, the OSTE is both a learning and evaluation instrument that has been used previously in assessing fellow teaching skills.[6] Two faculty members (EMM and DL) observed the fellows in a simulated encounter with a standardized learner acting in a scripted role of an internal medicine intern. Each fellow participated in one of two identical stations observed by one of the faculty preceptors. The OSTE was one of five total observed stations that all fellows rotated through; the other four stations assessed other clinical skills. In the OSTE, fellows were provided a scenario describing an inpatient consultation on a patient with suspected gout (Supplemental Material I). Every fellow received a preview of the scenario three days in advance to ensure knowledge was not a barrier to effective teaching. After reading the scenario, fellows were given up to seven minutes to relay their recommendations to the intern on the primary team (the standardized learner) and teach the intern about the evaluation and/or management of gout.

Assessment instruments. The OSTE preceptors (EMM and DL) scored each fellow participant according to a previously published rating instrument assessing fellows' proficiency as a clinical teacher in the following domains: learning environment, learner assessment, presenting material, feedback, and overall teaching ability (5-point scale; Supplemental Material II).[7] Both preceptors previously completed advanced training in medical education (certificate and master's level) to gain expertise in learner assessment, and EMM had previously developed and assessed OSTE scenarios. Before the OSTE, the preceptors completed two 60-minute training sessions during which they independently rated 14 prerecorded videos from a prior OSTE study and compared ratings to improve concordance.[7] These independent ratings were not compared statistically, as they were used to improve agreement through iterative discussion. Immediately following each OSTE scenario, the preceptors used the OSTE rating scale to provide individualized feedback to the fellows about their clinical teaching.

Before the OSTE, all fellow participants were asked to complete a self-assessment of their abilities as a clinical teacher. This self-assessment was created for this study by grouping domains of the OSTE into categories that could be self-assessed on a 5-point Likert scale (poor, fair, average, good, excellent; Supplemental Material III). Domain groupings are detailed in the Analysis section. Fellow responses to the self-assessment survey were paired with their individual ratings on the OSTE to determine the correlation between fellows' self-perceived teaching skills and their observed proficiency.

After the OSTE, fellows were asked to complete an additional survey assessing their experience with teaching during rheumatology fellowship training and their experience with the OSTE station itself (Supplemental Material IV). The survey was adapted from previously published surveys of fellow teaching experiences.[5,8]

Analysis. Correlational analysis on Likert data was performed using Spearman's rho. In the self-assessment analysis, to align with self-assessment domains, OSTE assessment domains were combined by averaging ratings as follows: item 1 (learner orientation) and item 2 (learner respect) were combined

into "learning environment"; item 3 (evaluate knowledge) and item 4 (evaluate synthesis) were combined into "learner assessment"; item 5 (determine objectives), item 6 (present well-organized material), and item 7 (manage time) were combined into "presenting material"; and item 8 (positive feedback), item 9 (corrective feedback), and item 10 (recommendations and closing the loop) were combined into "feedback." Item 11 (overall teaching) was not combined and stood alone.

RESULTS

Participants in the OSTE included 25 rheumatology fellows (12 first-years, 12 second-years, 1 third-year). Twenty-five (100%) completed the self-assessment. According to preceptor ratings on the OSTE, the domain with the highest average proficiency among all fellows was presenting material (4.16, SD 0.46), and the lowest was learner assessment (3.06, SD 1.56) (Table 1). In contrast, on the fellow self-assessment, the highest domain was learning environment (4.22, SD 0.58), and the lowest was feedback (3.22, SD 0.75). There was no significant correlation between OSTE ratings and fellow self-assessment in any domain (Figure 1). OSTE ratings of first-year fellows and second and third-year fellows were similar (Supplemental Material V). There was a strong positive correlation between average OSTE ratings of preceptors 1 and 2 ($r = 0.65$, $P = 0.03$).

Following the OSTE, 23 (92%) of 25 fellows completed the post-OSTE survey. The majority of fellows agreed or strongly agreed they had sufficient opportunities to teach on inpatient consultation services (n = 16, 70%) and in non-consultation experiences, such as didactics and clinic (n = 19, 82%); their teaching skills were more advanced in fellowship than during residency (n = 18, 78%); and they had received high-quality didactics focused on teaching skills during fellowship (n = 16, 70%) (Figure 2). Although the majority also agreed they had received high-quality feedback on their teaching skills during fellowship (n=13, 57%), no one strongly agreed with this statement, 6 (26%) felt neutral, and 4 (17%) disagreed. Regarding the OSTE itself, the majority agreed or strongly agreed they received effective feedback during the OSTE (n = 23, 100%), the OSTE was valuable for their learning (n = 22, 96%), and the OSTE station helped them better assess their teaching skills (n = 22, 96%).

Table 1. Average fellow ratings of clinical teaching on the OSTE (scale 1–5) vs their average self-assessment of clinical teaching (Likert scale 1–5)*

Teaching domain	OSTE ratings (N = 25), mean (SD)	Self-assessment (N = 25), mean (SD)
Learning environment	3.72 (0.79)	4.22 (0.58)
Learner orientation	3.36 (1.25)	–
Learner respect	4.08 (0.49)	–
Learner assessment	3.06 (1.56)	3.39 (0.70)
Evaluate knowledge	3.32 (1.49)	–
Evaluate synthesis	2.80 (1.73)	–
Presenting material	4.16 (0.46)	3.57 (0.77)
Determine objectives	4.28 (0.61)	–
Well-organized material	4.00 (0.58)	–
Manage time	4.20 (0.71)	–
Feedback	3.93 (0.78)	3.22 (0.75)
Positive feedback	4.08 (1.12)	–
Corrective feedback	3.76 (1.42)	–
Recommendations	3.96 (0.93)	–
Overall teaching	3.68 (0.90)	3.39 (0.65)

* A full description of each OSTE domain and the self-assessment survey is included in the Supplemental Materials. OSTE, observed structured teaching exercise.

DISCUSSION

In this study, we engaged 25 fellows from five different institutions in an observed simulation of teaching in the inpatient consultation setting. Fellows who participated in the OSTE found it to be a valuable mechanism for receiving effective feedback about their clinical teaching skills. The simulated consultation encounter takes about 10 minutes per fellow and can easily be expanded to other contexts as a low-resource, high-impact intervention. Our study revealed that fellows cannot accurately identify their strengths and areas for development and identified areas with the greatest growth potential for fellows as educators. These skills are critical for all fellows regardless of career direction, as all rheumatologists must be able to educate colleagues, trainees, and patients.

Although resident-as-teacher programs are common in US residency training programs, relatively few fellowship programs have well-developed teacher training curricula, despite data suggesting that most fellows anticipate teaching being an important part of their careers and professional development.[3,8–11] Opportunities for directed feedback on teaching diminish as learners become more independent in their clinical practice and decision-making. Accordingly, although most fellows in our study agreed they had received high-quality didactics on teaching skills during fellowship, fewer fellows reported receiving high-quality feedback on their teaching. On the post-OSTE survey, all respondents reported receiving effective feedback on their teaching during the OSTE, indicating that this intervention is one potential method to address this gap. However, creating feedback opportunities for fellows does not require a formal training curriculum or simulated teaching interactions. Encouraging and developing systems for "on-the-fly" feedback during clinical interactions in conferences, clinics, and consultation rotations are relatively low-resource interventions to effectively address this gap.[1,6,9,10] In addition, the OSTE can be paired with more formal instruction, such as didactics, simulation, and workshops, to help fellows improve the observed skills. Integrating this feedback into fellow training is a critical link in the chain of Kolb's experiential learning cycle.

According to faculty ratings on the OSTE, fellows consistently demonstrated an excellent ability to present well-organized material; however, they struggled with learner assessment—

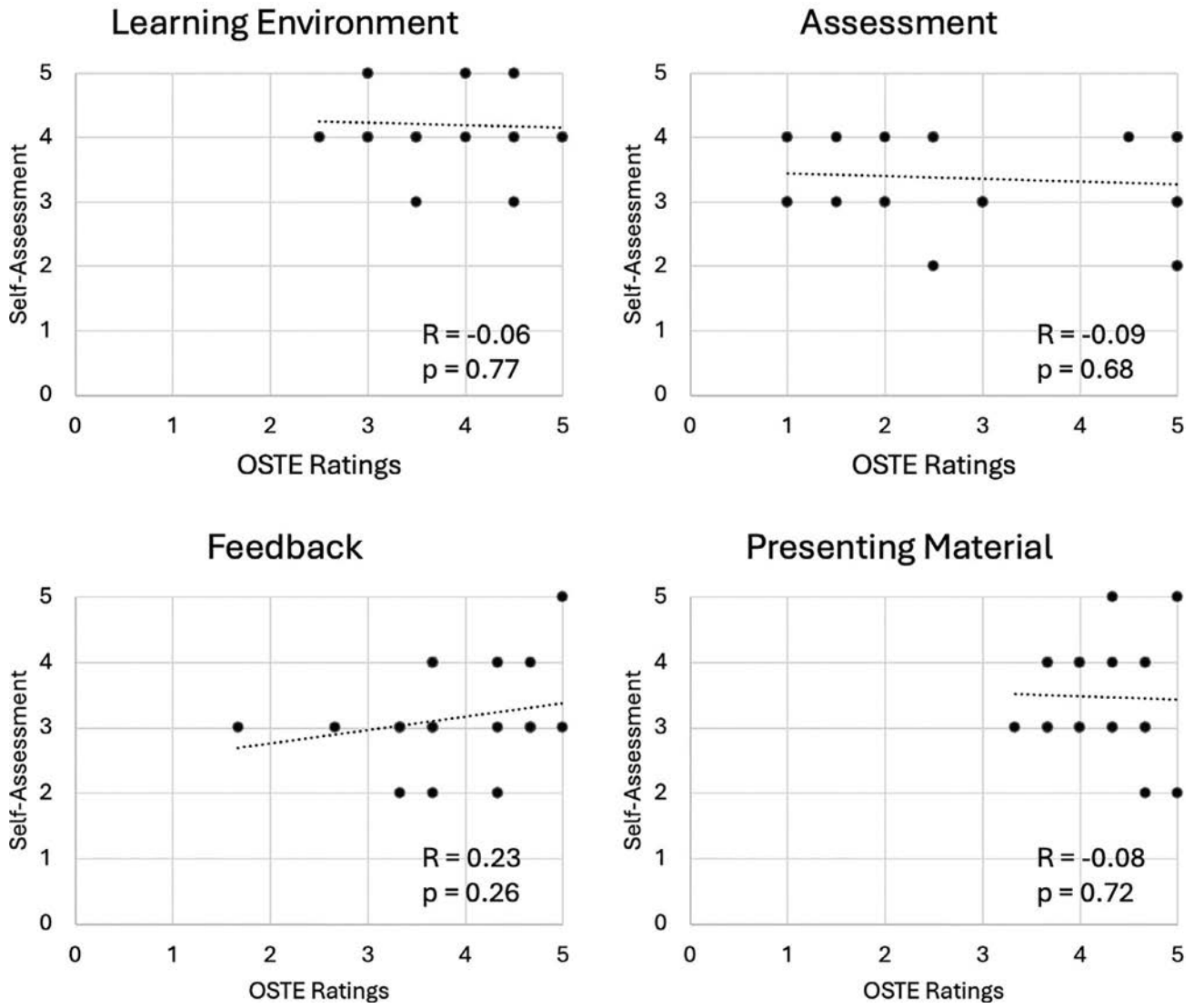

Figure 1. Correlation of OSTE ratings with learner self-assessment of clinical teaching ability. The overall teaching domain is not shown; however, there was no correlation in this domain either (r = 0.03, *P* = 0.88). A full description of each OSTE domain and the self-assessment survey is included in the Supplemental Materials. OSTE, observed structured teaching exercise.

particularly discerning the learners' level of knowledge. Although effective content delivery remains a critical teaching skill, ascertaining the content most important for an individual learner requires the ability to meet the learner where they are, recognizing their current level of knowledge and how they apply that knowledge. This skill is especially important during brief teaching interactions, such as those that occur in association with inpatient consultations, in which there is limited time available and the learner's knowledge level may be unknown to the consultant.[7,12] Likewise, the OSTE is a time-limited activity requiring fellows to assess learners quickly, similar to a realistic teaching interaction involving patient care and consultation. Lack of skill in learner assessment may lead fellows to avoid this aspect of the teaching interaction, particularly in time-constrained settings, thus missing effective teaching opportunities and limiting their professional growth as teachers. Our study demonstrated the highest variability in OSTE ratings in the domain of learner assessment (SD 1.56), indicating this is the domain with the largest opportunity for improvement. Assessment skills are some of the most challenging to develop among clinician educators, especially when deficiencies in this domain extend into independent practice. Therefore, it is crucial for educators to emphasize this domain when training fellows as teachers.

There was no significant correlation between fellows' self-assessment of their teaching abilities and faculty ratings in any teaching domain. Previous studies have demonstrated that learners may be unaware of the gaps in their teaching ability.[13] Equally important is an insufficient awareness of their strengths

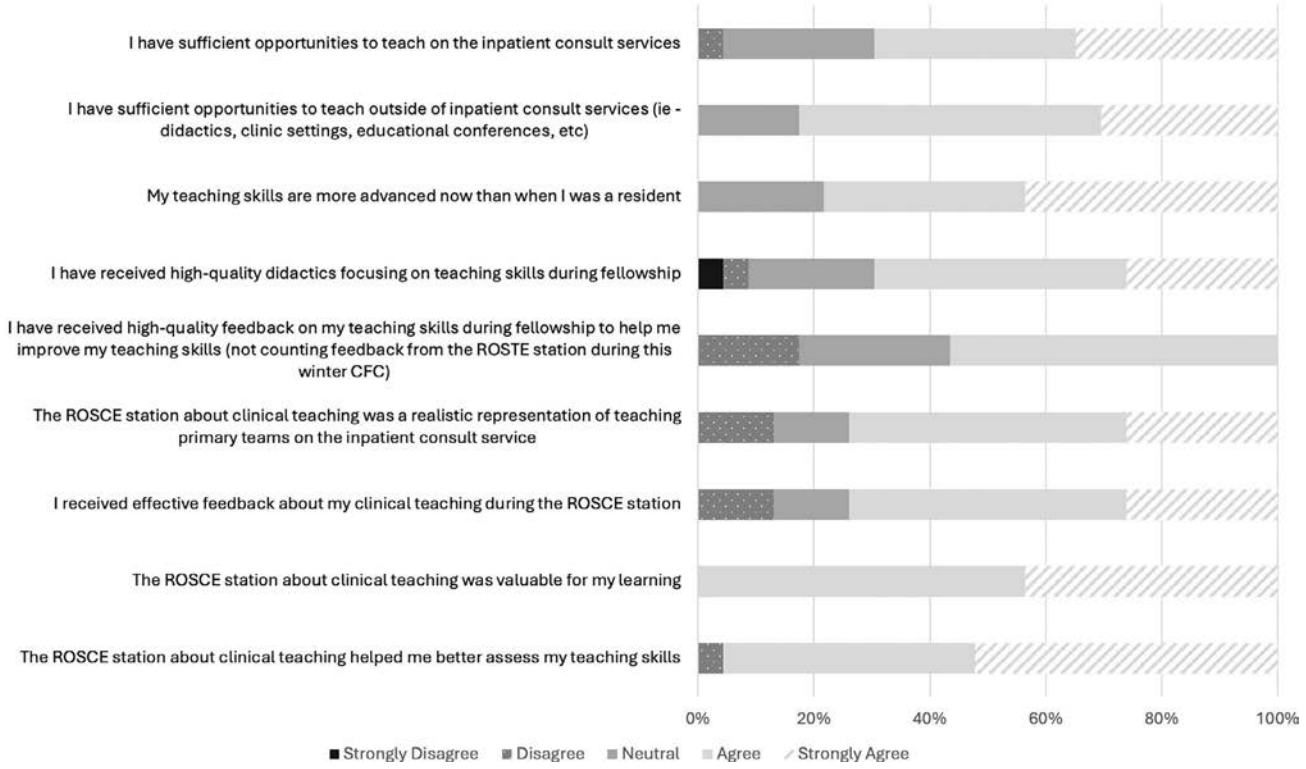

Figure 2. Fellow responses (n = 23) on the post-OSTE survey. Note, one fellow did not respond to the question about receiving high-quality feedback on teaching during fellowship. Their response was marked as "neutral." In the questions for this figure, the OSTE is also referred to as a ROSCE. CFC, Carolina Fellows Collaborative; OSTE, observed structured teaching exercise; ROSCE, rheumatology objective structured clinical examination; ROSTE, rheumatology observed structured teaching exercise.

as a teacher. Increased demands on faculty time have resulted in less engagement in teaching, and underrecognizing strengths in this domain may further decrease engagement in the setting of many competing responsibilities. Deliberate practice with expert feedback is important not only for skill development but also to help fellows more effectively self-assess.[14]

Our study has several limitations. We used a single OSTE station to assess teaching skills. Additional observations of the same domains would increase the certainty of our findings; however, this was not feasible in our study because fellows were also required to rotate through other observed stations assessing other clinical skills. The OSTE station focused on a straightforward case of gout to emphasize fellows' teaching skills, rather than content knowledge; however, disentangling these two competencies is challenging. We only assessed teaching during inpatient consultation; thus, our findings may not be generalizable to teaching skills in other venues, such as ambulatory teaching or lecture or small group–based didactics, although the domains investigated in our study are equally important for effective teaching in other settings. Although only a single rater scored each OSTE, both raters calibrated their scoring in prestudy training, and there was a strong positive correlation between raters. Additional potential sources of bias could have been introduced by the experimental setting or the standardized learner, but no differences in ratings were seen between stations. The strengths of the study included the relatively large number of fellows from five different academic institutions across multiple years of training, the use of a directly observed OSTE with trained raters to assess fellow teaching skills, timely and actionable feedback to the fellow learners on teaching skills, timely administration of the survey, and high survey completion rate.

An important future direction is to observe and provide feedback to learners about additional components of clinical teaching. The novel ACGME Clinician Educator Milestones provide a road map for assessing clinical teachers according to 5 competencies and 20 subcompetencies.[15] With these milestones in mind, our OSTE primarily assessed educational theory and practice. Future studies could use these Clinician Educator Milestones to develop more comprehensive assessments and interventions.

In conclusion, our study demonstrates there is an opportunity to enhance fellow teaching skills across multiple domains, especially regarding learner assessment and fellow self-assessment of their own teaching skills. Interventions such as our structured OSTE can improve fellows' skills as teachers, enhancing their impact on learners and other nonrheumatologists during clinical encounters.

AUTHOR CONTRIBUTIONS

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

REFERENCES

1. Serling-Boyd N, Miloslavsky EM. Enhancing the inpatient consultation learning environment to optimize teaching and learning. Rheum Dis Clin North Am 2020;46(1):73–83.
2. Miloslavsky EM, Bolster MB. Addressing the rheumatology workforce shortage: a multifaceted approach. Semin Arthritis Rheum 2020; 50(4):791–796.
3. Hill AG, Yu TC, Barrow M, et al. A systematic review of resident-as-teacher programmes. Med Educ 2009;43(12):1129–1140.
4. Sandars J. The use of reflection in medical education: AMEE Guide No. 44. Med Teach 2009;31(8):685–695.
5. Reid P, Miloslavsky EM, Dua A. Teacher training for rheumatology fellows: a national needs assessment of fellows and program directors. Clin Rheumatol 2020;39(3):673–680.
6. Miloslavsky EM, Degnan K, McNeill J, et al. Use of Fellow as Clinical Teacher (FACT) curriculum for teaching during consultation: effect on subspecialty fellow teaching skills. J Grad Med Educ 2017;9(3): 345–350.
7. Miloslavsky EM, McSparron JI, Richards JB, et al. Teaching during consultation: factors affecting the resident-fellow teaching interaction. Med Educ 2015;49(7):717–730.
8. McSparron JI, Huang GC, Miloslavsky EM. Developing internal medicine subspecialty fellows' teaching skills: a needs assessment. BMC Med Educ 2018;18(1):221–226.
9. Miloslavsky EM, Criscione-Schreiber LG, Jonas BL, et al. Fellow as teacher curriculum: improving rheumatology fellows' teaching skills during inpatient consultation. Arthritis Care Res (Hoboken) 2016; 68(6):877–881.
10. Adamson R, Goodman RB, Kritek P, et al; The Clinician-Educator Track of the University of Washington Pulmonary and Critical Care Medicine Fellowship Program. Training the teachers. The clinician-educator track of the University of Washington pulmonary and critical care medicine fellowship program. Ann Am Thorac Soc 2015;12(4): 480–485.
11. Richards JB, McCallister JW, Lenz PH. Pulmonary and critical care medicine program directors' attitudes toward training in medical education. A nationwide survey study. Ann Am Thorac Soc 2016;13(4): 475–480.
12. Pavitt S, Bogetz A, Blankenburg R. What makes the "perfect" inpatient consultation? A qualitative analysis of resident and fellow perspectives. Acad Med 2020;95(1):104–110.
13. Hodges B, Regehr G, Martin D. Difficulties in recognizing one's own incompetence: novice physicians who are unskilled and unaware of it. Acad Med 2001;76(10):S87–S89.
14. Ericsson KA. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. Acad Med 2004;79(10):S70–S81.
15. Mahan JD, Kaczmarczyk JM, Miller Juve AK, et al. Clinician Educator Milestones: assessing and improving educators' skills. Acad Med 2024;99(6):592–598.

Arthritis Care & Research
Vol. 77, No. 12, December 2025, pp 1459–1465
DOI 10.1002/acr.25600

American College of Rheumatology
Empowering Rheumatology Professionals

Prevalence and Incidence of Sjögren's Disease in Alaska Native and American Indian Peoples of Alaska

Tessalyn Morrison, MD, MPH, Peter Holck, PhD, MPH, Tammy L. Choromanski, MPH, Amy Wilson, PhD, RN, Flora Lee, RN, and Elizabeth D. Ferucci, MD, MPH

Objective. Our objective was to determine the prevalence, incidence, and clinical characteristics of Sjögren's disease (SjD) in Alaska Native and American Indian (AN/AI) peoples of Alaska.

Methods. We identified adults with SjD by querying electronic health records from participating tribal health organizations within the Alaska Tribal Health System (ATHS). Medical records were abstracted for adults with diagnostic codes for SjD. Individuals were included if they were diagnosed with SjD by a rheumatologist. Prevalence and incidence were calculated using the ATHS user population in 2019 (point prevalence) and from 2012 to 2019 (incidence), with direct age adjustment to the 2000 standard US population. We evaluated whether adults met modified criteria (positive Ro/SSA antigen with sicca symptoms), 2016 American College of Rheumatology (ACR)/EULAR, and 2012 ACR criteria.

Results. The age-adjusted prevalence of SjD was 199 per 100,000 adults (95% confidence interval [CI] 170–231); for primary SjD, it was 129 (106, 155), and for secondary SjD, it was 70 (95% CI 54–91). The age-adjusted incidence over the period was 16.6 (95% CI 13.7–20.0) per 100,000 person-years. Two-thirds (66%) of adults met modified criteria. Only 5% had a salivary gland biopsy performed, and only 3% met the 2016 ACR/EULAR or 2012 ACR criteria. The most common associated conditions in secondary SjD were rheumatoid arthritis and systemic lupus erythematosus.

Conclusion. The prevalence and incidence of SjD in AN/AI peoples is higher than other populations. These results may help clinicians to identify and treat this condition.

INTRODUCTION

Sjögren's disease (SjD) is an autoimmune disease characterized by reduced function of salivary and lacrimal glands that can be accompanied by extraglandular manifestations.[1] According to the most recent 2016 American College of Rheumatology (ACR)/EULAR classification criteria, primary SjD is diagnosed based on labial salivary gland biopsy showing lymphocytic sialadenitis, Ro/SSA antigen positivity, positive ocular staining, positive Schirmer's test, or reduced salivary flow, with the exclusion of a predisposing condition or other autoimmune disease that would designate disease as secondary SjD.[1] Compared with the 2012 ACR/Sjögren's International Collaborative Clinical Alliance Cohort (SICCA) and the 2002 American European Consensus Criteria classification criteria, the ACR/EULAR 2016 criteria added objective salivary flow testing and excluded subjective sicca symptoms, as well as La/SSB, rheumatoid factor (RF), and antinuclear antibody (ANA) antibody testing.[2,3] Over recent years, the field of rheumatology has moved toward labeling SjD as a disease rather than a syndrome, given its characteristic pathogenicity with correlating objective testing and the ongoing development of targeted treatments.[4] Rheumatologists have also moved away from defining SjD as primary or secondary, given the similar pathophysiology of the two and to recognize the full spectrum of the disease for future research and treatments.[5–7]

Owing to significant changes in classification criteria in recent years, estimating the current prevalence of SjD is challenging and most studies focus on primary SjD.[5,8,9] The prevalence of SjD in Indigenous populations has not been well described, although it is potentially higher than other populations based on previous studies documenting increased rates among Indigenous peoples of rheumatoid arthritis (RA), systemic lupus erythematosus (SLE), and spondyloarthropathies.[10] Estimating the prevalence of autoimmune disease in Indigenous communities, and among

The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

The authors were supported by the National Institute on Minority Health and Health Disparities, NIH, (grant R01-MD-014664) and the National Institute of General Medical Sciences, NIH, (grant P20GM152302).

[1]Alaska Native Tribal Health Consortium, Anchorage.

Author disclosures and graphical abstract are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25600.

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

Submitted for publication January 23, 2025; accepted in revised form June 23, 2025.

SIGNIFICANCE & INNOVATIONS

- This is the first study that describes the prevalence and incidence of Sjögren's disease (SjD) in Alaska Native/American Indian peoples.
- The prevalence and incidence of SjD is higher in Alaska Native and American Indian peoples than other populations previously studied.

persons living in rural areas in general, can be challenging due to limited access and lack of relative representation of Indigenous peoples. Our research on the epidemiology of autoimmune conditions was conducted within the Alaska Tribal Health System (ATHS), a unique comprehensive health system that provides specialty care for Alaska Native and American Indian (AN/AI) peoples in Alaska.

The prevalence and incidence of SjD in Indigenous North American populations has not been previously described. The primary objective of this study was to describe the prevalence and incidence of SjD among AN/AI peoples in Alaska. Secondary objectives were to describe the subtypes (primary vs secondary), clinical characteristics, classification criteria elements met, and extraglandular manifestations of SjD in this population. We hypothesized that the prevalence of SjD would be higher among this population than reported in other population-based studies and that objective ophthalmic and salivary gland testing would be limited.

MATERIALS AND METHODS

Study population. This study included adults receiving care within the ATHS, a statewide network of tribal health organizations operating self-governed health programs under a compact agreement with the United States Indian Health Service. Rheumatology health services are available at the Alaska Native Medical Center (ANMC), a tertiary health center in Anchorage, Alaska. People with rheumatologic disease within the ATHS are referred to ANMC or one of 11 regional field clinics conducted by rheumatologists from ANMC.

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

Case definition. Adults with potential SjD were identified from electronic health record encounter data in the ATHS between January 1, 2012, and December 31, 2019, based on *International Classification of Diseases, Ninth Version* code 710.2 and *International Classification of Diseases, Tenth Version* codes M35.0, M35.00, M35.01, M35.02, M35.03, M35.04, and M35.09. Adults were included if the diagnosis of SjD was endorsed by a rheumatologist and if the diagnosis was confirmed by December 31, 2019. Only individuals alive on December 31, 2019, contributed to the prevalence calculations; however, all adults with an SjD diagnosis during the entire study period were included in descriptive analysis of clinical characteristics regardless of date of death. The population denominator used for prevalence calculations was the 2019 ATHS adult (≥18 years of age) user population, defined as the number of adults receiving medical or dental care within the ATHS once during the previous three fiscal years. Similarly, ATHS user populations for the years 2012 through 2019 were used in incidence calculations. We used the ATHS user population to estimate the prevalence of SjD in AN/AI people in Alaska because the numerator was only drawn from people accessing the ATHS.

We defined the condition as primary SjD if there was no associated autoimmune condition suspected to be related to the SjD diagnosis, including RA, SLE, myositis, or mixed connective tissue disease; otherwise, the condition was defined as secondary SjD. All secondary SjD and associated diagnoses were recorded if confirmed by a rheumatologist. We used modified criteria set forth by Izmirly et al for population-based epidemiologic studies (SjD diagnosed by a medical provider with dry eyes and/or dry mouth and positive Ro/SSA).[9] We also evaluated adults for classification based on 2016 ACR/EULAR criteria, 2012 ACR/SICCA criteria, and 2002 American European Consensus Criteria criteria.[1–3]

Data collection. Data were obtained from the shared Cerner electronic health record platform maintained by ANMC or by using local electronic health record platforms for those tribal health organizations who do not participate in the shared platform. Data were compiled from multiple sources into REDCap hosted on a secure ANTHC server for analysis. Research nurses trained by the principal nvestigator performed medical record abstraction to confirm the diagnosis of SjD and determine associated features. The medical record abstraction followed a standardized data dictionary. Abstracted data elements included demographics, diagnoses, pathology reports, laboratory studies, clinical characteristics, and associated rheumatologic conditions reported in clinician notes. Data were reviewed by two research physician co-authors (THM and EDF) for quality assurance with respect to diagnosis of SjD, primary or secondary, and select clinical characteristics.

Statistical analysis. Analysis was performed using Excel (version 2405) and R (version 4.4.1). Unadjusted point prevalence

was calculated as the number of adults with the condition as of December 31, 2019, divided by the corresponding ATHS user population in 2019. Incidence was calculated as the number of adults diagnosed with the condition between 2012 and 2019 divided by the ATHS person-years at risk of developing the condition observed during the study period. Direct age-adjusted estimates of prevalence and incidence were also calculated, adjusting to the 2000 standard US population, recommended as the current standard for age adjustment in the United States.[11,12] Confidence intervals (95% CIs) were calculated using exact binomial tests for unadjusted prevalence. For age-adjusted estimates of prevalence and incidence (direct standardized rates), weighted sums of Poisson variables were assumed, and the mid-p γ-method of Fay and Kim[13] was used to estimate 95% CIs.

RESULTS

We identified 177 adults with SjD diagnosed by a rheumatologist (Figure 1). Among the 177 adults, 116 (66%) met modified criteria (SjD diagnosed by a medical provider with dry eyes and/or dry mouth with positive Ro/SSA).

The age-adjusted prevalence of SjD in this population was 199 per 100,000 persons (95% CI 170–231), whereas primary SjD age-adjusted prevalence was 129 per 100,000 (95% CI 106–155), and secondary SjD age-adjusted prevalence was 70 (95% CI 54–91; Table 1). Among female individuals, the age-adjusted prevalence of SjD was 335 per 100,000 persons (95% CI 285–393), whereas among male individuals, it was significantly lower at 46 per 100,000 (95% CI 27–72). The unadjusted prevalence of SjD was eight times higher in female individuals than male individuals. Using modified criteria, the overall prevalence was 127 per 100,000 (95% CI 105–153) and for the subset with primary SjD who meet modified criteria it was 91 per 100,000 (95% CI 72–114). The overall incidence was 15.1 per 100,000 person-years (95% CI 12.4–18.1), and the age-adjusted incidence of SjD was 16.6 per 100,000 person-years (95% CI 13.7–20.0).

We described the classification characteristics of adults with SjD diagnosed by a rheumatologist and those who met modified criteria (Table 2). Dry eyes and dry mouth were documented in 93% and 86% of adults diagnosed by a rheumatologist, respectively. Few people underwent objective ophthalmic and salivary gland testing: 10% had documentation of ocular staining scoring, 5% had a salivary gland biopsy (3% had biopsy results available), 2% had Schirmer testing (1% had testing results available), and no one underwent salivary flow testing. Of those who underwent salivary gland biopsy with available results, 80% had a characteristic positive test. Most adults had serology testing, and 81% had a positive Ro/SSA, 50% had a positive La/SSB antigen (La/SSB), 88% had a positive ANA (71% had positive ANA ≥ 1:320), and 73% had a positive RF (66% had an RF that was three times the upper limit of normal). Results are also listed for those diagnosed by modified criteria and are similar, noting that positive Ro/SSA was a criterion. Few adults met classification criteria for primary SjD: three people met 2002 ACR and 2016 ACR/EULAR criteria and six people met 2012 ACR SICCA criteria.[1–3]

We also summarize clinical features of adults and compare individuals with primary and secondary SjD (Table 3). Parotid swelling was observed in 13% of adults with SjD. Low complement 3 and low complement 4 was identified in 11% and 34% of people with SjD, respectively. Arthritis (46%) and lymphopenia (21%) were the most common extraglandular manifestations. The most common associated conditions in adults with secondary SjD were RA (83%) and SLE (23%). Osteoarthritis (46%) was the most common associated rheumatologic condition of adults with SjD.

DISCUSSION

In this study, we describe the prevalence and incidence of SjD in AN/AI adults in Alaska. The age-adjusted prevalence of SjD in 2019 was 199 per 100,000 (95% CI 170–231) and the age-adjusted incidence of SjD was 16.6 per 100,000 person-years (95% CI 13.7–20.0) for the period between 2012 to 2019. Two-thirds of cases (66%) were diagnosed as primary SjD. Most people were diagnosed clinically because very few adults underwent objective ophthalmic and salivary gland testing (only three people met 2002 ACR and 2016 ACR/EULAR criteria and six people met 2012 ACR SICCA criteria). SLE and RA were the most common conditions associated with diagnosed secondary SjD.

This is the first study describing prevalence and incidence of SjD in AN/AI peoples of Alaska and one of few studies describing SjD epidemiology in Indigenous North American populations. In contrast to most epidemiologic studies and reflecting current recommendations for describing SjD, we primarily reported an aggregate category of SjD (primary and secondary combined), although we did analyze persons with primary and secondary SjD separately for the purpose of comparing data with other studies. Compared with other estimates of SjD prevalence, we found that the prevalence of overall SjD, primary SjD, and secondary SjD, as well as the incidence of overall SjD, were all higher in AN/AI peoples than reported elsewhere in other populations.

Comparing the prevalence of SjD between studies is challenging due to differing methodologies and adjustments. One study from Spain measured the prevalence of physician-diagnosed SjD, finding a lower rate of SjD than observed in this study (crude prevalence of 84 vs 177 per 100,000).[14] Most studies focus on the prevalence of primary SjD. A recent 2024 review estimates a prevalence of 10 to 100 per 100,000 for primary SjD based on best evidence.[5] A 2015 systematic review of seven population studies found a prevalence of 43 per 100,000.[8] Studies since 2015 have estimated a lower prevalence of primary SjD by physician diagnosis or other algorithm compared with this study (7–55 vs 115 per

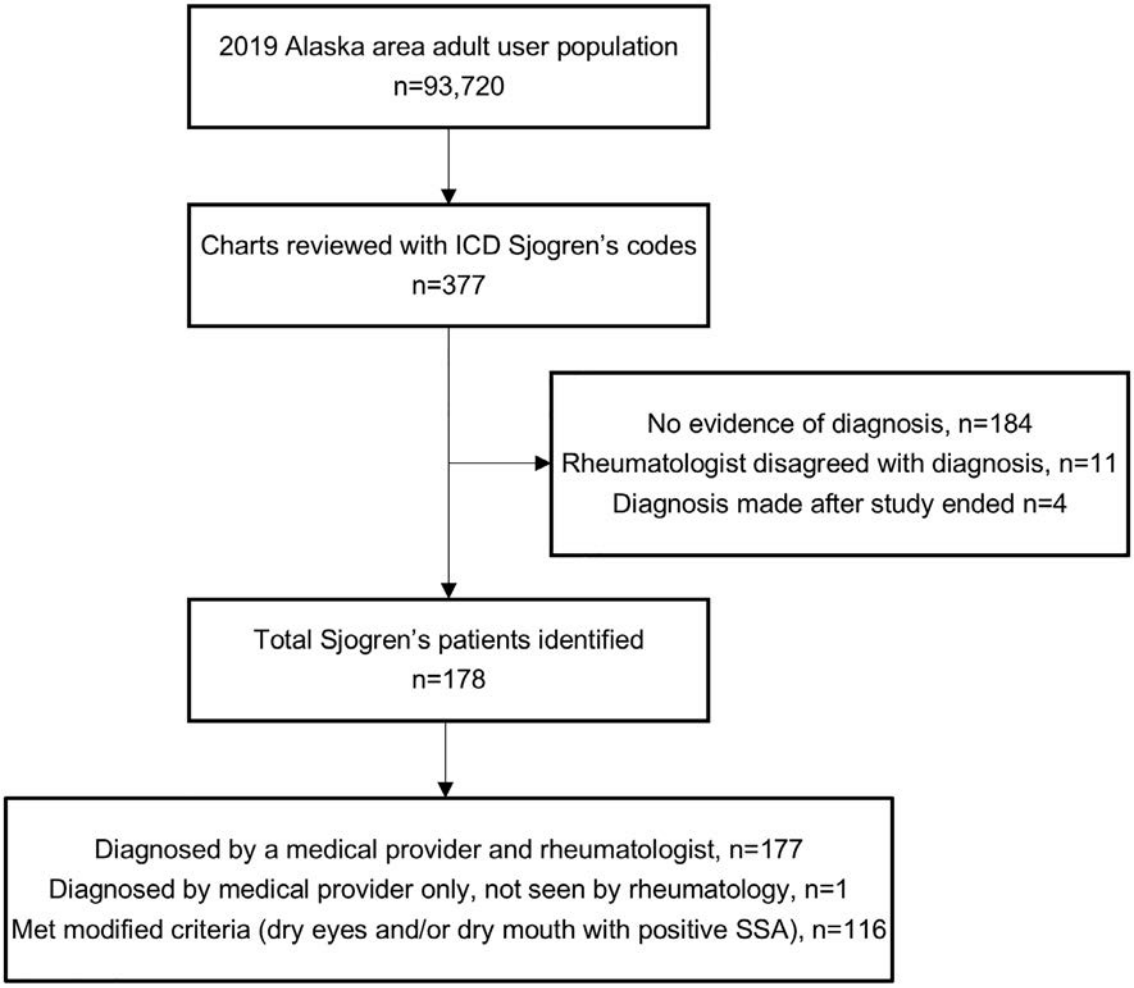

Figure 1. Flow diagram of included participants. ICD, International Classification of Diseases.

100,000).[9,14–17] Our study was most similar in methodology to that of Izmirly et al, in which the authors reported a lower age-adjusted prevalence of primary SjD diagnosed by rheumatologist (7.1 compared with 129 per 100,000) and of primary SjD diagnosed by modified criteria (1.1 vs 91 per 100,000).[9]

Our finding that two-thirds of adults had primary SjD as opposed to secondary SjD was similar to results of other studies that have reported both definitions (65%–72%).[14,18] Two studies, one from Spain and one from France, examined the prevalence of secondary SjD, and both report lower rates of secondary SjD compared with this study (prevalence of 16–28 vs 62 per 100,000).[14,16] Other studies have also shown that secondary SjD is most often associated with RA and SLE; however, our study found a higher prevalence of RA in adults with secondary SjD (83% versus 53%–55% in other studies).[14,16,18] Other studies have similarly shown SjD to be more prevalent among women, with rates ranging from 7 to 18 times those of men.[9,14,16,19,20] There were no other studies that evaluated incidence of aggregated SjD; however, we suspect that the incidence we found in this study is relatively high, given other estimates for primary SjD

Table 1. Prevalence of Sjögren's disease in AN/AI peoples

Characteristics	n	Crude prevalence (95% CI)	Age-adjusted prevalence (95% CI)
Overall	166	177 (151–206)	199 (170–231)
Primary	108	115 (95–139)	129 (106–155)
Secondary	58	62 (47–80)	70 (54–91)
Female	149	306 (259–359)	335 (285–393)
Male	17	38 (22–60)	46 (27–72)
Meets modified criteria	108	115 (95–139)	127 (105–153)
Primary	77	82 (65–203)	91 (72–114)

AN/AI, Alaska Native and American Indian; CI, confidence interval.

Table 2. Classification characteristics of AN/AI adults with Sjögren's disease*

Characteristics	Diagnosed by rheumatologist (n = 177)		Diagnosed by modified criteria (n = 116)	
	Tested, %	Positive, %	Tested, %	Positive, %
Symptoms				
Dry eyes	100	93	100	97
Dry mouth	100	86	100	92
Objective testing				
Salivary gland biopsy	5	–	6	–
Lymphocytic sialadenitis with a focus score of ≥1 foci/4 mm^2	3	80	3	75
Ocular staining score ≥3	10	39	12	29
Serologies				
Ro/SSA	84	81	100	100
La/SSB	80	50	92	63
ANA	90	88	97	92
ANA titer ≥ 1:320	67	71	78	74
Rheumatoid factor	79	73	84	70
Rheumatoid factor 3× ULN	49	66	48	57

* Percentage tested among those diagnosed by a rheumatologist or by modified criteria, respectively, and percentage found positive among those tested within each category. Fewer than five were those with Schirmer testing.
ANA, antinuclear antibody; AN/AI, Alaska Native and American Indian; ULN, upper limit of normal.

of 0.7 to 6.9 per 100,000 person-years and for secondary SjD of 0.3 to 2.0 per 100,000 person-years.[5,8,9,16,21,22]

There are a limited number of studies examining SjD prevalence in rural and Indigenous communities; however, sample sizes in these studies are too small to permit conclusive comparisons of rates to our study results. Rates reported include a 0.6% prevalence of primary SjD in rural Greece, 0.2% prevalence in a Wichi community of Argentina, and a 0.1% prevalence in a Qom

Table 3. Clinical characteristics of AN/AI adults with Sjögren's disease*

Characteristics	Diagnosed by rheumatologist (n = 177)		Primary (n = 114), %	Secondary (n = 64), %
	n	%		
Glandular symptoms				
Dry eyes	165	93	93	94
Dry mouth	152	86	88	83
Parotid swelling	23	13	13	13
Serologies[a]				
Ro/SSA	149	81	87	69
La/SSB	141	50	54	41
ANA positive	159	88	89	86
ANA titer >1:320	119	71	67	82
Rheumatoid factor positive	140	73	63	89
Rheumatoid factor 3× ULN	87	66	58	74
Low C3	7	11	7	20
Low C4	22	34	31	40
Extraglandular manifestations				
Arthritis	82	46	47	45
Lymphopenia	38	21	23	19
Peripheral neuropathy	20	11	12	9
Raynaud phenomenon	18	10	7	16
Photosensitivity	17	10	11	8
Interstitial lung disease	11	6	3	13
Renal tubular acidosis	9	5	6	3
Associated conditions				
Osteoarthritis	81	46	57	27
Rheumatoid arthritis	53	30	0	83
Systemic lupus erythematosus	15	8	0	23
Other conditions				
Cancer diagnosis	23	13	14	11

* Fewer than five have cutaneous vasculitis, interstitial nephritis, vasculitis, scleroderma, myositis, or a history of having an infant with a diagnosis of neonatal lupus or congenital heart block.
ANA, antinuclear antibody; AN/AI, Alaska Native and American Indian; ULN, upper limit of normal.
[a] Percentage listed as positive result among those tested.

community of Argentina.[19,23,24] A cross-sectional study of a diverse sample of adults with primary SjD from the US Midwest found American Indian (AI) people were overrepresented in their primary SjD cohort, suggesting that primary SjD may be more common among AI people.[25] These data found that compared with European Americans, AI people had lower rates of reduced tear and salivary flow and less antibody positivity, as well as higher levels of disease activity and more extraglandular manifestations.[25] When comparing our results to those of a large international study, we observed higher rates of antibody positivity in AN/AI peoples. In our study, 87% had Ro/SSA positivity compared with 75% internationally among adults with primary SjD, with similar differences among La/SSB and RF.[20] Conversely, we observed similar rates and types of extraglandular manifestations to results reported in other studies,[9,26] which found the most common extraglandular manifestation of primary SjD to be arthritis and lymphopenia.

The main strength of this study is that we were able to describe prevalence in a diverse population that included individuals from rural and remote areas, all receiving rheumatology care within one health care system. All adults with SjD in this study were evaluated by rheumatologists who confirmed a diagnosis. There were several limitations encountered in our study. First, we had limited ability to apply classification criteria due to a lack of objective ophthalmic and salivary gland testing, which is a limitation faced by most population-based studies. Although we were unable to find data from other studies with which to compare the rates of objective eye and salivary testing used in clinical practice, the study from which we adopted the modified criteria used these criteria because of limited objective eye and salivary gland testing data used for classification criteria.[9] We did not evaluate reasons why objective testing was absent in each case and as such we cannot assume lack of access was the cause. We suspect that clinicians may have opted not to use objective testing as it exposes patients to discomfort without significantly changing the treatment plan. Second, our study's reliance on clinical and serologic diagnosis may overestimate the prevalence of histologically evident SjD. Conversely, we may have missed people with SjD who did not receive a diagnosis by a health care provider. Third, our query only identified individuals who access health care in the ATHS for this condition during the period of our study. By using the ATHS user population as our denominator and identifying individuals with SjD from this population, we could miss AN/AI people in Alaska who either do not access ATHS for any of their health care or who are healthy and do not access health care services within the three-year period of user population calculation. We estimate that approximately 95% of the AN/AI population living in Alaska is represented within the ATHS user population (unpublished data). Fourth, our study used electronic health record data and was limited by lack of available data from outside institutions or early paper records, especially for individuals diagnosed years before 2012. We used *International Classification of Diseases* codes related to SjD to identify those adults with SjD, which may not have captured all cases of secondary SjD if they were coded as the primary rheumatologic disease alone. Finally, we did not collect data related to disease activity and so were unable to compare our data to previous studies citing higher disease activity rates in AI people.[25]

Our research offers a methodology for evaluating the population-based prevalence of SjD in a health care system with a population dispersed over a vast geographic area. These findings will be useful in guiding clinicians and health care systems delivering care to AN/AI populations. We hope that future research will describe the prevalence of SjD in different rural and Indigenous populations. Given the inability to apply current SjD classification criteria, it could be useful to develop modified criteria or standardized probability metrics based on characteristic features to assess for the presence of SjD where objective testing is impractical. Future criteria considering new diagnostic methods should also consider the feasibility of applying these approaches in rural or underserved populations. Like SjD, many other rheumatologic diseases are more prevalent in AN/AI peoples, and more research is needed to understand the etiology of the increased prevalence of these diseases, as well as evaluating for protective factors that may exist.

ACKNOWLEDGMENTS

We would like to thank the Alaska Native Tribal Health Organizations statewide for their time reviewing our study and manuscripts. Specifically, we thank the following organizations who approved and assisted with local electronic health record data collection: Yukon Kuskokwim Health Corporation, Tanana Chiefs Conference, Maniilaq Association, and Annette Island Service Unit. In addition, we thank the following organizations who approved and allowed us to access electronic health record data for their region within the shared electronic health record platform: Alaska Native Tribal Health Consortium, Southcentral Foundation, Norton Sound Health Corporation, Southeast Alaska Regional Health Consortium, Bristol Bay Area Health Corporation, and Kodiak Area Native Association.

AUTHOR CONTRIBUTIONS

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

REFERENCES

1. Shiboski CH, Shiboski SC, Seror R, et al. 2016 American College of Rheumatology/European League Against Rheumatism classification criteria for primary Sjögren's syndrome: a consensus and data-driven

methodology involving three international patient cohorts. Arthritis Rheumatol 2017;69(1):35–45.

2. Shiboski SC, Shiboski CH, Criswell LA, et al. American College of Rheumatology classification criteria for Sjögren's syndrome: a data-driven, expert consensus approach in the Sjogren's International Collaborative Clinical Alliance cohort. Arthritis Care Res (Hoboken) 2012;64(4):475–487.
3. Vitali C, Bombardieri S, Jonsson R, et al. Classification criteria for Sjögren's syndrome: a revised version of the European criteria proposed by the American-European Consensus Group. Ann Rheum Dis 2002;61(6):554–558.
4. Baer AN, Hammitt KM. Sjögren's disease, not syndrome. Arthritis Rheumatol 2021;73(7):1347–1348.
5. Beydon M, McCoy S, Nguyen Y, et al. Epidemiology of Sjögren syndrome. Nat Rev Rheumatol 2024;20(3):158–169.
6. Kollert F, Fisher BA. Equal rights in autoimmunity: is Sjögren's syndrome ever 'secondary'? Rheumatology (Oxford) 2020;59(6):1218–1225.
7. Mavragani CP, Moutsopoulos HM. Primary versus secondary Sjögren syndrome: is it time to reconsider these terms? J Rheumatol 2019; 46(7):665–666.
8. Qin B, Wang J, Yang Z, et al. Epidemiology of primary Sjögren's syndrome: a systematic review and meta-analysis. Ann Rheum Dis 2015;74(11):1983–1989.
9. Izmirly PM, Buyon JP, Wan I, et al. The incidence and prevalence of adult primary Sjögren's syndrome in New York county. Arthritis Care Res (Hoboken) 2019;71(7):949–960.
10. Ferucci ED. Understanding the disproportionate burden of rheumatic diseases in Indigenous North American populations. Rheum Dis Clin North Am 2020;46(4):651–660.
11. Anderson RN, Rosenberg HM. Report of the Second Workshop on Age Adjustment. Vol. 4. Vital and Health Statistics, US Dept of Health and Human Services; 1998.
12. Klein RJ, Schoenborn CA. Age adjustment using the 2000 projected U.S. population. Healthy People 2000 Stat Notes 2001(20):1–9.
13. Fay MP, Kim S. Confidence intervals for directly standardized rates using mid-p gamma intervals. Biom J 2017;59(2):377–387.
14. Barrio-Cortes J, López-Rodríguez JA, Gómez-Gascón T, et al. Prevalence and comorbidities of Sjögren's syndrome patients in the community of Madrid: a population-based cross-sectional study. Joint Bone Spine 2023;90(4):105544.
15. Maciel G, Crowson CS, Matteson EL, et al. Prevalence of primary Sjögren's syndrome in a US population-based cohort. Arthritis Care Res (Hoboken) 2017;69(10):1612–1616.
16. Seror R, Chiche L, Beydon M, et al. Estimated prevalence, incidence and healthcare costs of Sjögren's syndrome in France: a national claims-based study. RMD Open 2024;10(1):e003591.
17. Albrecht K, Binder S, Minden K, et al. Systematic review to estimate the prevalence of inflammatory rheumatic diseases in Germany. Z Rheumatol 2024;83(Suppl 1):20–30.
18. Dézsi AJ, Erdei C, Demeter T, et al. Prevalence of Sjögren's syndrome in patients with dry mouth in the region of Central Hungary. Oral Dis 2023;29(7):2756–2764.
19. Juárez V, Quintana R, Crespo ME, et al. Prevalence of musculoskeletal disorders and rheumatic diseases in an Argentinean indigenous Wichi community. Clin Rheumatol 2021;40(1):75–83.
20. Brito-Zerón P, Acar-Denizli N, Zeher M, et al. Influence of geolocation and ethnicity on the phenotypic expression of primary Sjögren's syndrome at diagnosis in 8310 patients: a cross-sectional study from the Big Data Sjögren Project Consortium. Ann Rheum Dis 2017; 76(6):1042–1050.
21. Maciel G, Crowson CS, Matteson EL, et al. Incidence and mortality of physician-diagnosed primary Sjögren syndrome: time trends over a 40-year period in a population-based US cohort. Mayo Clin Proc 2017;92(5):734–743.
22. Kim HJ, Kim KH, Hann HJ, et al. Incidence, mortality, and causes of death in physician-diagnosed primary Sjögren's syndrome in Korea: a nationwide, population-based study. Semin Arthritis Rheum 2017; 47(2):222–227.
23. Dafni UG, Tzioufas AG, Staikos P, et al. Prevalence of Sjögren's syndrome in a closed rural community. Ann Rheum Dis 1997;56(9): 521–525.
24. Quintana R, Silvestre AMR, Goñi M, et al. Prevalence of musculoskeletal disorders and rheumatic diseases in the indigenous Qom population of Rosario, Argentina. Clin Rheumatol 2016;35(Suppl 1):5–14.
25. Scofield RH, Sharma R, Pezant N, et al. American Indians have a higher risk of Sjögren's syndrome and more disease activity than Caucasians and African-Americans. Arthritis Care Res (Hoboken) 2020; 72(8):1049–1056.
26. Maldini C, Seror R, Fain O, et al. Epidemiology of primary Sjögren's syndrome in a French multiracial/multiethnic area. Arthritis Care Res (Hoboken) 2014;66(3):454–463.

Arthritis Care & Research
Vol. 77, No. 12, December 2025, pp 1466–1473
DOI 10.1002/acr.25595

American College of Rheumatology
Empowering Rheumatology Professionals

Cancer Risk in Patients With Systemic Sclerosis: A Nationwide Cohort Study in South Korea 2004 to 2021

Jihyun Na,[1] Su Jin Choi,[1] Sojeong Park,[2] Minae Park,[2] and Doo-Ho Lim[1]

Objective. Systemic sclerosis (SSc) is a rare autoimmune disease characterized by tissue fibrosis, vasculopathy, and immune dysregulation. Our objectives were to quantify the overall and site-specific cancer risks in patients with SSc compared to the general population, examine temporal trends in cancer incidence following SSc diagnosis, and explore potential associations with immunosuppressive agent use.

Methods. Using data from the Korean National Health Insurance Service, we identified 6,386 patients newly diagnosed with SSc between 2004 and 2020. Standardized incidence ratios (SIRs) were calculated to compare the risk of cancer between patients with SSc and the general population. Subgroup analyses were performed based on age at diagnosis, follow-up duration, and immunosuppressive agent use.

Results. Most patients (80.2%) were women, with a mean ± SD age at diagnosis of 53.1 ± 13.4 years. Patients with SSc had higher risks of overall cancer (SIR 3.12; 95% confidence interval [CI] 2.92–3.33), solid cancer (SIR 3.07; 95% CI 2.86–3.28), and hematologic cancer (SIR 5.70; 95% CI 4.50–7.11) compared to the general population. Myelodysplastic syndrome (SIR 12.52; 95% CI 7.00–20.65) had the highest risk, followed by multiple myeloma, eye, Hodgkin lymphoma, and larynx cancers. Cancer risk peaked among patients aged 20 to 39 years (SIR 5.27; 95% CI 4.31–6.38) and during the first year after diagnosis (SIR 4.44; 95% CI 3.81–5.14).

Conclusion. In this study, we revealed that the incidence of cancer is higher in patients with SSc in South Korea compared to the general population. The strong association between SSc and cancer occurrence prompts clinicians to conduct careful cancer screening for patients with SSc.

INTRODUCTION

Systemic sclerosis (SSc) is a rare autoimmune disease characterized by tissue fibrosis, vasculopathy, and immune dysregulation.[1] The disease is associated with significant morbidity and mortality, with a spectrum of clinical manifestations ranging from localized skin involvement to severe internal organ dysfunction. In recent years, evidence has emerged highlighting an increased risk of malignancy in patients with SSc, particularly cancers of the lung, esophagus, and breast.[1,2] This association underscores the need for heightened vigilance in cancer screening and prevention strategies in this population.

The relationship between SSc and cancer is complex and multifactorial. Autoimmune-mediated tissue damage, chronic inflammation, and the use of immunosuppressive therapies are all thought to contribute to carcinogenesis. Additionally, specific autoantibodies, such as anti–RNA polymerase III, have been implicated in a heightened cancer risk, particularly in the early phases of SSc diagnosis.[3] However, data on cancer incidence in patients with SSc from Asian populations, including South Korea, remain limited, and the interplay across environmental, genetic, and treatment-related factors is not fully understood.

In this study, we analyzed data from the Korean National Health Insurance Service (NHIS) to investigate the cancer risk among patients with SSc in South Korea. Our objectives were to quantify the overall and site-specific cancer risks in patients with SSc compared to the general population, examine temporal trends in cancer incidence following SSc diagnosis, and explore potential associations with immunosuppressive agent use. By elucidating these patterns, we aim to provide insights into optimizing cancer screening and surveillance strategies for patients with SSc.

[1]Jihyun Na, MD, Su Jin Choi, MD, Doo-Ho Lim, MD, PhD: Ulsan University Hospital, University of Ulsan College of Medicine, Ulsan, South Korea; [2]Sojeong Park, MPH, Minae Park, MPH: Hanmi Pharmaceutical, Co, Ltd, Seoul, South Korea.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25595).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25595.

Address correspondence via email to Doo-Ho Lim, MD, PhD, at dlaengh@hanmail.net.

Submitted for publication February 11, 2025; accepted in revised form June 23, 2025.

SIGNIFICANCE & INNOVATIONS

- The Korea National Cancer Registry provides data on cancer incidence across the entire population of Korea. The Korean National Health Insurance Service includes all patients diagnosed with systemic sclerosis (SSc), ensuring reliable data from a large-scale source.
- The study revealed higher overall standardized incidence ratios in patients with SSc compared to the general population in South Korea.
- Increased cancer risk was observed in the 20 to 39 years age group and during the first year after diagnosis of SSc.

MATERIALS AND METHODS

Data resource. To identify patients diagnosed with SSc in Korea, we used a customized database from Korean NHIS, a mandatory and nationwide health insurance system that covers approximately 97% of the entire Korean population.[4] The remaining 3% are covered by the Medical Aid Program. The NHIS database provides comprehensive information, including demographic characteristics, diagnostic codes, and drug prescriptions from both inpatient and outpatient care. *International Classification of Disease, 10th Revision* (*ICD-10*) codes were used to identify relevant diagnostic codes during the study period.

Cancer incidence data for the general population were obtained from the Statistics Korea (http://kosis.kr, Korean National Cancer Registry data for year 2019), which tracks annual cancer incidence by age, sex, and cancer type for the entire population of Korea.

Study population. We identified eligible patients diagnosed with SSc between January 1, 2002, and December 31, 2020, from the NHIS database. Patients with SSc were defined as those who had at least one claim with the diagnostic code for SSc (*ICD-10* code M34) and were registered in the Rare and Intractable Disease (RID) program for SSc (RID code V138). The RID program, managed by the NHIS, provides financial support to patients with certain RIDs. To be enrolled in the RID program for SSc, patients have to meet the 2013 American College of Rheumatology/EULAR collaborative initiative for classification of SSc.[5] Because the RID program is administered through the NHIS and is accessible across the country—including rural and underserved regions—the likelihood of regional disparities in enrollment is minimized. Patients must fulfill standardized diagnostic criteria verified by specialists, ensuring consistent case identification nationwide. We included patients aged 20 years or older and excluded those with a history of cancer. A two-year washout period (2002–2003) was also applied to exclude the patients who had been previously diagnosed with SSc. The final study population consisted of patients diagnosed with SSc between January 1, 2004, and December 31, 2020. The index date was defined as the first registration date of the SSc diagnostic code. Patients were observed from the index date until the occurrence of the cancer, death, or the end of the study (December 31, 2021).

The study was conducted in accordance with the principles of the Declaration of Helsinki. The institutional review board (IRB) of Ulsan University Hospital approved the study (Ulsan, Korea; IRB number: 2023-03-030) and waived the requirement for informed consent because this study used a public database and did not collect personally identifiable information.

Cancer case ascertainment and patient characteristics. Cancer cases were identified based on the first occurrence of relevant *ICD-10* diagnostic codes (M34.0, M34.1, M34.2, M34.8, M34.9) listed in the Korean National Cancer Registry data. Baseline patient characteristics, including age, sex, and Charlson comorbidity index (CCI) scores, were collected.[6] CCI scores were calculated using *ICD-10* codes for comorbidities documented within one year before the index date (Supplementary Table S1). Prescription data for immunosuppressive agents, including hydroxychloroquine, methotrexate, azathioprine, tacrolimus, mycophenolate mofetil, and cyclophosphamide, were also collected after the index date.

Statistical analysis. Continuous and categorical variables are presented as mean (±SD) and as number (%), respectively. To compare the cancer incidence between patients with SSc and the general population, we calculated the standardized incidence ratios (SIRs) for overall and site-specific cancers, using incidence data stratified by five-year age intervals and sex. SIRs were estimated by dividing the observed number of cancer cases in patients with SSc by the expected number of cancer cases. The expected number of cancer cases was calculated by multiplying the age- and sex-specific cancer incidence rates of the general population by the person-years of patients with systemic lupus erythematosus (SLE). The 95% confidence interval (CI) was calculated using Byar's approximation of the Poisson distribution. Subgroup analyses were conducted by cancer type, age at diagnosis, follow-up duration, and the use of immunosuppressive agents. Solid cancers were identified using *ICD-10* codes C00 through C80, and hematologic cancers included Hodgkin lymphoma (C81), non-Hodgkin lymphoma (C82–C86, C96), multiple myeloma (C90), leukemia (C91–C95), myelodysplastic syndrome (MDS; D46), and myeloproliferative disease (D47.1). SIRs were calculated for all types of hematologic cancers and 20 types of solid cancers. Age at diagnosis was categorized into 20 to 39 years, 40 to 59 years, 60 to 79 years, and ≥80 years. Follow-up duration was grouped into less than one year, one to two years, three to four years, five to six years, seven to eight years, and nine years or more. SAS Enterprise Guide version 9.4 (SAS Institute Inc) was used for all statistical analyses.

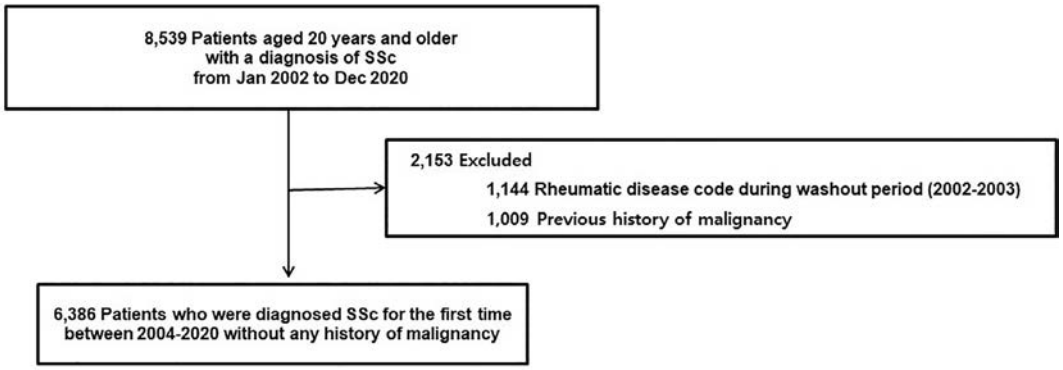

Figure 1. Flow diagram of study population. SSc, systemic sclerosis.

RESULTS

Baseline characteristics of patients with SSc. Among 8,539 patients with SSc identified from the NHIS database between 2002 and 2020, 1,144 patients were excluded due to the presence of an SLE code during the washout period, and 1,009 patients were excluded due to a history of cancer. Finally, a total of 6,386 patients were included for further analyses (Figure 1).

Table 1 summarizes the baseline characteristics of patients with SSc. Most patients (80.2%) were women, and the mean ± SD age at diagnosis was 53.1 ± 13.4 years. The mean + SD follow-up duration was 7.8 ± 5.1 years. Glucocorticoid was the most frequently prescribed immunosuppressant (93.7%). Specifically, 32.4% of all patients had been prescribed glucocorticoids for more than a year, followed by hydroxychloroquine (36.0%), methotrexate (22.9%), azathioprine (16.6%), and cyclophosphamide (10.9%). According to the CCI, 1,945 patients (30.5%) were classified as having rheumatologic disease, and 78.5% of the patients had a CCI score of ≥1 (Table 1 and Supplementary Table S2).

Risk of cancer incidence in patients with SSc compared to the general population. The observed number and SIRs of overall and site-specific cancers in patents with SSc are presented in Table 2. During the follow-up period (49,916.28 person-years), 910 patients with SSc developed cancer, with 833 solid cancer cases and 78 hematologic cancer cases. Lung cancer (n = 166) was the most common, followed by thyroid (n = 85), liver (n = 68), and breast cancers (n = 64). Among hematologic cancers, non-Hodgkin lymphoma was most common (n = 33), followed by multiple myeloma (n = 22), MDS (n = 15), and leukemia (n = 6).

The risk of overall cancer was significantly higher in patients with SSc compared to the general population (SIR 3.12; 95% CI 2.2–3.33). Elevated risks were also observed for both solid cancer (SIR 3.07; 95% CI 2.86–3.28) and hematologic cancer (SIR 5.70; 95% CI 4.50–7.11). Patients with SSc showed higher risks for most site-specific cancers, except for myeloproliferative disease. Among solid cancers, the organ with the highest risk of developing cancer compared to the general population was the eye (SIR 9.90; 95% CI 0.13–55.07), followed by the larynx, lung, prostate, and lip, oral cavity, and pharynx. Among hematologic cancers, MDS (SIR 12.52; 95% CI 7.00–20.65) showed the highest risk in patients with SSc, followed by multiple myeloma, Hodgkin lymphoma, and non-Hodgkin lymphoma.

Table 1. Clinical characteristics of study population*

Characteristics	Total (N = 6,386)
Age at diagnosis, mean ± SD, y	53.1 ± 13.4
Women, n (%)	5,123 (80.2)
Follow-up duration, mean ± SD, y	7.82 ± 5.06
Medication, n (%)	
Hydroxychloroquine	2,299 (36.0)
Methotrexate	1,463 (22.9)
Azathioprine	1,059 (16.6)
Mycophenolate	692 (10.8)
Cyclophosphamide	696 (10.9)
Glucocorticoids	5,981 (93.7)
Glucocorticoids ≥1 y	2,067 (32.4)
Charlson comorbidities, n (%)	
Myocardial infarction	93 (1.5)
Congestive heart failure	375 (5.9)
Peripheral vascular disease	1,511 (23.7)
Cerebrovascular disease	481 (7.5)
Dementia	94 (1.5)
Chronic pulmonary disease	2,127 (33.3)
Rheumatologic disease	1,945 (30.5)
Peptic ulcer disease	1,797 (28.1)
Mild liver disease	1,797 (28.1)
Diabetes without chronic complication	776 (12.2)
Diabetes with chronic complications	487 (7.6)
Hemiplegia or paraplegia	70 (1.1)
Renal disease	125 (2.0)
Moderate to severe liver disease	35 (0.6)
CCI score, n (%)	
0	1,746 (21.6)
1	1,730 (27.1)
2	1,339 (21.0)
≥3	1,571 (24.6)

* CCI, Charlson comorbidity index.

Table 2. Estimation of cancer risk in patients with SSc according to sex*

Cancer (*ICD-10* code)	Total (N = 6,386; 49,916.28 PY)			Women (n = 5,123; 39,874.89 PY)			Men (n = 1,263; 10,041.39 PY)		
	Observed, n	Expected, n	SIR (95% CI)	Observed, n	Expected, n	SIR (95% CI)	Observed, n	Expected, n	SIR (95% CI)
All cancer	910	291.7	3.12 (2.92–3.33)	678	231.0	2.94 (2.72–3.17)	232	60.7	3.82 (3.34–4.34)
Solid cancer (C00–C80)	833	271.6	3.07 (2.86–3.28)	619	215.8	2.87 (2.65–3.10)	214	55.9	3.83 (3.33–4.38)
Hepatobiliary cancer									
Liver (C22)	68	12.1	5.62 (4.37–7.13)	55	6.4	8.66 (6.52–11.27)	13	5.7	2.26 (1.20–3.87)
Pancreas (C25)	48	7.5	6.41 (4.73–8.50)	33	5.7	5.83 (4.01–8.19)	15	1.8	8.21 (4.59–13.54)
Gallbladder/biliary tract (C23–C24)	20	5.9	3.39 (2.07–5.24)	17	4.4	3.88 (2.26–6.21)	3	1.5	1.98 (0.40–5.77)
Reproductive cancer									
Breast (C50)	64	57.2	1.12 (0.86–1.43)	64	57.1	1.12 (0.86–1.43)	–	–	–
Ovary (C56)	31	6.0	5.13 (3.48–7.28)	31	6.0	5.13 (3.48–7.28)	–	–	–
Cervix uterus (C53)	25	6.7	3.71 (2.40–5.47)	25	6.7	3.71 (2.40–5.47)	–	–	–
Corpus uterus (C54)	13	7.5	1.72 (0.92–2.95)	13	7.5	1.72 (0.92–2.95)	–	–	–
Prostate (C61)	46	6.6	6.98 (5.11–9.31)	–	–	–	46	6.6	6.98 (5.11–9.31)
Respiratory cancer									
Lung (C33–C34)	166	23.7	7.01 (5.99–8.17)	129	15.6	8.25 (6.89–9.80)	37	8.0	4.61 (3.24–6.35)
Lip, oral cavity, and pharynx (C00–C14)	23	3.5	6.56 (4.16–9.84)	17	2.1	8.27 (4.82–13.25)	6	1.5	4.13 (1.51–8.99)
Larynx (C32)	5	0.6	7.90 (2.55–18.43)	2	0.1	19.28 (2.17–69.61)	3	0.5	5.67 (1.14–16.56)
Gastrointestinal cancer									
Colon and rectum (C18–C20)	72	27.7	2.60 (2.04–3.28)	50	19.5	2.56 (1.90–3.38)	22	8.2	2.69 (1.69–4.08)
Stomach (C16)	60	26.4	2.27 (1.73–2.93)	38	16.9	2.25 (1.59–3.09)	22	9.5	2.32 (1.45–3.51)
Esophagus (C15)	9	1.7	5.30 (2.42–10.06)	6	0.5	11.12 (4.06–24.20)	3	1.2	2.59 (0.52–7.56)
Urinary cancer									
Kidney (C64)	15	5.8	2.59 (1.45–4.27)	10	3.7	2.74 (1.31–5.03)	5	2.1	2.34 (0.75–5.46)
Bladder (C67)	14	2.8	4.98 (2.72–8.36)	8	1.2	6.51 (2.80–12.82)	6	1.6	3.79 (1.39–8.26)
Skin cancer									
Melanoma (C43)	12	1.6	7.38 (3.81–12.89)	9	1.4	6.53 (2.98–12.40)	3	0.2	12.08 (2.43–35.28)
Nonmelanoma of the skin (C44)	53	13.0	4.08 (3.05–5.33)	47	11.0	4.26 (3.13–5.66)	6	2.0	3.05 (1.12–6.65)
Other									
Thyroid (C73)	85	53.7	1.58 (1.26–1.96)	79	49.9	1.58 (1.25–1.97)	6	3.9	1.55 (0.57–3.37)
Central nervous system (C70–C72)	8	1.9	4.15 (1.78–8.17)	6	1.5	4.08 (1.49–8.88)	2	0.5	4.36 (0.49–15.73)
Thymus, heart, or pleura (C37–C38)	6	1.1	5.24 (1.91–11.41)	2	0.9	2.34 (0.26–8.45)	4	0.3	13.79 (3.71–35.29)
Eye (C69)	1	0.1	9.90 (0.13–55.07)	0	–	–	1	0	57.94 (0.76–322.39)
Hematologic cancer (C81–C86, C90–C96, D46, D47.1)	78	13.7	5.70 (4.50–7.11)	60	10.2	5.88 (4.48–7.56)	18	3.5	5.18 (3.07–8.18)
Non-Hodgkin lymphoma (C82–C86, C96)	33	5.2	6.31 (4.34–8.86)	25	3.8	6.52 (4.22–9.63)	8	1.4	5.72 (2.46–11.28)
MDS (D46)	15	1.2	12.52 (7.00–20.65)	11	0.9	12.70 (6.33–22.72)	4	0.3	12.05 (3.24–30.86)
Multiple myeloma (C90)	22	1.9	11.73 (7.35–17.76)	17	1.4	11.86 (6.91–19.00)	5	0.4	11.30 (3.64–26.36)
Leukemia (C91–C95)	6	3.5	1.73 (0.63–3.76)	5	2.6	1.92 (0.62–4.47)	1	0.9	1.16 (0.02–6.46)
Hodgkin lymphoma (C81)	2	0.3	7.96 (0.89–28.74)	1	0.2	5.79 (0.08–32.20)	1	0.1	12.74 (0.17–70.91)
Myeloproliferative disease (D47.1)	1	1.7	0.60 (0.01–3.35)	1	1.3	0.77 (0.01–4.29)	0	–	–

* CI, confidence interval; *ICD-10*, *International Classification of Disease, 10th Revision*; MDS, myelodysplastic syndrome; PY, person-year; SIR, standardized incidence ratio; SSc, systemic sclerosis.

Table 2 also showed the risk of cancer in patients with SSc by sex. Both women and men with SSc had higher overall cancer risks than the general population (women: SIR 2.94 [95% CI 2.72–3.17]; men: SIR 3.82 [95% CI 3.34–4.34]). In women with SSc, the highest SIR was observed for laryngeal cancer, followed by multiple myeloma, MDS, liver cancer, esophageal cancer, and lung cancer. In men with SSc, the highest SIR was observed for eye cancer, followed by thymus, heart, and pleura cancer and Hodgkin lymphoma, melanoma, and MDS.

Risk of cancer in patients with SSc by age at diagnosis and follow-up duration. Next, we evaluated the risk of cancer according to age at diagnosis and follow-up duration (Table 3). In the subgroup analysis by age at diagnosis, cancer risk was significantly higher in all age groups compared to the general population. The highest overall cancer risk was observed in the 20 to 39 years age group (SIR 5.27; 95% CI 4.31–6.38), with the risk decreasing gradually with age and being lowest in the ≥80 years age group (SIR 1.97; 95% CI 1.08–3.31).

In the follow-up duration analysis, cancer risk was elevated throughout all time intervals. The highest SIR was observed within the first year of the follow-up period (SIR 4.44; 95% CI 3.81–5.14). The SIR was lowest during a follow-up duration of three to four years (SIR 2.36; 95% CI 1.97–2.81) and then gradually increased. When the follow-up duration exceeded nine years, the SIR increased to 3.87 (95% CI 3.36–4.44).

Risk of cancer in patients with SSc based on immunosuppressive agent use. We identified ever-users of methotrexate (11,132.11 person-years), hydroxychloroquine (18,264.23 person-years), azathioprine (8,904.41 person-years), cyclophosphamide (5,585.91 person-years), mycophenolate mofetil (4619.64 person-years), and tacrolimus (2,454.35 person-years) during the follow-up period. We then evaluated the SIRs for both overall and site-specific cancers in comparison to the general population, stratified by immunosuppressive agent use (Supplementary Table S3). For all cancers, cyclophosphamide exhibited the highest SIR compared to the general population (SIR 3.05; 95% CI 2.46–3.76). Similarly, for solid tumors, cyclophosphamide had the highest SIR (SIR 2.94; 95% CI 2.34–3.64), followed by hydroxychloroquine and methotrexate. For hematologic malignancies, tacrolimus demonstrated the highest SIR (SIR 9.45; 95% CI 3.04–22.05), followed by methotrexate and cyclophosphamide.

Table 3. Estimation of cancer risk in patients with SSc by age at diagnosis and years of follow-up*

Characteristics	Observed, n	Expected, n	SIR (95% CI)
Age at diagnosis, y			
20–39	105	19.9	5.27 (4.31–6.38)
40–59	464	150.5	3.08 (2.81–3.38)
60–79	327	114.2	2.86 (2.56–3.19)
≥80	14	7.1	1.97 (1.08–3.31)
Follow-up duration, y			
<1	179	40.3	4.44 (3.81–5.14)
1–2	204	69.9	2.92 (2.53–3.35)
3–4	129	54.6	2.36 (1.97–2.81)
5–6	105	42.2	2.49 (2.03–3.01)
7–8	88	31.7	2.78 (2.23–3.42)
≥9	205	53.0	3.87 (3.36–4.44)

* CI, confidence interval; SIR, standardized incidence ratio; SSc, systemic sclerosis.

DISCUSSION

In this study, we revealed that the incidence of cancer is higher in patients with SSc in South Korea compared to the general population. Patients with SSc consistently have higher incidence rates of not only solid tumors but also hematologic malignancies. We reported an SIR value of 3.12 when comparing the cancer incidence rates in patients with SSc to that of the general population. Compared to the results of other studies in which the SIR values ranged from 1.48 to 2.15, this finding appears relatively high.[2,7,8] This discrepancy may be partially explained by differences in study populations and methodologies. In contrast to previous research, the mean follow-up period in our study was relatively long at nine years. Other methodologic variations, such as difference in comparator, inclusion and exclusion criteria, matching variables, and study periods, may also have contributed to the differing results.

The SIR value was highest in the 20 to 39 years age group in this study, but caution is needed in interpreting this result. If we only look at the absolute values, the cancer incidence rate was highest in the 40 to 59 years age group, which is consistent with the general population trend.[9] However, considering that the cancer incidence rate is lower in the 20 to 39 years general population, it can be interpreted that the relative value, the SIR, is highest in this group. Likewise, the highest SIR value was observed in eye cancer, specifically ocular melanoma. When interpreting the high SIR values of rare cancers, such as eye cancer, caution is required as well. Because these cancers are uncommon in the general population, their prevalence may have been calculated as relatively high. Of course, the higher SIR values for cancers of the eye (SIR 9.90), larynx (SIR 7.90), and lip, oral cavity, and pharynx (SIR 6.56) could be attributed to the physiologic aspects of SSc.[10] However, the fact that these SIR values are higher than those typically known for breast or lung cancer warrants attention, considering the relative characteristics of SIR values.

The mechanism linking cancer and SSc can be broadly divided into two categories. The first is malignant transformation within individual organs, which can be exemplified by esophageal cancer and lung cancer. The high risk of esophageal cancer is said to be associated with severe reflux or Barrett's esophagus.[11] In South Korea, esophageal cancer ranks as the 15th most common cancer and has been on a declining trend in recent years.[12]

Considering that the male group accounts for 92% of all cases, the SIR of 11.12 for esophageal cancer in the female subgroup represents a relatively very high risk. Traditionally, cancers known to be associated with SSc include not only esophageal cancer but also lung cancer. Lung cancer is known to be associated with the long-standing presence of SSc interstitial lung disease,[13,14] specifically pulmonary involvement. Traditional risk factors include smoking[15] and the male sex.[16] The risk is reported to increase with a longer duration of SSc and when SSc is diagnosed at a younger age.[17] Additionally, lung cancer is associated with the presence of anti–topoisomerase I antibody and a history of scleroderma renal crisis.[18]

Second, a possible explanation for the association between cancer development and SSc is the use of cytotoxic agents for treating SSc.[11] Immunosuppressive agents used in SSc may increase the risk of cancer occurrence, and different SSc clinical and serologic features may influence malignancy risk. Although 93.7% of patients in our cohort had at least one prescription for glucocorticoids, only 32.4% received these medications for more than one year, indicating that long-term use was relatively uncommon. This suggests that glucocorticoids were often prescribed on a short-term basis, particularly during the early phase of SSc when inflammatory features are more prominent. In clinical practice, glucocorticoids are generally used selectively in patients with specific organ involvement, such as interstitial lung disease, diffuse cutaneous disease, or myositis, in which anti-inflammatory effects can provide symptomatic benefit.[19] Moreover, emerging evidence suggests that early glucocorticoid use in very early or early SSc may help control inflammation and potentially influence disease progression.[20] Therefore, the high rate of any glucocorticoid use in our study likely reflects initial or episodic treatment, rather than prolonged therapy, and may differ from Western cohorts due to differences in clinical practice and treatment strategies.

Although most cytotoxic agents showed high SIR values, some notable distinctions can be identified. First, relatively high SIR values for lung cancer were associated with cyclophosphamide (SIR 9.84) and azathioprine (SIR 8.27). Cyclophosphamide, an alkylating agent, is known to be an effective treatment for lung disease in SSc.[21] Lung fibrosis and a long-standing SSc pulmonary involvement have been proposed as lung cancer risk factors.[22] Second, the high SIR value for esophageal cancer, at 13.97, is associated with cyclophosphamide. Exposure to cyclophosphamide and upper gastrointestinal involvement have been traditionally liked to esophagus cancer.[2] Third, the high SIR values for lip, oral cavity, and pharynx cancers were associated with tacrolimus (SIR 15.34), azathioprine (SIR 13.09), and cyclophosphamide (SIR 11.86). There have been studies indicating that the use of tacrolimus following organ transplantation is strongly associated with malignancy, particularly skin cancer.[23,24] There are also studies suggesting that azathioprine is associated with an increased risk of skin cancer.[25] Cyclophosphamide is known to increase the risk of carcinogenesis in a dose-dependent manner and with prolonged exposure, and it is typically associated with skin cancer, bladder cancer, and leukemia.[26] This pattern may further reflect the possibility that immunosuppressive agents impair tumor immune surveillance, especially in malignancies in which T cell–mediated immunity plays a central role. Tacrolimus may interfere with antitumor immune responses, which could partly explain the elevated SIR observed in this study. Fourth, interestingly, the SIR for breast cancer incidence with the use of all agents was <1, suggesting a potential protective effect. Cyclophosphamide is a widely used chemotherapeutic agent for the treatment of breast cancer. In fact, there are very few studies suggesting that cyclophosphamide has a preventive effect. The protective SIR values identified in this study may have been influenced by other factors, such as patient characteristics or effects related to immune modulation. The current results are difficult to interpret, so further research seems necessary.

It is known that the risk of cancer in SSc exhibits a biphasic temporal relationship.[27] Cancers occurring within a few years of SSc onset are thought to reflect specific immunologic mechanisms, possibly indicating a paraneoplastic phenomenon or shared pathogenic triggers. The elevated cancer incidence early after SSc diagnosis may also be influenced by increased medical surveillance, such as more frequent imaging studies. In contrast, late-onset cancer is known to result from a subtle interplay across repeated target organ inflammation, immunosuppressant use, mesenchymal cell dysfunction, subsequent genetic alterations, and/or coincidental age-related risks.[28] A similar trend was observed in this study. We stratified cancer incidence by detailed intervals of follow-up duration, which allowed for a nuanced assessment of temporal patterns. The SIR was highest in the first year of follow-up (4.44), suggesting a possible early immune-mediated link, whereas a secondary increase after nine years (SIR 3.87) raises the possibility of long-term effects. These findings emphasize the importance of longitudinal monitoring and may point to different pathogenic mechanisms depending on the timing of cancer onset.

Compared to previous studies,[29,30] our findings show some discernible differences in the spectrum of malignancies observed among patients with SSc. We observed a higher incidence of hematologic malignancies and a relatively lower frequency of gynecological cancers. Although this pattern contrasts with many prior reports,[29,30] some studies have also reported similar trends, including elevated rates of hematologic malignancies or fewer gynecologic cancers.[8,31] These discrepancies may stem from a combination of factors such as regional epidemiologic variations, differences in cancer screening and detection practices, and population-specific cancer prevalence. Additionally, immunologic characteristics, particularly the presence or absence of specific autoantibodies such as anti–RNA polymerase III, may influence cancer risk profiles. Although prior research has linked these autoantibodies to certain malignancies, including ovarian

and breast cancer, our study was limited by the lack of serologic data, preventing direct assessment of such associations. This highlights the need for future investigations incorporating detailed immunologic and molecular profiling to better understand the complex interplay across SSc, autoantibody status, and cancer risk.

Among serologic features, the presence of anti–RNA polymerase III and anti–topoisomerase I autoantibodies seems to increase malignancy frequency in patients with SSc.[2] SSc may be a paraneoplastic phenomenon in patients with anti–RNA polymerase III antibodies or those who are negative for centromere, topoisomerase I, or RNA polymerase III autoantibodies.[32] Patients with SSc with anti–RNA polymerase III autoantibodies exhibit a higher prevalence of cancer compared to those with other autoantibodies.[33] Notably, the incidence of breast cancer was particularly elevated during the early stages following the diagnosis of SSc.[3] Furthermore, the genetic and epigenetic mechanisms linking SSc and cancer can be explained by factors such as telomere shortening,[34] microRNAs,[35] and long noncoding RNAs.[36] Additionally, the causes can also be attributed to functional abnormalities in signaling pathways, including glycolysis,[37] oxidative stress,[38] and the PI3K/Akt pathway.[39]

Our study is a retrospective analysis and, as such, has several limitations. First, although we identified an increased cancer risk among patients with SSc, the possibility of surveillance bias cannot be excluded. Patients with SSc typically have more frequent hospital visits and undergo more extensive diagnostic evaluations—including imaging studies such as computed tomography scans and endoscopic procedures—than the general population. This increased medical surveillance may lead to earlier or more frequent detection of malignancies, potentially inflating observed cancer incidence rates compared to the general population. Therefore, caution is needed when interpreting our findings because some of the increased cancer risk may reflect this ascertainment bias rather than a true biologic increase. Second, important carcinogenic factors such as smoking, alcohol consumption, dietary habits, and information on family history were not accounted for in this study. Third, there is no information on the dosage or duration of immunosuppressant use, which necessitates caution when interpreting the data. Fourth, this study did not include data on serologic markers such as specific autoantibodies or other detailed clinical features of SSc (eg, organ involvement, disease subtype), which are known to influence malignancy risk. This limitation reflects the nature of the NHIS claims database, which does not contain laboratory or immunologic test results. Consequently, we were unable to directly evaluate potential associations between autoantibody profiles, including anti–RNA polymerase III, and specific cancer types. Fifth, patients with a history of cancer were excluded, and a two-year washout period was applied. Although this may slightly limit generalizability, it helped reduce reverse causality and ensured a clearer temporal relationship between SSc diagnosis and subsequent cancer development, thereby strengthening internal validity.

In conclusion, newly diagnosed patients with SSc require thorough physical examination and cancer screening based on risk stratification considering age, sex, and prescribed medications. Furthermore, if this epidemiologic relationship is further clarified, it will greatly contribute to establishing cancer screening guidelines for patients with SSc.

AUTHOR CONTRIBUTIONS

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

REFERENCES

1. Hachulla E, Launay D. Diagnosis and classification of systemic sclerosis. Clin Rev Allergy Immunol 2011;40(2):78–83.
2. Lepri G, Catalano M, Bellando-Randone S, et al. Systemic sclerosis association with malignancy. Clin Rev Allergy Immunol 2022;63(3):398–416.
3. Lazzaroni MG, Cavazzana I, Colombo E, et al; EUSTAR co-authors. Malignancies in patients with anti-RNA polymerase iii antibodies and systemic sclerosis: analysis of the EULAR Scleroderma Trials and Research Cohort and possible recommendations for screening. J Rheumatol 2017;44(5):639–647.
4. Song SO, Jung CH, Song YD, et al. Background and data configuration process of a nationwide population-based study using the Korean national health insurance system. Diabetes Metab J 2014;38(5):395–403.
5. van den Hoogen F, Khanna D, Fransen J, et al. 2013 classification criteria for systemic sclerosis: an American College of Rheumatology/-European League Against Rheumatism collaborative initiative. Arthritis Rheum 2013;65(11):2737–2747.
6. Thabet M, Anoun J, Najeh A, et al. Systemic scleroderma and comorbidity: evaluation of the Charlson Comorbidity Index. J Hypertens 2019;37:e222.
7. Olesen AB, Svaerke C, Farkas DK, et al. Systemic sclerosis and the risk of cancer: a nationwide population-based cohort study. Br J Dermatol 2010;163(4):800–806.
8. Carbonell C, Marcos M, Guillén-Del-Castillo A, et al; RESCLE Investigators, Autoimmune Diseases Study Group (GEAS). Standardized incidence ratios and risk factors for cancer in patients with systemic sclerosis: data from the Spanish Scleroderma Registry (RESCLE). Autoimmun Rev 2022;21(10):103167.
9. Jung KW, Kang MJ, Park EH, et al. Prediction of cancer incidence and mortality in Korea, 2023. Cancer Res Treat 2023;55(2):400–407.
10. Coček A, Hahn A, Ambruš M, et al. Complications in the treatment of oropharyngeal carcinoma in patients with systemic sclerosis: a case report. Oncol Lett 2015;9(1):25–28.
11. Weeding E, Casciola-Rosen L, Shah AA. Cancer and scleroderma. Rheum Dis Clin North Am 2020;46(3):551–564.

12. Park SY, Kim DJ. Esophageal cancer in Korea: epidemiology and treatment patterns. J Chest Surg 2021;54(6):454–459.
13. Abu-Shakra M, Guillemin F, Lee P. Cancer in systemic sclerosis. Arthritis Rheum 1993;36(4):460–464.
14. Jeon H, Nam BD, Yoon CH, et al. Radiologic approach and progressive exploration of connective tissue disease-related interstitial lung disease: meeting the curiosity of rheumatologists. J Rheum Dis 2024;31(1):3–14.
15. Kaşifoğlu T, Yaşar Bilge Ş, Yıldız F, et al. Risk factors for malignancy in systemic sclerosis patients. Clin Rheumatol 2016;35(6):1529–1533.
16. Peng H, Wu X, Wen Y, et al. Association between systemic sclerosis and risk of lung cancer: results from a pool of cohort studies and Mendelian randomization analysis. Autoimmun Rev 2020;19(10):102633.
17. Colaci M, Giuggioli D, Sebastiani M, et al. Lung cancer in scleroderma: results from an Italian rheumatologic center and review of the literature. Autoimmun Rev 2013;12(3):374–379.
18. Catano J, Guedon A, Riviere S, et al. Cancers in systemic sclerosis: risk factors, impact on survival and literature review. Article in French. Rev Med Interne 2019;40(10):637–644.
19. Iudici M, van der Goes MC, Valentini G, et al. Glucocorticoids in systemic sclerosis: weighing the benefits and risks - a systematic review. Clin Exp Rheumatol 2013;31(2 suppl 76):157–165.
20. Blagojevic J, Legendre P, Matucci-Cerinic M, et al. Is there today a place for corticosteroids in the treatment of scleroderma? Autoimmun Rev 2019;18(12):102403.
21. Tashkin DP, Elashoff R, Clements PJ, et al; Scleroderma Lung Study Research Group. Cyclophosphamide versus placebo in scleroderma lung disease. N Engl J Med 2006;354(25):2655–2666.
22. Chen M, Liu X, Xu Y, et al. Clinicopathological characteristics of lung cancer in patients with systemic sclerosis. Clin Respir J 2020;14(12):1131–1136.
23. Rodríguez-Perálvarez M, Colmenero J, González A, et al; Chronic immunosuppression, cancer Spanish consortium. Cumulative exposure to tacrolimus and incidence of cancer after liver transplantation. Am J Transplant 2022;22(6):1671–1682.
24. Wang L, Ma K, Yao Y, et al. Carcinogenicity risk associated with tacrolimus use in kidney transplant recipients: a systematic review and meta-analysis. Transl Androl Urol 2022;11(3):358–366.
25. Jiyad Z, Olsen CM, Burke MT, et al. Azathioprine and risk of skin cancer in organ transplant recipients: systematic review and meta-analysis. Am J Transplant 2016;16(12):3490–3503.
26. Calatroni M, Buzio C, Vaglio A. The evolving paradigm of cancer risk related to cyclophosphamide therapy in granulomatosis with polyangiitis. Rheumatology (Oxford) 2015;54(8):1339–1341.
27. Partouche L, Goulabchand R, Maria ATJ, et al. Biphasic temporal relationship between cancers and systemic sclerosis: a clinical series from Montpellier University Hospital and review of the literature. J Clin Med 2020;9(3):9.
28. Greten FR, Grivennikov SI. Inflammation and cancer: triggers, mechanisms, and consequences. Immunity 2019;51(1):27–41.
29. Nemes-Tömöri D, Jász DK, Tari D, et al. Malignancy in systemic sclerosis: a multicenter retrospective study. Biomedicines 2025;13(4):13.
30. Derk CT, Rasheed M, Artlett CM, et al. A cohort study of cancer incidence in systemic sclerosis. J Rheumatol 2006;33(6):1113–1116.
31. Bonifazi M, Tramacere I, Pomponio G, et al. Systemic sclerosis (scleroderma) and cancer risk: systematic review and meta-analysis of observational studies. Rheumatology (Oxford) 2013;52(1):143–154.
32. Maria ATJ, Partouche L, Goulabchand R, et al. Intriguing relationships between cancer and systemic sclerosis: role of the immune system and other contributors. Front Immunol 2019;9:3112.
33. Liu C, Hou Y, Xu D, et al. Analysis of anti-RNA polymerase III antibodies in Chinese Han systemic sclerosis patients. Clin Rheumatol 2020;39(4):1191–1197.
34. Artlett CM, Black CM, Briggs DC, et al. Telomere reduction in scleroderma patients: a possible cause for chromosomal instability. Br J Rheumatol 1996;35(8):732–737.
35. Zhang L, Wu H, Zhao M, et al. Meta-analysis of differentially expressed microRNAs in systemic sclerosis. Int J Rheum Dis 2020;23(10):1297–1304.
36. Wu GC, Pan HF, Leng RX, et al. Emerging role of long noncoding RNAs in autoimmune diseases. Autoimmun Rev 2015;14(9):798–805.
37. Dolcino M, Pelosi A, Fiore PF, et al. Gene profiling in patients with systemic sclerosis reveals the presence of oncogenic gene signatures. Front Immunol 2018;9:449.
38. Schumacker PT. Reactive oxygen species in cancer cells: live by the sword, die by the sword. Cancer Cell 2006;10(3):175–176.
39. Janku F. Phosphoinositide 3-kinase (PI3K) pathway inhibitors in solid tumors: from laboratory to patients. Cancer Treat Rev 2017;59:93–101.

Arthritis Care & Research
Vol. 77, No. 12, December 2025, pp 1474–1486
DOI 10.1002/acr.25603

Where, How, and How Much? A Multicenter Cohort Study of the Relationship Across Lupus Decision-Aid Modality, Place of Administration, Interruption and Viewing Completeness, and Patient-Reported Outcomes

Jasvinder A. Singh,[1] Mark Beasley,[2] and Larry R. Hearld,[2] for the Implementing Decision-Aid for Lupus in Clinics (IDEAL) Consortium

Objective. We assessed whether shared decision-making (SDM) and patient acceptability, feasibility, and overall satisfaction with a computerized patient decision aid (PtDA) for patients with systemic lupus erythematosus (SLE) differs by PtDA setting, modality, and the viewing experience.

Methods. Patients with SLE were invited to view a self-administered computerized SLE PtDA during regular clinic visits at 15 rheumatology clinics in an implementation trial. Patients completed a survey that included SDM measures including the decision conflict scale (DCS), Preparation for Decision Making (PDM) scale, and CollaboRATE scale, and we measured perceived patient acceptability, feasibility, and satisfaction. Patients viewed the SLE PtDA in two settings/places, in clinic or at home (telemedicine visits), using one of three modalities, a touchpad computer, smart phone, or computer (desktop or laptop computer). We also assessed the effects of interruptions while viewing the PtDA and incomplete viewings.

Results. We had a cohort of 813 patients with SLE (43% of 1,895 total) who completed the PtDA modality and setting questions, which were added midway after the COVID-19 pandemic started. In a multivariable-adjusted logistic regression analysis, the setting or modality of viewing the SLE PtDA were not associated with SDM or patient outcomes except the association of place of viewing with feasibility. We noted important significant associations of interruption while viewing the SLE PtDA with lower feasibility, acceptability, and PDM and DCS scores and incomplete viewing of the SLE PtDA with worse PDM and DCS scores.

Conclusion. The SLE PtDA was effective regardless of setting and modality of delivery. Uninterrupted and complete viewing of the SLE PtDA is desirable for better SDM and higher acceptability.

INTRODUCTION

Systemic lupus erythematosus (SLE) affects young women and is associated with high morbidity and mortality.[1,2] Active SLE is effectively treated with immunosuppressive medications, biologics, and glucocorticoids.[3,4] These drugs prevent end organ damage, such as kidney failure and dialysis, which are more common in racial and ethnic minorities compared to White patients.[5–7] Racial and ethnic minority patients are more likely to decline the effective SLE medications due to a greater fear of harms and the lack of recognition of benefits.[8,9] Shared decision-making (SDM) is critical to overcoming these barriers.

SDM is a process by which patients and their health care providers work together to address patients' clinical conditions

ClinicalTrials.gov identifier: NCT03735238.
The contents of this article are solely the responsibility of the authors and do not necessarily represent the views of the Patient-Centered Outcomes Research Institute (PCORI), its Board of Governors, or the Methodology Committee.
Supported by PCORI (award SDM-2017C2-8224). Dr Singh's work was supported by research grants from the National Institutes of Arthritis, Musculoskeletal, and Skin Diseases, NIH (grant P50-AR-060772) and PCORI (grant CER-2020C1-19193).
[1]Jasvinder A. Singh, MBBS, MPH: Michael E. DeBakey VA Medical Center and Baylor College of Medicine, Houston, Texas, and University of Alabama at Birmingham; [2]Mark Beasley, PhD, Larry R. Hearld, PhD: University of Alabama at Birmingham.
Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25603).
Author disclosures and graphical abstract are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25603.
Address correspondence via email to Jasvinder A. Singh, MBBS, MPH, at Jasvinder.md@gmail.com.
Submitted for publication February 16, 2025; accepted in revised form June 23, 2025.

SIGNIFICANCE & INNOVATIONS

- The setting and the modality of delivery of the systemic lupus erythematosus (SLE) patient decision aid (PtDA) had no or minimal effect on patient shared decision-making (SDM), perceived patient acceptability, feasibility, and satisfaction.
- Uninterrupted and complete viewing of the SLE PtDA was associated with better SDM outcomes and higher patient acceptability.
- These findings indicate that dissemination of the SLE PtDA may likely be as effective during telemedicine and video-medicine rheumatology visits as in-person clinic visits and thereby allow its widespread use.
- Multiple modalities (touchpad, smart phone, computer) are effective for SLE PtDA dissemination.

that require action by considering all the information, discussing the advantages and disadvantages, and choosing the best course of action given a patient's unique clinical condition, values and preferences, and context.[10] SDM is facilitated by using patient decision aids (PtDAs). PtDAs reduce decisional conflict by improving knowledge, facilitating patient–provider communication and patient participation in the decision-making[11,12] without increasing the visit duration.[13]

The setting and modality for PtDAs may influence their use. For example, although providing PtDAs during routine clinical visits can increase PtDA feasibility,[14,15] organizational barriers, such as lack of clinic staff time and perceived clinic workflow issues for the healthcare provider, may reduce its perceived feasibility and acceptability by clinic teams.[13] Thus, in-clinic PtDA introduction has some barriers. How this compares to PtDA access at home is unknown. Thus, large knowledge gaps exist in PtDA implementation.

Patients are witnessing a rapidly evolving role of technology in their lives via smart phones, electronic health record messaging, and the emerging use of artificial intelligence.[16] These tools offer unique patient education opportunities. We developed, tested, and implemented an evidence-based SLE PtDA during regular clinic visits in private and academic rheumatology practices in the US.[17–19] The SLE PtDA was adapted to a phone application during the COVID-19 pandemic.[20] To our knowledge, it is not known how in-clinic PtDA use compares to at-home PtDA use, and how the device used for its delivery conrtibutes to its impact on patient outcomes.

The specific aims of our cohort study were to examine key questions in PtDA implementation: Do modality of administration, the place of viewing, and the patient experience of viewing the SLE PtDA impact (a) patient's SDM outcomes and (b) patient acceptability, feasibility, and satisfaction? We hypothesized that in-clinic tablet computer administration would be superior to all other options.

METHODS

Sampling frame. This was an ancillary cohort study of the baseline trial visit from a multicenter prospective implementation trial focused on the dissemination of an SLE PtDA during regular outpatient visits in 15 geographically diverse rheumatology clinics in the US from 2019 to 2024. A detailed study protocol was published.[18] Patient assessments in the main trial were done over 6 months[21] and clinic personnel outcomes[22] were assessed over 24 months, both described previously. The main study findings are reported elsewhere.[22] This cohort study included the data collected during the baseline visit of our implementation trial. See Appendix A for Implementing Decision-Aid for Lupus in Clinics (IDEAL) Consortium authors.

During the COVID-19 pandemic, we received an additional Patient-Centered Outcomes Research Institute (PCORI) supplement grant to adapt the SLE PtDA to both in-clinic and home settings, including providing it through various modalities (smart phone app or desktop computer) in addition to tablet or touchpad computer. The key objective of our ancillary study was to assess whether SDM and patient acceptability, satisfaction, and feasibility of PtDA varied across modalities, settings, and PtDA viewing experience. This article reports the results of this ancillary study. The SLE PtDA provided information on disease manifestations, disease management, treatment options, and the benefits and harms of each treatment option.[20] All participants completed a brief self-administered Qualtrics survey after viewing the SLE PtDA through a link provided at the last page of the electronic SLE PtDA. The study was registered at ClinicalTrials.gov (ID: NCT03735238) and was approved by each of the participating sites' institutional review boards (IRBs), including the University of Alabama at Birmingham (UAB) Coordinating Center (UAB Coordinating Center IRB: 300002554).

The implementation trial enrolled adults (≥18 years) of all races and ethnicities with a clinical diagnosis of SLE during routine rheumatology outpatient clinic visit. We surveyed patients and clinic personnel to understand the barriers and facilitators to DA implementation and sustainment.[18] Study participants viewed an individualized, computerized, culturally tailored SLE PtDA, while in the waiting room or clinic examination room (for in-clinic visits, on a tablet or touchpad computer provided to each participating clinic[17]) or at home (for telemedicine visits). Before the COVID-19 pandemic, all patient participants viewed the SLE PtDA in person in the clinic on a touchpad and completed the survey right after viewing the SLE PtDA, usually within an hour. During the initial COVID-19 pandemic, telemedicine visits required that we make our SLE PtDA remotely available to participants. They were reminded by the study coordinator to view the SLE PtDA before the telemedicine visit and to complete the survey. As the COVID-19 pandemic progressed, patient participants were enrolled during in-person and telemedicine visits. We added the SLE PtDA

as phone apps for smart phones and touchpad, in addition to the SLE PtDA website version for touchpad. Study coordinators assisted patients in accessing the SLE PtDA and the survey, as needed.

SLE PtDA. We developed an individualized, computerized, culturally-tailored, self-administered SLE PtDA in 2014 to facilitate decision-making for immunosuppressive medications in SLE based on extensive qualitative work with the target population including racial and ethnic minorities[23–25] and on the comparative effectiveness research data on medication benefits and risks.[26–28] The SLE PtDA is now available free of cost to the public; its development and testing was funded by PCORI.

This SLE PtDA is superior to the American College of Rheumatology (ACR) SLE pamphlet in a multicenter randomized trial[17] of 301 women with SLE nephritis completed in 2017. The SLE PtDA development was based on the International Patient Decision-Aid Standards (IPDAS)[29] with multistakeholder group input (patients with SLE, patient co-investigators, clinicians, researchers, lupus advocacy leaders) and underwent iterative modification and pilot testing. It was tailored to the target population's numeracy and health and graphical literacy levels.[30]

The SLE PtDA was developed initially for women with SLE nephritis, based on patient and provider feedback and tested in a trial.[17] Subsequently, its content has been expanded to all SLE manifestations, including but not limited to nephritis, and to both men and women. Specific content added to the SLE PtDA included the following: (1) treatments for nonrenal manifestations of lupus, including SLE skin and joint disease; (2) treatment of nonrenal vital organ involvement; (3) the use of biologics (belimumab, rituximab, anifrolumab); (4) second-line treatment of active renal disease despite cyclophosphamide or mycophenolate mofetil with voclosporin or belimumab; and (5) male sex-specific issues.

Per IPDAS principles, we performed regular updates of the content of the SLE PtDA. The SLE PtDA was iteratively updated with last update in 2023 as new medications were approved for SLE (voclosporin, anifrolumab) or existing SLE medications received new indications (belimumab). For updates, we focused primarily on US Food and Drug Administration (FDA) approvals.

Outcomes of interest. *SDM.* We assessed SDM with three measures: the low literacy version of the decision conflict scale (DCS), Preparation for Decision Making (PDM) scale, and the CollaboRATE scale.

The low literacy version DCS is a validated self-administered instrument[31] that has 10 items with three response categories: yes = 0, unsure = 2, and no = 4. Item scores were summed, divided by 10 and multiplied by 25. The DCS scores range from 0 (best; no decisional conflict) to 100 (worst; extremely high decisional conflict).[32–37] DCS scores ≥25 indicate clinically significant residual decisional conflict.[38]

PDM is a validated scale that assesses a patient's perception of how useful a decision aid is in preparing the respondent to communicate with their practitioner at a consultation focused on making a health decision.[39] It consists of 10 questions scored on an ordinal scale from 1 to 5 (1 = "not at all" to 5 = "a great deal").[39] We summed the score of the 10 items and divided the sum by 10 and converted it to a 0 to 100 scale by subtracting 1 from this summed score and multiplying by 25. Higher PDM scores indicate higher perceived levels of preparation for decision-making.

The CollaboRATE scale is a validated three-item scale (understand, listen, include what matters) that assesses the extent to which the clinical team understood patient's health issues during an appointment.[40] The score ranges from 0 to 9 (0 = no effort was made; 9 = every effort was made) for each item, which are summed to get an overall scale score (0–27). Higher scores indicate more effective SDM.[40]

Perceptions of PtDA. Patient acceptability of the SLE PtDA was assessed using a validated acceptability measure developed by the University of Ottawa.[41] This instrument allows specification of survey items to the PtDA being tested. Patient acceptability was assessed using the five-item ordinal scale with response options ranging from 1 to 4 (1 = poor to 4 = excellent). Individual item responses were aggregated into an overall response by averaging across all items on a 1 to 4 scale.[41]

Feasibility of the SLE PtDA was assessed with a validated instrument[33] that included four items, with response options ranging from not feasible at all (1) to extremely feasible (5). Summated scores were created by averaging across the four items on a 1 to 5 scale.[42]

Patient satisfaction with the SLE PtDA was assessed using an ordinal single item scale,[42] "The decision aid was easy to use" with response ranging from strongly disagree (1) to strongly agree (5). A higher score on each of these three outcome measures indicates high patient acceptability, feasibility, and satisfaction.

Covariates. We included several important sociodemographic, clinic, disease, and PtDA covariates in our analysis, accounting for important covariates previously known to be associated with PtDA outcomes and potential confounders of outcomes (age, race and ethnicity, sex, marital status, education, clinic site, flare, PtDA version),[43–46] to obtain adjusted estimates of associations between independent variables of interest and outcomes. Patient demographic characteristics included age, race and ethnicity, sex, and marital status. Insurance payer type was categorized as Medicare, Medicaid, private insurance, and other. Social determinants of health (SDOH) included the highest education level achieved, categorized as less than high school (HS), HS degree or general education development (GED), or greater than HS and patient residence, categorized as rural, suburban, or urban. Payer type, residence, and education questions

were added during the COVID-19 pandemic, around the time the modality question was added; therefore, they have missingness for the overall sample surveyed before their implementation. We examined clinic level variables, including the clinic site (each site modeled as a random effect) and clinic type, categorized as private, academic, or hybrid practice rheumatology clinic site. We included disease characteristics, presence versus absence of an SLE flare, and PtDA version viewed (recommended by rheumatology provider), full scenario version (included immunosuppressive treatment options for active SLE kidney or end organ disease; scenario A–D) versus lupus lite version (a shorter version for less active SLE disease including medications other than immunosuppressives). Lupus Lite had a short viewing time of 10 to 15 minutes, whereas the full versions took 15 to 25 minutes, based on patient choice for viewing the optional sections.

Independent variable of interest. There were four independent variables of interest. Modality of SLE PtDA delivery was either tablet or touchpad computer, participant smartphone, or a computer (desktop or laptop computer; three-level variable). The setting of SLE PtDA delivery was either at home or in the clinic (two-level variable). For multivariable models, we examined a six-level modality and setting composite variable with in-clinic computer viewing as the reference category. We assessed patient experience of viewing related to interruption and complete viewing of the SLE PtDA. Patient-reported interrupted viewing indicated whether a patient was interrupted while viewing the SLE PtDA or not. Finally, patient-reported complete viewing reflected whether the patient completed the viewing of the SLE PtDA or not.

Statistical analysis. We calculated summary statistics as number and proportions for categorical variables and means with standard deviations or standard errors for continuous variables. We examined unadjusted associations by the key variables of interest. Modality, setting, complete viewing, and interruptions of an SLE PtDA viewing were examined as key independent variables of interest. In the main multivariable mixed model regression analysis, we examined the association of these variables with outcomes of interest while adjusting for covariates, including the clinic level variables (each site modeled as a random effect), patient characteristics (sex, race and ethnicity, marital status, and age), SDOH (patient residence, education, insurance payer type), patient-reported current SLE flare, and the lupus PtDA version (full scenario version vs lupus lite) viewed. We considered age, race and ethnicity, sex, education, current lupus flare, lupus DA version, and clinic characteristics as potential confounders for the association of primary variables of interest with all SDM outcomes, and lupus DA version, clinic characteristics, and lupus flare as potential confounders for patient acceptability, feasibility, and satisfaction outcomes. We calculated the multivariable-adjusted least squares estimates from the mixed models to understand the clinical relevance of each of these independent factors. This was a complete case analysis, that is, participants with missing observations on key variables included in our multivariable analysis were not included. We tested for collinearity between interruption and complete SLE PtDA viewing to ensure this was not a concern for these analyses.

Sensitivity analyses (SAs) included the following: SA1, modality and setting combined into a four-level composite variable, in-clinic tablet or computer, in-clinic phone, home tablet or computer, and home phone; SA2 modeled the interruption of the viewing of the SLE PtDA into a three-level composite variable, no interruption, interrupted at home, or in clinic; and SA3 adjusted for each clinic site. We examined whether the "where viewed by how viewed" cross-classification had a statistically significant interaction when using the composite variable in models. Because collinearity diagnostics for mixed models are an unresolved issue, we ran models treating site as a fixed effect and computed standard and generalized collinearity diagnostics. These analyses showed that although some of the covariates (eg, site, clinic type) had relatively high levels of collinearity, the independent variables of main interest had low levels of collinearity based on generalized collinearity statistics. All analyses were performed using SAS version 9.4 (SAS Institute Inc).

RESULTS

Study cohort. Of the 1,895 patient participants in the implementation trial, the mean age was 44.68 years (SD, 14.35), 91% were female, 40% were White, and 44% were African American and 8% reported Hispanic or Latino ethnicity (Table 1). This study included the 813 participants (43%) who answered the modality study questions, which were added midway of the implementation trial. They were similar to the overall sample in demographics. Less than half of the participants (43%) were married, 81% had higher than HS education, and 22% resided in rural areas (Table 1). There was low correlation coefficient of –0.09 between SLE PtDA completion and interruption. The study flow chart shows the study cohort for this study (Figure 1).

Unadjusted difference by SLE PtDA delivery settings and modality. There were significant differences in patient demographics when comparing the subgroups of modality and place of viewing of the SLE PtDA. Compared to participants in the other three subgroups, those who viewed the SLE PtDA at home on a smartphone app were significantly younger; less likely to be male, White, or married; and less likely to have greater than HS education and have commercial or private insurance (Table 2).

Comparing the four participant groups based on SLE PtDA delivery settings of clinic versus home and of phone versus tablet or other computer, we found no differences in patient acceptability or key SDM outcomes, that is, DCS, CollaboRATE and PDM scores were not significantly different in unadjusted analyses

Table 1. Baseline demographic characteristics of patients who answered the modality question and of the overall study sample who viewed the SLE PtDA*

Variable	All patients who viewed the SLE PtDA, N = 1,895	Patients who answered the modality question, n = 813
Age, y		
Mean ± SD	44.68 ± 14.35	44.27 ± 13.95
Median (IQR)	44 (33–56)	44 (33–54)
Sex		
Female	1,731 (91.3)	766 (94.2)
Male	124 (6.5)	47 (5.8)
Missing	40 (2.1)	0 (0.0)
Race		
Asian	80 (4.2)	33 (4.1)
African American	827 (43.6)	366 (45.0)
Hispanic	149 (7.9)	62 (7.6)
Other/mixed	22 (1.2)	11 (1.4)
White	754 (39.8)	340 (41.8)
Missing	63 (3.3)	1 (0.1)
Ethnicity		
Hispanic or Latino	149 (7.9)	62 (7.6)
Not Hispanic or Latino	1,685 (88.9)	750 (92.3)
Missing	61 (3.2)	1 (0.1)
Insurance type		
Commercial/private	655 (34.6)	379 (46.6)
Medicare	340 (17.9)	187 (23.0)
Medicaid	221 (11.7)	120 (14.8)
Other	234 (12.3)	119 (14.6)
Missing	445 (23.5)	8 (1.0)
Marital status		
Married	443 (23.4)	347 (42.7)
Separated	19 (1.0)	18 (2.2)
Divorced	87 (4.6)	69 (8.5)
Single	372 (19.6)	307 (37.8)
Living with partner	45 (2.4)	39 (4.8)
Widowed	33 (1.7)	27 (3.3)
Other	7 (0.4)	2 (0.2)
Missing	889 (46.9)	4 (0.5)
Residence		
Urban	305 (16.1)	246 (30.3)
Suburban	483 (25.5)	387 (47.6)
Rural	218 (11.5)	176 (21.6)
Missing	889 (46.9)	4 (0.5)
Education level		
Less than HS	27 (1.4)	19 (2.3)
HS degree or GED	152 (8.0)	129 (15.9)
Greater than HS	827 (43.6)	661 (81.3)
Missing	889 (46.9)	4 (0.5)
SDM Outcomes for SLE PtDA		
DCS score		
Mean ± SD	19.48 ± 23.78	21.86 ± 24.93
Median (IQR)	10 (0–35)	15 (0–40)
DCS ≥25	638 (34.5)	317 (39.0)
DCS <25	1,211 (65.5)	496 (61.0)
Missing, n	46	0
PDM total		
Mean ± SD	74.44 ± 23.76	71.34 ± 24.74
Median (IQR)	77.5 (60–95)	75 (55–92)
Missing, n	68	4
CollaboRATE		
Mean ± SD	25.17 ± 4.10	25.18 ± 3.92
Median (IQR)	30 (26–30)	30 (26–30)
Missing, n	26	0

(Continued)

Table 1. (*Cont'd*)

Variable	All patients who viewed the SLE PtDA, N = 1,895	Patients who answered the modality question, n = 813
PROs for implementation of SLE PtDA		
Patient acceptability		
Mean ± SD	3.26 ± 0.63	3.19 ± 0.63
Median (IQR)	3.2 (3–4)	3.2 (2.8–3.8)
Missing, n	37	4
Patient feasibility scale		
Mean ± SD	4.33 ± 0.77	4.33 ± 0.74
Median (IQR)	4.5 (4–5)	4.5 (4–5)
Missing, n	73	4
Patient satisfaction scale		
Mean ± SD	4.40 ± 1.08	4.56 ± 1.01
Median (IQR)	5 (4–5)	5 (4–5)
Missing, n	40	0

* Values are n (%), unless specified otherwise. DCS, decision conflict scale; GED, general education development; HS, high school; IQR, interquartile range; PDM, Preparation for Decision Making; PRO, patient-reported outcome; PtDA, patient decision aid; SDM, shared decision-making; SLE, systemic lupus erythematosus.

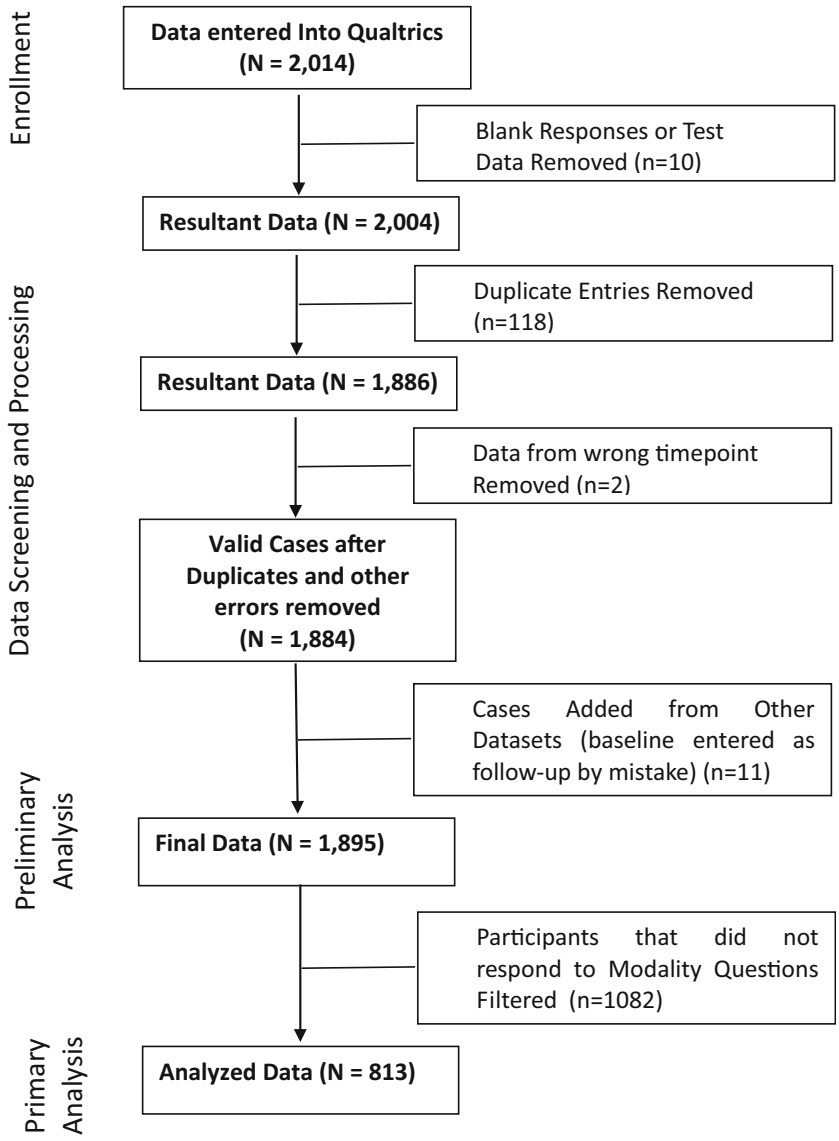

Figure 1. Flowchart of study enrollment at baseline, data screening, and filtering for nonresponse to modality questions.

Table 2. Demographics by site and modality of lupus DA review*

Variable	At-home tablet	At-home phone app	At-home computer	In-clinic tablet	In-clinic phone app	In-clinic computer
n	36	180	123	388	57	29
Age, y						
Mean ± SD	48.58 ± 15.24	41.80 ± 13.36	44.92 ± 14.14	44.98 ± 13.92	43.75 ± 13.27	42.97 ± 15.33
Median (IQR)	47.50 (39–59)	41 (32–52)	44 (33–55)	45 (33–56)	43 (34–54)	41 (29–52)
Sex						
Female	35 (97.2)	176 (97.8)	116 (94.3)	360 (92.8)	55 (96.5)	24 (82.8)
Male	1 (2.8)	4 (2.2)	7 (5.7)	28 (7.2)	2 (3.5)	5 (17.2)
Missing, n	0	0	0	0	0	0
Race						
Asian	2 (5.6)	5 (2.8)	10 (8.1)	16 (4.1)	0 (0.0)	0 (0.0)
African American	20 (5.4)	95 (7.5)	33 (3.6)	173 (6.6)	27 (5.4)	18 (5.4)
Hispanic	1 (2.8)	14 (7.8)	12 (9.8)	29 (7.5)	3 (5.3)	3 (10.3)
Other/mixed	0 (0.0)	4 (2.2)	0 (0.0)	6 (1.6)	0 (0.0)	1 (3.4)
White	13 (36.1)	62 (34.4)	68 (55.3)	163 (42.1)	27 (47.4)	7 (24.1)
Missing, n	0	0	0	1	0	0
Ethnicity						
Hispanic or Latino	1 (2.8)	14 (7.8)	12 (9.8)	29 (7.5)	3 (5.3)	3 (10.3)
Not Hispanic or Latino	35 (97.2)	166 (92.2)	111 (90.2)	358 (92.5)	54 (94.7)	26 (0.0030)
Missing, n	0	0	0	1	0	0
Insurance type						
Commercial/private	11 (31.4)	69 (39.2)	78 (63.9)	183 (47.3)	28 (50.0)	10 (34.5)
Medicare	13 (37.1)	41 (23.3)	20 (16.4)	90 (23.3)	14 (25.0)	9 (31.0)
Medicaid	2 (5.7)	30 (17.0)	15 (12.3)	60 (15.5)	7 (12.5)	6 (20.7)
Other	9 (25.7)	36 (20.5)	9 (7.4)	54 (14.0)	7 (12.5)	4 (13.8)
Missing, n	1	4	1	1	1	0
Marital status						
Married	11 (31.4)	69 (38.5)	55 (45.1)	173 (44.7)	32 (56.1)	7 (24.1)
Separated	3 (8.57)	4 (2.23)	5 (4.10)	6 (1.55)	0 (0.00)	0 (0.00)
Divorced	2 (5.7)	13 (7.3)	5 (4.1)	45 (11.6)	3 (5.3)	1 (3.4)
Single	14 (40.0)	82 (45.8)	43 (35.2)	132 (34.1)	19 (33.3)	17 (58.6)
Living with partner	1 (2.9)	7 (3.9)	11 (9.0)	17 (4.4)	1 (1.8)	2 (6.9)
Widowed	4 (11.4)	4 (2.2)	3 (2.5)	12 (3.1)	2 (3.5)	2 (6.9)
Other	0 (0.0)	0 (0.0)	0 (0.0)	2 (0.5)	0 (0.0)	0 (0.0)
Missing, n	1	1	1	1	0	0
Residence						
Urban	11 (31.4)	59 (33.0)	36 (29.5)	107 (27.6)	11 (19.3)	22 (75.9)
Suburban	17 (48.6)	81 (45.3)	62 (50.8)	188 (48.6)	32 (56.1)	7 (24.1)
Rural	7 (20.0)	39 (21.8)	24 (19.7)	92 (23.8)	14 (24.6)	0 (0.0)
Missing, n	1	1	1	1	0	0
Education level						
Less than HS	1 (2.9)	3 (1.7)	1 (0.8)	13 (3.4)	1 (1.8)	5 (17.2)
HS degree or GED	6 (17.1)	43 (24.0)	9 (7.4)	57 (14.7)	9 (15.8)	24 (82.8)
Greater than HS	28 (80.0)	133 (74.3)	112 (91.8)	317 (81.9)	47 (82.5)	0 (0.0)
Missing, n	1	1	1	1	0	0
SDM outcomes for the SLE PtDA						
DCS score						
Mean ± SD	24.03 ± 24.92	19.39 ± 23.24	24.02 ± 27.52	21.51 ± 24.57	23.77 ± 25.18	26.38 ± 28.06
Median (IQR)	15 (0–47.5)	10 (0–35)	15 (0–45)	15 (0– 40)	20 (0–40)	15 (0–50)
PDM Total						
Mean ± SD	73.14 ± 25.19	71.01 ± 23.71	67.89 ± 26.02	71.58 ± 25.20	73.16 ± 24.28	78.97 ± 18.13
Median (IQR)	77.5 (60–95)	75 (55–90)	73.8 (55–85)	75 (55–92.5)	75 (55–100)	82.5 (75–92.5)
CollaboRATE						
Mean ± SD	25.39 ± 3.22	25.29 ± 3.58	24.34 ± 5.61	25.28 ± 3.65	25.25 ± 3.37	26.45 ± 1.74
Median (IQR)	27 (26–27)	27 (25–27)	27 (24–27)	27 (25–27)	27 (24–27)	27 (27–27)
DCS ≥25	17 (47.2)	61 (33.9)	49 (39.8)	153 (39.4)	24 (42.1)	13 (44.8)
PROs for implementation of the SLE PtDA						
Patient acceptability						
Mean ± SD	3.25 ± 0.60	3.21 ± 0.60	3.05 ± 0.72	3.21 ± 0.61	3.06 ± 0.70	3.49 ± 0.55
Median (IQR)	3.17 (2.8–4)	3.17 (2.8–3.8)	3 (2.5–3.7)	3.17 (3–3.8)	3 (2.7–3.7)	3.7 (3–4)

(Continued)

Table 2. (*Cont'd*)

Variable	At-home tablet	At-home phone app	At-home computer	In-clinic tablet	In-clinic phone app	In-clinic computer
Patient feasibility scale						
Mean ± SD	4.32 ± 0.75	4.20 ± 0.77	4.38 ± 0.72	4.39 ± 0.71	4.26 ± 0.86	4.47 ± 0.53
Median (IQR)	4.50 (4–5)	4.25 (4–5)	4.50 (4–5)	4.5 (4–5)	4.25 (4–5)	4.5 (4–5)
Patient satisfaction scale						
Mean ± SD	4.58 ± 0.94	4.59 ± 0.68	4.33 ± 1.15	4.38 ± 1.09	4.47 ± 0.87	4.48 ± 1.15
Median (IQR)	5 (5–5)	5 (4–5)	5 (4–5)	5 (4–5)	5 (4–5)	5 (5–5)

* Values are n (%), unless specified otherwise. Two *P* values are provided for variables in which both mean and median are presented due to the comparison of means using parametric tests and medians using nonparametric tests. DCS, decision conflict scale; GED, general education development; HS, high school; IQR, interquartile range; PDM, Preparation for Decision Making; PRO, patient-reported outcome; PtDA, patient decision aid; SDM, shared decision-making; SLE, systemic lupus erythematosus.

(Table 2). We found small significant differences in perceptions of feasibility and satisfaction (Table 2).

Similarly, we found that younger age, non-White race, Hispanic or Latino ethnicity, and lower education level were associated with higher unadjusted risk of complete viewing of the PtDA. Younger age, non-White race, Hispanic or Latino ethnicity, and urban residence were associated with higher unadjusted risk of uninterrupted viewing of the PtDA (Supplementary Material 1).

Multivariable-adjusted association of decision-aid characteristics with SDM and other patient-reported outcomes. In multivariable-adjusted mixed model regression analysis, we found important significant association of interruption while viewing the SLE PtDA with lower feasibility, acceptability, and PDM and DCS scores and incomplete viewing of the SLE PtDA with worse PDM and DCS scores.

Specifically, we noted that the following variables were significantly associated with a higher PDM score: minority race or ethnicity, Medicaid or Medicare insurance payer, uninterrupted viewing of the SLE PtDA, and complete viewing of the SLE PtDA (Table 3). The following were significantly associated with lower decisional conflict, that is, lower DCS score: male sex, minority race or ethnicity, viewing the lupus lite version of the SLE PtDA (vs others), uninterrupted viewing of the SLE PtDA, and complete viewing of the SLE PtDA (Table 4). None of the variables were associated with the CollaboRATE score or the overall patient satisfaction.

We found that the modality/location composite of viewing of the SLE PtDA and whether the patient was interrupted while viewing SLE DA were each significantly associated with feasibility and patient acceptability (Table 4). The adjusted differences in means were 5 to 8 points worse for the PDM score (Supplementary Material 2) and the DCS score (Supplementary Material 3). Numerically smaller adjusted differences were noted for patient acceptability (Supplementary Material 4) and feasibility (Supplementary Material 5–6). Full main models are also provided (Supplementary Material 7–8). Figure 2 shows that patients who viewed the PtDA in the clinic on a computer or a touchpad had higher degrees of acceptability.

Sensitivity Analyses (SAs). In SAs in which modality and setting were combined into a four-level composite variable, we reproduced most findings from the main analyses with minimal

Table 3. Mixed multivariable-adjusted regression model results for DCS, CollaboRATE, and PDM surveys for key characteristics*

Independent variables	df	DCS score			CollaboRATE score			PDM score		
		β	SE	*P* value	β	SE	*P* value	β	SE	*P* value
DA completed (Ref, not completed)	1	−7.874	2.179	**0.0003**	0.000	0.353	0.9998	5.106	2.172	**0.0190**
DA interrupted (Ref, not interrupted)	1	7.304	2.126	**0.0006**	−0.628	0.345	0.0689	−6.004	2.119	**0.0047**
Where by how (Ref, clinic computer)	5			0.1923[a]			0.1945[a]			0.1767[a]
Home tablet		−3.773	6.866	0.5828	−0.848	1.113	0.4464	−11.283	6.842	0.0995
Home phone		−7.073	5.802	0.2232	−0.743	0.941	0.4302	−14.207	5.782	0.0142
Home computer		−2.335	5.930	0.6938	−1.722	0.962	0.0737	−13.207	5.909	0.0257
Clinic tablet		−8.679	5.508	0.1155	−0.611	0.893	0.4943	−10.743	5.489	0.0507
Clinic phone		−7.915	6.363	0.2139	−0.399	1.032	0.6992	−8.128	6.340	0.2003

* Bold values represent statistically significant associations with a *P* value of <0.05. Additionally adjusted for sex, race and ethnicity, age, marital status, insurance, patient residence, education, patient-reported current SLE flare, and the version of the systemic lupus erythematosus (SLE) PtDA. Other variables significantly associated with a higher Preparation for Decision Making (PDM) score not shown in the table: minority race or ethnicity, Medicaid or Medicare insurance payer. Higher PDM, higher CollaboRATE and lower decision conflict scale (DCS) scores correspond to higher levels of Shared Decision-Making (SDM), and therefore are most desirable. b, beta estimate; df, degree of freedom; PtDA, patient decision aid; SE, standard error.

[a] *P* value from Wald Type 3 test of omnibus main effect.

Table 4. Mixed multivariable-adjusted regression model results for patient feasibility, acceptability of computerized information, and overall satisfaction for key characteristics*

Independent variables	df	Feasibility			Acceptability			Overall satisfaction		
		β	SE	*P* value	β	SE	*P* value	β	SE	*P* value
DA completed (Ref, not completed)	1	0.079	0.066	0.2312	0.040	0.056	0.4749	0.096	0.091	0.2927
DA interrupted (Ref, not interrupted)	1	−0.180	0.064	**0.0050**[a]	−0.197	0.055	**0.0004**	−0.048	0.089	0.5922
Where by how (Ref, clinic computer)	5			**0.0029**			**0.0161**			0.6537
Home tablet		−0.427	0.206	0.0391	−0.388	0.178	0.0293	0.049	0.287	0.8642
Home phone		−0.583	0.174	0.0009	−0.373	0.150	0.0131	0.075	0.243	0.7586
Home computer		−0.404	0.178	0.0238	−0.513	0.153	0.0009	−0.117	0.248	0.6379
Clinic tablet		−0.325	0.166	0.0504	−0.330	0.142	0.0206	−0.076	0.230	0.7420
Clinic phone		−0.475	0.191	0.0132	−0.440	0.164	0.0076	−0.019	0.266	0.9441

* Bold values represent statistically significant associations with a *P* value of <0.05. PtDA, patient decision aid; SLE, systemic lupus erythematosus. Other variables significantly associated with a lower DCS score not shown in the table: male sex, minority race or ethnicity, viewing the lupus lite versions of the SLE PtDA (versus other versions). Higher scores on patient feasibility, acceptability and satisfaction scales are most optimal and desirable. Additionally adjusted for sex, race and ethnicity, age, marital status, insurance, patient residence, education, patient-reported current SLE flare, and the version of the SLE PtDA.
[a] *P* value from Wald Type 3 test of omnibus main effect. The "where viewed" by "how viewed" cross-classification had a statistically significant interaction (*P* = 0.0003).

differences (SA1; Supplementary Material 9–10). For SAs in which we modeled the interruption of the viewing of the SLE PtDA into a three-level composite variable (no interruption, interrupted at home, interrupted in clinic) or examined the composite DA interrupted or completed variable, we reproduced the main study findings (SA2; Supplementary Material 11–12). SAs that additionally adjusted for clinic site confirmed the main study findings (SA3; Supplementary Material 13–14).

DISCUSSION

Whether and to what extent the modality, setting, and patient-reported actual SLE PtDA delivery (uninterrupted and complete viewing) impacts SDM and patient outcomes in implementation initiatives is not known. In this multicenter SLE PtDA implementation trial, we examined the association of these important PtDA implementation variables with the patient's SDM

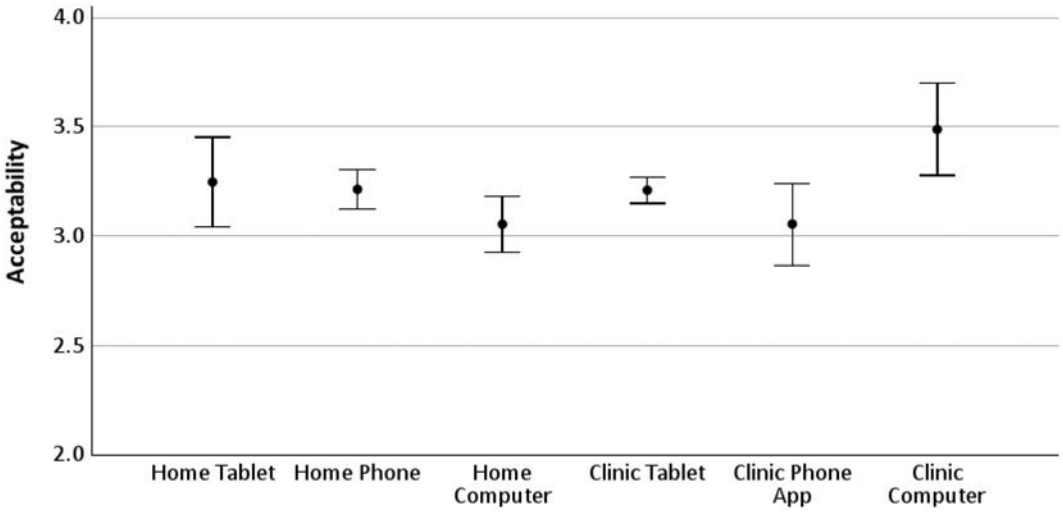

Figure 2. Mean patient acceptability of systemic lupus erythematosus PtDA information for the "where DA viewed" (home versus clinic) by "how DA viewed" (modality: phone, tablet, or computer or laptop) interaction with 95% confidence interval error bars. Each point estimate represents the mean patient acceptability score on a scale from 1 to 4. The x-axis shows each setting modality of viewing the SLE PtDA, and the y-axis shows abbreviated acceptability scale scores from 2 to 4 (full range of the scale is 1–4). Acceptability was assessed using a validated acceptability measure developed by the University of Ottawa, which is a five-item ordinal scale with response options ranging from 1 to 4 (1 = poor to 4 = excellent). Individual item responses were aggregated into an overall response by averaging across all items on a 1–4 scale. Each point estimate is represented by a dot, and whiskers represent the 95% confidence intervals for each point estimate. PtDA, patient decision aid; SLE, systemic lupus erythematosus.

outcomes and perceptions. We made several observations that merit further discussion.

We observed that seeing the full content as intended by the patient and having no interruption when viewing the SLE PtDA were each independently and significantly associated with better SDM outcomes and more positive patient perception of the PtDA. Specifically, full-content viewing of the SLE PtDA was associated with better PDM and DCS scores. Uninterrupted viewing of the SLE PtDA was associated with better PDM and DCS scores, feasibility, and acceptability. The adjusted mean differences ranged from 5 to 8 points for DCS and PDM (0–100 scales) and 0.16 to 0.18 for patient acceptability and feasibility scores (1–5 scales). These differences approximated the clinically meaningful difference thresholds for SDM outcomes but seemed less impressive for patient acceptability and feasibility.

The main implementation question is how would and could organizations (or the PtDA developer) mitigate the negative effects of interruptions and incomplete viewings? Is there technology or clinic or patient-focused solutions that can help overcome this issue? Additional research is needed to understand why patients were interrupted and/or did not complete their viewing of the SLE PtDA and whether it is possible to reduce interruptions. Based on the informal briefing with some participants, some in-clinic interruptions were at least partially related to the patient transition during the clinic visit, including the patient check-in process, nurse check-in, and vital-sign recording; rooming; provider assessment; check-out; and completion of laboratory or radiology tests after the check-out.

On the other hand, the setting (home vs in clinic) and the modality (touchpad vs computer vs smartphone) used to access the SLE PtDA were not associated with differences in SDM outcomes (DCS, PDM, CollaboRATE scores) or patient perceptions of the PtDA, with one notable exception of a positive association between in-clinic use of a tablet and feasibility. Thus, a key study hypothesis was rejected. This is an important negative finding. The fact that SLE PtDA is feasible, acceptable, and effective at fostering SDM, independent of the device or the setting, implies potential opportunities for widespread dissemination of the SLE PtDA. This finding provides high confidence for future dissemination of the SLE PtDA in clinic and home settings using various modalities by engaging patients, providers, and advocacy organizations to disseminate the SLE PtDA. Whether and to what extent social media and active advertising versus provider-based dissemination can help with widespread dissemination of SLE PtDA needs further evaluation. The modality and setting were chosen by each study participant. However, the study period, which included social distancing during the early COVID-19 period and avoidance of in-person visit during the early COVID-19 vaccination period, likely impacted the setting; patient's socioeconomic status likely impacted the availability of and/or access to a desktop or touchpad computer at home. The contribution of these factors to the selection of setting and modality cannot be determined.

Another key study finding was that female sex, Hispanic ethnicity, and African American race were each independently associated with lower, better DCS scores, and Hispanic ethnicity and African American race were each independently associated with higher, better PDM scores compared to their counterparts. These important findings confirm the effectiveness of our SLE PtDA in facilitating SDM in racial and ethnic minorities with SLE, similar to observations of the effectiveness of the earlier version of this SLE PtDA in women with SLE kidney disease in our original randomized trial.[17]

Our SLE PtDA is available as a free phone app in the Android Play Store and the iOS App Store. Patients with SLE across the spectrum of socioeconomic status, literacy, numeracy, and health care access can be educated. The widespread use of SLE PtDA has the potential to reduce disparities in SLE care and outcomes. SLE is more prevalent and is associated with worse outcomes in racial and ethnic minorities.[5–7,47–49] Our SLE PtDA could be an important tool to reduce racial disparities in SLE patient education.

In a systematic review, the use of PtDAs led to better informed, more knowledgeable patients who were clearer about their values,[12] had more accurate risk perceptions, and made value-congruent choices across a wide variety of decision contexts. We did not observe any adverse effects related to the use of SLE PtDA during routine clinic visits on patient outcomes.

Providers can use SLE PtDA as an educational tool to enhance patient–physician communication. They can reallocate the time they would have spent providing basic facts (in the absence of use of patient self-administered, free-of-cost SLE PtDA) to more advanced specific treatment-related risks/benefits and patient value/preference-centered SDM discussions with patients who have reviewed the PtDA. This can help patient–provider dyads achieve desired improved SLE outcomes. A separate analysis of time and cost of using our SLE PtDA showed the important economic implications of using this free tool, which are favorable to its widespread use, since the SLE PtDA has the potential to decrease time spent by the provider during an outpatient rheumatology visit when patients have viewed the SLE PtDA.[50]

Our results must be interpreted considering several limitations. This was a prospective single-arm trial that enrolled consecutive patients in rheumatology clinics with a diagnosis of SLE. The SLE PtDA was superior to the ACR SLE pamphlet in a randomized trial previously.[17] Due to an observational study design, residual confounding bias is likely. We therefore controlled for multiple patient, clinical, and geographic factors to account for potential confounders (current lupus flare, lupus DA version, and patient education level) in our study. We also performed several SAs, which support the robustness of our study findings. Due to the focus on implementation of SLE PtDA in this trial and not

clinical aspects of SLE, additional validated SLE measures of potential interest, including provider-assessed disease activity, severity, specific organ characteristics, and treatment were not collected. Whether many statistically significant associations are clinically meaningfully different or not is unclear.

Our findings may not be generalizable to all clinic settings in the US because we only examined people at 15 US sites. We aimed to enhance generalizability by including geographically diverse sites, including a mix of private and academic rheumatology sites. The key modality and setting variables were added to data collection halfway through the study, and therefore, this study reports on a subgroup of all patients. However, demographic characteristics of the those with modality and setting data were similar to the overall sample. Because participants would generally be unlikely to respond to PDM and DCS immediately after an incomplete viewing at home, the effect of the incomplete viewing of SLE PtDA may have diminished over time, or their perspectives may have changed due to additional information exposure. Therefore, their responses may only accurately reflect their initial impressions immediately after viewing the SLE PtDA, and not long-term effects. Another study limitation was that participants did not provide specific reasons for interruptions (eg, technical issues, time constraints, or types of distractions) because our objective was to keep the responder burden low to modest. This could have provided important information. This study was funded after the start of the COVID-19 pandemic, which likely influenced the study results in many ways. The COVID-19 pandemic was characterized by social distancing, immunization, and patient care patterns that changed rapidly and dramatically during its different phases. Our study was not designed to capture the nuances of this effect. However, the impact of the COVID-19 pandemic on many parts of our study results cannot be denied. We did not adjust *P* values for multiple comparisons of the descriptive analyses, and we prespecified the SAs that we present in this study.

Our study also has several strengths. We examined the SLE PtDA at several US clinics, enrolled a large sample of patients before and during the COVID-19 pandemic, included several covariates and potential confounders, examined both SDM outcomes and patient-reported outcomes (PROs), and tested the robustness of our findings by performing multiple SAs.

In conclusion, we completed a multicenter implementation study during which an evidence-based effective SLE PtDA was provided to patients in two settings (home versus in clinic) using three potential modalities (a smartphone, touchpad tablet, or computer). We found no differences in SDM outcomes, patient feasibility, acceptability, and satisfaction by setting or modality of SLE PtDA delivery. Our trial provides evidence to further support regular and appropriate use of the SLE PtDA by offering it to SLE patients during routine care in different settings via multiple modalities. Complete viewing and uninterrupted viewing of the SLE PtDA were associated with SDM outcomes and other PROs. Studies are needed to enhance ways for uninterrupted and complete viewing of the SLE PtDA for widespread dissemination.

ACKNOWLEDGMENTS

We thank Implementing Decision-Aid for Lupus in Clinics (IDEAL) site principal investigators, site coordinators, and clinic personnel who were part of the implementation team at each clinic site and patient participants for their contributions, time, and participation, without whom this study would not be possible.

AUTHOR CONTRIBUTIONS

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

REFERENCES

1. Barber MRW, Drenkard C, Falasinnu T, et al. Global epidemiology of systemic lupus erythematosus. Nat Rev Rheumatol 2021;17(9):515–532.
2. Fava A, Petri M. Systemic lupus erythematosus: diagnosis and clinical management. J Autoimmun 2019;96:1–13.
3. Durcan L, O'Dwyer T, Petri M. Management strategies and future directions for systemic lupus erythematosus in adults. Lancet 2019;393(10188):2332–2343.
4. Fanouriakis A, Tziolos N, Bertsias G, et al. Update on the diagnosis and management of systemic lupus erythematosus. Ann Rheum Dis 2021;80(1):14–25.
5. Alarcón GS, McGwin G Jr, Petri M, et al; PROFILE Study Group. Baseline characteristics of a multiethnic lupus cohort: PROFILE. Lupus 2002;11(2):95–101.
6. Costenbader KH, Desai A, Alarcón GS, et al. Trends in the incidence, demographics, and outcomes of end-stage renal disease due to lupus nephritis in the US from 1995 to 2006. Arthritis Rheum 2011;63(6):1681–1688.
7. Kasitanon N, Magder LS, Petri M. Predictors of survival in systemic lupus erythematosus. Medicine (Baltimore) 2006;85(3):147–156.
8. Chambers SA, Raine R, Rahman A, et al. Why do patients with systemic lupus erythematosus take or fail to take their prescribed medications? A qualitative study in a UK cohort. Rheumatology (Oxford) 2009;48(3):266–271.
9. Kumar K, Gordon C, Barry R, et al. 'It's like taking poison to kill poison but I have to get better': a qualitative study of beliefs about medicines in rheumatoid arthritis and systemic lupus erythematosus patients of South Asian origin. Lupus 2011;20(8):837–844.
10. Hargraves I, LeBlanc A, Shah ND, et al. Shared decision making: the need for patient-clinician conversation, not just information. Health Aff (Millwood) 2016;35(4):627–629.
11. Légaré F, Turcotte S, Stacey D, et al. Patients' perceptions of sharing in decisions: a systematic review of interventions to enhance shared decision making in routine clinical practice. Patient 2012;5(1):1–19.

12. Stacey D, Légaré F, Lewis K, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2017;4(4):CD001431.
13. Scalia P, Durand MA, Berkowitz JL, et al. The impact and utility of encounter patient decision aids: systematic review, meta-analysis and narrative synthesis. Patient Educ Couns 2019;102(5):817–841.
14. Elwyn G, Pickles T, Edwards A, et al. Supporting shared decision making using an Option Grid for osteoarthritis of the knee in an interface musculoskeletal clinic: a stepped wedge trial. Patient Educ Couns 2016;99(4):571–577.
15. Oostendorp LJ, Ottevanger PB, van der Graaf WT, et al. Exploring the routine administration of decision aids for palliative chemotherapy to patients by nurses: a qualitative study among nurses and oncologists. Cancer Nurs 2015;38(4):270–278.
16. Nestor Maslej, Loredana Fattorini, Raymond Perrault, et al. The AI Index 2024 Annual Report. Stanford Institute for Human-Centered Artificial Intelligence, 2024. Accessed July 22, 2024.
17. Singh JA, Fraenkel L, Green C, et al. Individualized decision aid for diverse women with lupus nephritis (IDEA-WON): a randomized controlled trial. PLoS Med 2019;16(5):e1002800.
18. Singh JA, Hearld LR, Hall AG, et al. Implementing the DEcision-Aid for Lupus (IDEAL): study protocol of a multi-site implementation trial with observational, case study design. Implement Sci Commun 2021; 2(1):30.
19. Singh JA, Shah N, Green C. Individualized patient decision-aid for immunosuppressive drugs in women with lupus nephritis: study protocol of a randomized, controlled trial. BMC Musculoskelet Disord 2017;18(1):53.
20. Robinson K. SMILE, an Online Lupus Tool, Can Aid Shared Decision Making. The Rheumatologist. Published March 6, 2024. https://www.the-rheumatologist.org/article/the-smile-online-lupus-tool-can-aid-shared-decision-making/
21. Singh JA, Hearld LR, Eisen S, et al. Patient outcomes from implementing a shared decision-making aid for systemic lupus erythematosus: a prospective implementation study. Lancet Rheumatol 2025. doi:10.1016/S2665-9913(25)00130-4
22. Singh JA, Hearld LR, Eisen S, et al. Implementing DEcision-Aid for Lupus in clinics (IDEAL) Consortium. Implementation outcomes of a patient decision-aid in a diverse population with systemic lupus erythematosus in 15 US rheumatology clinics. Rheumatology (Oxford) 2025;64(8):4631–4640.
23. Qu H, Shewchuk RM, Alarcón G, et al. Mapping perceptions of lupus medication decision-making facilitators: the importance of patient context. Arthritis Care Res (Hoboken) 2016;68(12):1787–1794.
24. Singh JA, Qu H, Yazdany J, et al. Barriers to medication decision making in women with lupus nephritis: a formative study using nominal group technique. J Rheumatol 2015;42(9):1616–1623.
25. Singh JA, Qu H, Yazdany J, et al. Minorities with lupus nephritis and medications: a study of facilitators to medication decision-making. Arthritis Res Ther 2015;17(1):367.
26. Singh JA, Hossain A, Kotb A, et al. Treatments for lupus nephritis: a systematic review and network metaanalysis. J Rheumatol 2016;43(10): 1801–1815.
27. Singh JA, Hossain A, Kotb A, et al. Risk of serious infections with immunosuppressive drugs and glucocorticoids for lupus nephritis: a systematic review and network meta-analysis. BMC Med 2016; 14(1):137.
28. Singh JA, Hossain A, Kotb A, et al. Comparative effectiveness of immunosuppressive drugs and corticosteroids for lupus nephritis: a systematic review and network meta-analysis. Syst Rev 2016; 5(1):155.
29. Elwyn G, O'Connor A, Stacey D, et al; International Patient Decision Aids Standards (IPDAS) Collaboration. Developing a quality criteria framework for patient decision aids: online international Delphi consensus process. BMJ 2006;333(7565):417.
30. Hibbard JH, Slovic P, Peters E, et al. Strategies for reporting health plan performance information to consumers: evidence from controlled studies. Health Serv Res 2002;37(2):291–313.
31. O'Connor AM. User Manual - Decisional Conflict Scale. The Ottawa Hospital Research Institute. 1993. Updated 2010. http://decisionaid.ohri.ca/docs/develop/User_Manuals/UM_Decisional_Conflict.pdf.
32. Cohan D, Gomez E, Greenberg M, et al. Patient perspectives with abbreviated versus standard pre-test HIV counseling in the prenatal setting: a randomized-controlled, non-inferiority trial. PLoS One 2009;4(4):e5166.
33. O'Connor AM. Validation of a decisional conflict scale. Med Decis Making 1995;15(1):25–30.
34. Brehaut JC, O'Connor AM, Wood TJ, et al. Validation of a decision regret scale. Med Decis Making 2003;23(4):281–292.
35. O'Connor AM, Bennett CL, Stacey D, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2009;(3):CD001431.
36. Stacey D, Bennett CL, Barry MJ, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2011;(10):CD001431.
37. Sun Q. Predicting Downstream Effects of High Decisional Conflict: Meta-analyses of the Decisional Conflict Scale. Master of Science thesis. University of Ottawa; 2005.
38. Thompson-Leduc P, Turcotte S, Labrecque M, et al. Prevalence of clinically significant decisional conflict: an analysis of five studies on decision-making in primary care. BMJ Open 2016;6(6):e011490.
39. Bennett C, Graham ID, Kristjansson E, et al. Validation of a preparation for decision making scale. Patient Educ Couns 2010;78(1):130–133.
40. Elwyn G, Barr PJ, Grande SW, et al. Developing CollaboRATE: a fast and frugal patient-reported measure of shared decision making in clinical encounters. Patient Educ Couns 2013;93(1):102–107.
41. O'Connor AM. User Manual – Acceptability. The Ottawa Hospital Research Institute. 1996. Updated 2002. Accessed December 21, 2017. http://decisionaid.ohri.ca/docs/develop/User_Manuals/UM_Acceptability.pdf. Ottawa Hospital Research Institute.
42. Mudano AS, Gary LC, Oliveira AL, et al. Using tablet computers compared to interactive voice response to improve subject recruitment in osteoporosis pragmatic clinical trials: feasibility, satisfaction, and sample size. Patient Prefer Adherence 2013;7:517–523.
43. Hess TJ, Fuller MA, Mathew J. Involvement and Decision-Making Satisfaction with a Decision Aid: The Influence of Social Multimedia, Gender, and Playfulness. In: Burstein F, Holsapple C. Handbook on Decision Support Systems 1. Springer Berlin Heidelberg; 2008:731–761.
44. Mhaimeed N, Mhaimeed N, Mhaimeed O, et al. Shared decision making with black patients: a scoping review. Patient Educ Couns 2023; 110:107646.
45. Ashana DC, Welsh W, Preiss D, et al. Racial differences in shared decision-making about critical illness. JAMA Intern Med. 2024; 184(4):424–432.
46. Lane GI, Dupati A, Qi J, et al. Factors associated with decision aid use in localized prostate cancer. Urol Pract 2022;9(1):108–115.
47. Lim SS, Helmick CG, Bao G, et al. Racial disparities in mortality associated with systemic lupus erythematosus - Fulton and DeKalb Counties, Georgia, 2002-2016. MMWR Morb Mortal Wkly Rep. 2019;68(18):419–422.
48. Maningding E, Dall'Era M, Trupin L, et al. Racial and ethnic differences in the prevalence and time to onset of manifestations of systemic lupus erythematosus: the California Lupus Surveillance Project. Arthritis Care Res (Hoboken) 2020;72(5):622–629.
49. Drenkard C, Parker S, Aspey LD, et al. Racial disparities in the incidence of primary chronic cutaneous lupus erythematosus in the

southeastern US: The Georgia Lupus Registry. Arthritis Care Res (Hoboken) 2019;71(1):95–103.
50. Carroll NW, Karabukayeva A, Hearld LR, et al. Time required to implement a computerized patient decision aid for lupus in outpatient visits. Implement Sci Commun 2025;6(1):41.

APPENDIX A: IMPLEMENTING DECISION-AID FOR LUPUS IN CLINICS CONSORTIUM AUTHORS

Members of the Implementing Decision-Aid for Lupus in Clinics (IDEAL) Consortium are as follows (site name listed from highest to lowest enrollment):

University of Alabama at Birmingham, Drs Winn Chatham and Jasvinder Singh;
Northwell Hospital, Dr Sonali Narain;
Vanderbilt University, Dr Narender Annapureddy;
Medical University of South Carolina, Dr Diane Kamen;
University of Chicago, Dr Kimberly Trotter;
University of Mississippi Medical Center, Drs Vikas Majithia and Cathy Ching;
Loyola University, Dr Zineb Aouhab;
Cedars-Sinai Medical Center, Dr Swamy Venturupalli;
Northwestern University, Dr Rosalind Ramsey-Goldman;
Washington University School of Medicine, Dr Alfred Kim;
University of California at Los Angeles, Dr Maureen McMahon;
Emory University, Dr Sam Lim;
Baylor College of Medicine, Dr Kalpana Bhairavarasu;
Ohio State University, Dr Alexa Meara;
University of California at San Diego, Dr Kenneth Kalunian.

Arthritis Care & Research
Vol. 77, No. 12, December 2025, pp 1487–1492
DOI 10.1002/acr.25604

American College of Rheumatology
Empowering Rheumatology Professionals

BRIEF REPORT

Use of Apremilast for the Treatment of Immune Checkpoint Inhibitor Psoriasis and Psoriatic Arthritis

Nilasha Ghosh,[1] Pankti Reid,[2] Jeffrey A. Sparks,[3] Kaitlin McCarter,[3] Kyle Ge,[4] and Anne R. Bass[1]

Objective. The objective of this study was to present effectiveness and tolerability of apremilast in a cohort of 21 patients with immune checkpoint inhibitor psoriatic arthritis (ICI-PsA) and/or immune checkpoint inhibitor psoriasis (ICI-PsO).

Methods. This multicenter study combined data from patients treated with apremilast after experiencing ICI-PsO and/or ICI-PsA. Patients taking apremilast before ICI initiation and patients with preexisting autoimmune disease before ICI therapy were also included. Response to apremilast was determined as complete, partial, or none as determined by improvement in Common Terminology Criteria for Adverse Events grading after drug initiation.

Results. There were 21 patients who used apremilast for either ICI-PsO and/or ICI-PsA, but only five of these patients had de novo ICI-PsO and/or ICI-PsA. Of these five patients, four had partial response or improvement in their immune-related adverse event with apremilast, although there were intolerances in three of these patients. Of the 21 total patients, 16 had a relevant preexisting autoimmune disease, indicating a likely flare of the underlying disease with ICI therapy. Flares occurred much sooner for patients with ICI-PsA (4 weeks) compared to patients with ICI-PsO only (39.7 weeks), although the majority of both groups had grade II severity. Among the 13 patients with preexisting disease and no exposure to apremilast before ICI therapy, all patients in the ICI-PsO–only group (100%) responded to apremilast with either a complete or partial response, whereas only 57% of patients in the ICI-PsA group had complete or partial response. Twenty-nine percent of patients in the entire cohort had to discontinue apremilast due to intolerability. Thirty-eight percent of the entire cohort had progression of cancer or death at last follow-up after being on apremilast.

Conclusion. This study highlights the potential benefit of apremilast for the treatment of ICI-PsO, both de novo and PsO flare, with less of an apparent benefit for ICI-PsA. Thirty percent of patients in the whole group had to discontinue apremilast due to intolerance. Apremilast may be an attractive therapeutic option for either condition given that it is not immunosuppressive, but further prospective observational studies with larger patient numbers are needed.

INTRODUCTION

Immune-related adverse events (irAEs) occur commonly after immune checkpoint inhibitors (ICIs), which are used for the treatment of many different types and stages of cancers, including locally advanced and metastatic cancers.[1] Rheumatic irAEs, such as inflammatory arthritis (IA), myositis, and sicca syndrome, account for about 5% to 10% of all irAEs.[2,3] However, unlike many other irAEs, rheumatic irAEs, particularly IA, tend to persist for longer periods of time, sometimes even after ICI has been

Supported by Clinical Translational Science Award (grant 1UL1-TR-002541-01) to Harvard University and Brigham and Women's Hospital from the National Center for Research Resources. The Hospital for Special Surgery received support from Clinical and Translational Science Center (CTSC) (grant UL1-TR-002384). Dr Reid's work was supported by the University of Chicago Institute of Translational Medicine Clinical and Translational Science Award K12/KL2 (grant 5KL2-TR-002387-05). Dr Sparks's work was supported by the National Institute of Arthritis and Musculoskeletal and Skin Diseases, NIH (grants R01-AR-080659, R01-AR-077607, P30-AR-070253, and P30-AR-072577) and National Heart, Lung, and Blood Institute, NIH (grant R01-HL-155522), the Rheumatology Research Foundation, the Arthritis Foundation, the R. Bruce and Joan M. Mickey Research Scholar Fund, and the Llura Gund Award funded by the Gordon and Llura Gund Foundation.

[1]Nilasha Ghosh, MD, MS, Anne R. Bass, MD: Hospital for Special Surgery and Weill Cornell Medical College, New York City, New York; [2]Pankti Reid, MD, MPH: University of Chicago, Chicago, Illinois; [3]Jeffrey A. Sparks, MD, MS, Kaitlin McCarter, MD: Brigham and Women's Hospital, Boston, Massachusetts; [4]Kyle Ge, BS: Hospital for Special Surgery, New York City, New York.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25604.

Address correspondence via email to Nilasha Ghosh, MD, MS, at ghoshn@hss.edu.

Submitted for publication April 1, 2025; accepted in revised form July 1, 2025.

SIGNIFICANCE & INNOVATIONS

- This study highlights data about effectiveness and tolerability of apremilast for the treatment of immune checkpoint inhibitor psoriasis (ICI-PsO) and/or immune checkpoint inhibitor psoriatic arthritis (ICI-PsA) from the largest collected cohort of patients published thus far.
- All patients in the ICI-PsO–only group had either complete or partial response, whereas half of the patients with ICI-PsA had no response at all.
- Almost one-third of patients had to discontinue apremilast prematurely due to intolerability, such as gastrointestinal side effects.

discontinued.[4,5] Thus, there has been an unmet need to find the most appropriate treatment for ICI-IA, especially a treatment that would not alter the efficacy of the ICI and tumor response.

So far, the most used medications for the treatment of ICI-IA include glucocorticoids, disease-modifying antirheumatic drugs, such as hydroxychloroquine and methotrexate, and biologics, such as tumor necrosis factor inhibitors or interleukin-6 inhibitors.[6–8] However, there are little data available about other treatments. Apremilast is a phosphodiesterase-4 inhibitor that is approved for the treatment of psoriasis (PsO), psoriatic arthritis (PsA), and oral ulcers from Behçet disease. Its use for the treatment of irAEs from ICI, specifically ICI-PsO and ICI-PsA, has not been widely published. Incidence of ICI-PsO has been reported in a pharmacovigilance database as 4%, of which 70% of patients had preexisting PsO.[9] The incidence of ICI-PsA is more difficult to ascertain given heterogenous presentations. However, it has been suggested that an oligoarticular IA resembling PsA or another spondylitis variant could account for one-fifth of ICI-IA cases and that the reported incidence of ICI-IA could be about 4% of all irAEs.[6,10] Theoretically, apremilast would serve as an attractive choice for the treatment of either ICI-PsO and/or ICI-PsA given that it does not suppress the immune system or potentially "counteract" the efficacy of the checkpoint inhibitor's enhanced T cell activity because it likely works downstream at the cytokine level. A recent systematic literature review from a dermatology group identified 16 individual cases of patients with ICI-PsO treated with apremilast and only two patients with concomitant ICI-PsO and ICI-PsA.[11] Here, we present the largest cohort of apremilast-treated patients with ICI-PsA and/or ICI-PsO collected from registries of patients with rheumatic irAEs, highlighting the effectiveness and tolerability of apremilast.

PATIENTS AND METHODS

This was a multicenter, observational study from three different major US academic medical centers (Hospital for Special Surgery [HSS]/Weill Cornell Medical College, University of Chicago Medical Center, and Brigham and Women's Hospital), each with a dedicated rheumatic irAE registry (one prospective registry and two retrospective registries). To be included, patients had to have received an ICI for cancer and had to have been treated with apremilast after ICI initiation for any relevant immune-related indication. Patients taking apremilast before ICI initiation and patients with preexisting autoimmune disease before ICI therapy were included, although they were singled out given potential confounding. Patients taking treatments specifically for other irAEs were excluded to avoid attribution conflicts. The study was approved by the HSS Internal Review Board, and the other individual institutions and data transfer agreements were obtained. Deidentified data were sent to the coordinating site (HSS) for analysis. The dataset was closed with an end date of October 31, 2024. Variables gathered included demographics, cancer type and stage, ICI (+/– chemotherapy) regimen, irAE time of onset, other irAEs experienced, severity per Common Terminology Criteria for Adverse Events (CTCAE),[12] response and tolerability to apremilast for irAE treatment, concurrent therapies, cancer status after apremilast initiation, and follow-up time. Response to apremilast was determined as change in CTCAE grading: complete response was an improvement to grade 0 (no disease), partial response was any improvement in grading that did not achieve grade 0, or none was as determined by the treating rheumatologist. Descriptive statistics were used to describe variables.

RESULTS

Study sample. We identified 23 patients who used apremilast for an irAE indication: 21 used it for either ICI-PsO or ICI-PsA (Table 1), and two used it for oral ulcers and lichen planus, who were excluded from analysis. We report the outcomes of the 21 patients with ICI-PsO and/or ICI-PsA: 10 used it for ICI-PsO alone, three for ICI-PsA alone, and eight for ICI-PsA and ICI-PsO. Mean age for the entire cohort was 59 years (SD 12), 71% were men, and 90% of the cohort was White. Melanoma and lung and genitourinary cancers were most prevalent, and about half the patients were being treated for stage IV cancer. Sixty-two percent of this cohort used anti–programmed death 1 and/or anti–programmed death ligand 1 (anti-PD1/PDL-1) monotherapy, and the rest used combination therapy (anti-PD1 plus anti–CTLA-4). Only five patients did not have any preexisting autoimmune disease, indicating de novo ICI-PsO and/or ICI-PsA (Table 2, top of the table). Seventy-six percent of patients (16 patients) had a preexisting autoimmune disease such as PsO, PsA, or inflammatory bowel disease who likely experienced a flare of their underlying disease with ICI. Of these patients, one patient (patient 11) was on apremilast for PsO and PsA at baseline before ICI initiation, and two other patients (patients 15 and 18) were on the drug before ICI for PsO. Eliminating these three individuals left 13 patients who experienced a flare of disease, either

Table 1. Patient demographics*

Patient number	Age, years	Sex	Race or ethnicity	Cancer/stage	ICI + chemotherapy regimen	Total ICI duration, days	Other irAEs experienced	ICI discontinuation reason
1	49	F	White	Breast/IV	Pembrolizumab + paclitaxel/carboplatin	91	None	Therapy completed
2	61	M	White	Melanoma/IV	Ipilimumab/nivolumab	98	Vitiligo, colitis, pneumonitis	Therapy completed
3	62	M	White	Bladder/IV	Atezolizumab + etoposide/carboplatin	125	None	Cancer progression
4	72	M	White	Hepatocellular/IIIC	Durvalumab/tremelimumab → durvalumab → atezolizumab + bevacizumab	903	Hyperthyroidism	Cancer progression
5	59	M	White	Melanoma/IV	Nivolumab → ipilimumab/nivolumab	939	Colitis	Cancer progression
6	46	F	White	Melanoma/IV	Ipilimumab/nivolumab	456	Adrenal insufficiency, rash, colitis	Therapy completed
7	41	M	White	Melanoma/IIIB	Nivolumab/relatlimab	112	Colitis	Ongoing
8	62	F	White	Endometrial/IIIC	Pembrolizumab	165	Hypothyroidism	Therapy completed
9	77	M	White	Mesothelioma/II	Ipilimumab/nivolumab	518	None	irAE
10	70	M	Hispanic	RCC/IV	Ipilimumab/nivolumab	259	None	irAE
11	71	M	White	Bladder/II	Pembrolizumab	1	None	irAE
12	70	M	White	RCC/IV	Nivolumab	1,505	None	Ongoing
13	48	M	White	Lung /IV	Ipilimumab/nivolumab → nivolumab	1,773	None	Ongoing
14	48	M	White	H&N/III	Nivolumab + paclitaxel/carboplatin	355	None	irAE
15	47	F	Black	H&N/II	Pembrolizumab + docetaxel/carboplatin	41	None	Therapy completed
16	53	M	White	Skin SCC/unknown	Cemiplimab	336	Adrenal insufficiency	Therapy completed
17	73	M	White	Lung/IV	Pembrolizumab + paclitaxel/carboplatin → nivolumab	657	None	Cancer progression/deceased
18	78	M	White	Lung/IA	Ipilimumab/nivolumab	21	None	Cancer progression/deceased
19	46	F	White	Melanoma/IV	Ipilimumab/nivolumab → nivolumab	462	None	irAE
20	61	M	White	H&N/IV	Nivolumab	91	Acute interstitial nephritis	Cancer progression
21	52	F	White	Lung/IV	Pembrolizumab + carboplatin/pemetrexed	126	Pneumonitis	irAE/deceased

* F, female; H&N, head and neck; ICI, immune checkpoint inhibitor; irAE, immune-related adverse event; M, male; RCC, renal cell carcinoma; SCC, squamous cell carcinoma.

PsO, PsA, or PsA and PsO (Table 3). Median onset of the PsO and/or PsA flare for this group was 20.4 weeks from ICI initiation, but it was 39.7 weeks for the ICI-PsO–only group and 4 weeks for the ICI-PsA group. Most of the patients experienced grade II irAE per the CTCAE rubric.

Effectiveness and tolerability. Among the patients with de novo ICI-PsO and ICI-PsA (Table 2, top), there were four out of five patients with partial improvement with apremilast, although apremilast was eventually stopped either due to intolerability (three patients) or being ultimately ineffective long term (one patient). One patient that had improvement in PsA was also on methotrexate. Among the patients with preexisting disease and no exposure to apremilast before ICI (Table 3), all patients in the ICI-PsO–only group (100%) responded to apremilast as either having complete or partial response, whereas only 57% of patients in the ICI-PsA group had complete or partial response. Of note, apremilast was not effective treatment for arthritis for the three patients with preexisting PsA (patients 7, 8, and 11). Twenty-nine percent of patients in the entire cohort had to discontinue apremilast due to intolerability, such as diarrhea (four patients) and nausea (two patients). Cancer status at the last follow-up for the entire cohort revealed progression or death for eight patients (38%) and was similar for patients with flares of underlying PsO only and PsA, although numbers were small. Apremilast was not effective for either the patient with oral ulcers

Table 2. Clinical presentation of immune-related adverse event and apremilast efficacy and tolerability*

Patient number	Clinical presentation	CTCAE grade (joints/skin)	Preexisting history[a]	irAE onset after ICI, weeks	Concurrent therapies used with apremilast	Time on apremilast, weeks	Response to apremilast (joints/skin)	Reasons for stopping apremilast	Cancer status at last follow-up after apremilast	Follow-up time after apremilast, weeks
1	Peripheral arthritis	2	No	18	MTX, prednisone	11	None	Intolerable	CR	32
2	Peripheral arthritis, dactylitis	2	No	15	MTX	43	Partial	Improved	CR	43
3	10% BSA	3	No	18	Topical	13	Partial	Ineffective	Progression	12
4	Gluteal cleft	2	No	102	Topical	36	Partial	Intolerable	Progression	42
5	Peripheral arthritis + 10% BSA	2/2	No	6	Prednisone	5	Partial/partial	Intolerable	Progression	79
6	Peripheral arthritis	2	UC	137	None	17	None	Ineffective	CR	23
7	Peripheral arthritis, axial disease, tendonitis, + scalp PsO	2/1	PsA, PsO	4	None	5	None/none	Intolerable	PR	30
8	Peripheral arthritis + PsO	2/1	PsA, PsO	26	Prednisone	13	None/none	Intolerable	CR	59
9	Peripheral arthritis + 30% BSA	2/3	PsO	22	SSZ	6	Partial/none	Ineffective	Stable	68
10	Peripheral arthritis + 70% BSA	2/3	PsO	3	UV	31	Partial/none	Ineffective	Stable	283
11	Peripheral arthritis + <10% BSA	1/2	PsA, PsO	2	Prednisone	223 (started before ICI)	None/partial	Ineffective	Progression	223
12	10% BSA	2	PsO	8	Topical	22	Partial	Intolerable	Stable	98
13	<1% BSA, nails	2	PsO	61	None	99	None	Ineffective	Stable	133
14	10% BSA	2	PsO	45	Topical	19	Complete	irAE resolution	CR	4
15	30% BSA	2	PsO	16	Topical	7 (started before ICI)	Partial	irAE resolution	PR	37
16	Peripheral arthritis + 50% BSA	3/3	PsO	2	Topical	211	Partial/partial	Improved	Stable	211
17	10% BSA	2	PsO	12	Topical, UV, prednisone	85	Complete	irAE resolution	Progression/death	84
18	10% BSA	2	PsO	Present before ICI	None	97 (started before ICI)	Partial	Ongoing	Progression/death	97
19	15% BSA	2	PsO	77	Topicals, prednisone	Unknown	Partial	irAE resolution	CR	350
20	Peripheral arthritis + <10% BSA	1/2	PsO	1	Topical	60	Complete/complete	irAE resolution	Progression	57
21	Single plaque	1	PsO	34	Topical	17	Partial	Ineffective	Death	25

* BSA, body surface area; CR, complete response; CTCAE, Common Terminology Criteria for Adverse Events; ICI, immune checkpoint inhibitor; irAE, immune-related adverse event; MTX, methotrexate; PR, partial response; PsA, psoriatic arthritis; PsO, psoriasis; SSZ, sulfasalazine; UC, ulcerative colitis; UV, UV light therapy.
[a] Patients 1–5 are without preexisting autoimmune disease.

Table 3. Cohort-level data of preexisting autoimmune disease flares*

Characteristics	All patients (n = 13)	PsO only (n = 6)	PsA ± PsO (n = 7)
Age, mean (SD), years	57.5 (12)	56.2 (12)	58.6 (12.8)
Male, %	69	67	71
White, %	69	100	86
Cancer type, n (%)			
Melanoma	3 (23)	1 (17)	2 (29)
Lung	3 (23)	3 (50)	0 (0)
RCC/bladder	2 (15)	1 (17)	1 (14)
H&N	2 (15)	1 (17)	1 (9)
Skin (SCC)	1 (8)	0 (0)	1 (9)
Endometrial	0 (0)	0 (0)	1 (14)
Mesothelioma	0 (0)	0 (0)	1 (9)
Other	2 (15)	0 (0)	0 (0)
Stage, n (%)			
Stage IV	8 (62)	5 (83)	3 (43)
Stage III	3 (23)	1 (17)	2 (29)
Stage II	1 (8)	0 (0)	1 (14)
Unknown	1	0 (0)	1
ICI regimen, n (%)			
Monotherapy	8 (62)	4 (67)	4 (57)
Combination	5 (38)	2 (33)	3 (43)
Median onset of irAE, weeks	22	39.7	4
CTCAE Grade irAE,[a] n			
Grade I	2	1	1
Grade II	10	5	5
Grade III	1	0	1
Median duration of apremilast use, weeks	20.4	21.9	17.3
Response to apremilast,[a] n (%)			
Complete	3 (23)	2 (33)	1 (14)
Partial	7 (54)	4 (67)	3 (43)
None	3 (23)	0 (0)	3 (43)
Discontinuation of apremilast due to intolerability, n (%)	3 (23)	1 (17)	2 (29)
Cancer status at last follow-up, n (%)			
CR	4 (31)	2 (33)	2 (29)
PR	1 (8)	0 (0)	1 (14)
Stable	5 (38)	2 (33)	3 (43)
Progression or death	3 (23)	2 (33)	1 (14)

* CR, complete response; CTCAE, Common Terminology Criteria for Adverse Events; H&N, head and neck; ICI, immune checkpoint inhibitor; irAE, immune-related adverse event; PR, partial response; PsA, psoriatic arthritis; PsO, psoriasis; RCC, renal cell carcinoma; SCC, squamous cell carcinoma.
[a] For specific irAE only (PsO or PsA).

or the one with lichen planus (excluded from analysis), and it was discontinued for both due to inefficacy.

This is the largest collected cohort of patients who were treated with apremilast for an irAE, primarily ICI-PsO and ICI-PsA, including those with preexisting disease and/or flare of underlying disease. Of those without preexisting disease, apremilast delivered mostly partial responses, although there were a fair number of gastrointestinal intolerances requiring discontinuation. Of the rest of the patients, over half had preexisting PsO, and three patients had preexisting PsA, suggesting that their irAE was a flare of their autoimmune disease, which is line with previous literature.[9,13] These patients are inherently different than the patients that develop de novo irAE. Apremilast was not helpful for those that experienced flares of their previous PsA, but there was greater clinical response seen for patients with ICI-PsO (100%). Unfortunately, these numbers are too small to draw significant conclusions from, but they do suggest a potential benefit for skin efficacy over joint efficacy. Adverse events in this cohort were similar to previous studies of apremilast in non-ICI settings,[14] although drug discontinuation was higher in this unique population with an advanced malignancy. This could be because they were experiencing other additive side effects from their cancer treatments, such as side effects from concurrent chemotherapies or other irAEs.

As in the study by Kaur et al,[11] all patients with skin manifestations had complete resolution or significant clinical improvement on apremilast. That study did not clearly report the response of the patients with ICI-PsA only, although it implies there was some improvement in disease activity. Although patient numbers are small for both of our studies, there was numerically less cancer progression among their collected cases relative to our cohort. This could reflect the limitation of literature review data and highlight the longer and more detailed follow-up data our registry patients provided. Neither study can broach long-term safety of

apremilast in regard to cancer progression given the small numbers of patients and other confounding variables, such as variable baseline characteristics, concurrent medication, and immortal time bias, which would make attribution of progression to the drug indeterminable. Another limitation of the paper is the determination of effectiveness, which was documented by the treating physician as a change in the CTCAE score. This can be seen as subjective, although this practice is commonly used in oncologic studies as well as studies involving irAEs. Still, a validated measure of an outcome measure such as an American College of Rheumatology 50% improvement criteria (ACR50) response for joint involvement or a Psoriasis Area and Severity Index score for skin involvement could be used in future prospective observational studies for more enhanced effectiveness outcomes.

In summary, this study highlights the benefit of apremilast for the treatment of ICI-PsO, including those with preexisting disease and/or flare, with less of a benefit for ICI-PsA. Almost one-third of patients in both groups had to discontinue apremilast due to intolerance, which may be attributed to the patient population with malignancy being treated with other therapies and experiencing other irAE. Nonetheless, apremilast remains an attractive therapeutic option, even for ICI-PsA, given that it is not immunosuppressive and therefore not thought to impair cancer responses to ICI.

AUTHOR CONTRIBUTIONS

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

REFERENCES

1. Postow MA, Callahan MK, Wolchok JD. Immune checkpoint blockade in cancer therapy. J Clin Oncol 2015;33(17):1974–1982.
2. Angelopoulou F, Bogdanos D, Dimitroulas T, et al. Immune checkpoint inhibitor-induced musculoskeletal manifestations. Rheumatol Int 2021;41(1):33–42.
3. Cappelli LC, Gutierrez AK, Bingham CO III, et al. Rheumatic and musculoskeletal immune-related adverse events due to immune checkpoint inhibitors: a systematic review of the literature. Arthritis Care Res (Hoboken) 2017;69(11):1751–1763.
4. Braaten TJ, Brahmer JR, Forde PM, et al. Immune checkpoint inhibitor-induced inflammatory arthritis persists after immunotherapy cessation. Ann Rheum Dis 2020;79(3):332–338.
5. Ladouceur A, Barnetche T, Prey S, et al. Resolution of immune checkpoint inhibitors-induced inflammatory arthritis while maintaining active treatment with checkpoint inhibitors and after its discontinuation: an observational study. Joint Bone Spine 2025;92(1):105795.
6. Ghosh N, Tiongson MD, Stewart C, et al. Checkpoint inhibitor–associated arthritis: a systematic review of case reports and case series. J Clin Rheumatol 2021;27(8):e317–e322.
7. Bass AR, Abdel-Wahab N, Reid PD, et al. Comparative safety and effectiveness of TNF inhibitors, IL6 inhibitors and methotrexate for the treatment of immune checkpoint inhibitor-associated arthritis. Ann Rheum Dis 2023;82(7):920–926.
8. Chan KK, Tirpack A, Vitone G, et al. Higher checkpoint inhibitor arthritis disease activity may be associated with cancer progression: results from an observational registry. ACR Open Rheumatol 2020;2(10):595–604.
9. Cutroneo P, Ingrasciotta Y, Isgrò V, et al. Psoriasis and psoriasiform reactions secondary to immune checkpoint inhibitors. Dermatol Ther 2021;34(2):e14830.
10. Kostine M, Rouxel L, Barnetche T, et al; FHU ACRONIM. Rheumatic disorders associated with immune checkpoint inhibitors in patients with cancer-clinical aspects and relationship with tumour response: a single-centre prospective cohort study. Ann Rheum Dis 2018;77(3):393–398.
11. Kaur M, Feldman AE, Dulmage B, et al. Apremilast for immune checkpoint inhibitor-associated psoriasis and psoriatic arthritis: a systematic review. JAAD Rev 2024;2:149–152.
12. CTCAE and Adverse Event Reporting. National Cancer Institute Division of Cancer Treatment & Diagnosis. Accessed August 22, 2025. https://dctd.cancer.gov/research/ctep-trials/for-sites/adverse-events
13. Abdel-Wahab N, Suarez-Almazor ME. Frequency and distribution of various rheumatic disorders associated with checkpoint inhibitor therapy. Rheumatology (Oxford) 2019;58(suppl 7):vii40–vii48.
14. Mease PJ, Hatemi G, Paris M, et al. Apremilast long-term safety up to 5 years from 15 pooled randomized, placebo-controlled studies of psoriasis, psoriatic arthritis, and Behçet's syndrome. Am J Clin Dermatol 2023;24(5):809–820.

Arthritis Care & Research
Vol. 77, No. 12, December 2025, pp 1493–1495

AMERICAN COLLEGE of RHEUMATOLOGY
Empowering Rheumatology Professionals

LETTER

DOI 10.1002/acr.25545

Lower or higher 25-hydroxyvitamin D levels associated with adverse pregnancy outcomes: comment on the article by Madanchi et al

To the Editor:

We read with interest the recent publication by Madanchi and colleagues on the association between 25-hydroxyvitamin D (25[OH]D) levels and adverse pregnancy outcomes in systemic lupus erythematosus (SLE).[1] The authors concluded that there is an association of 25(OH)D at both lower and higher levels with adverse pregnancy outcomes. We recommend the monitoring of maternal serum 25(OH)D levels during SLE pregnancies, aiming for the ideal range[1] of 40 to 59 ng/mL. We support and appreciate the authors' work and agree with their conclusions, but we have some concerns about some of the details in the article.

In patients with SLE, pharmacological treatment is a crucial component of disease management. Many patients with SLE require long-term use of multiple medications, including glucocorticoids, immunosuppressants (such as tacrolimus, cyclosporine, mycophenolate mofetil, etc), and other supportive treatment drugs. These medications play a significant role in controlling disease activity and preventing complications. However, they may also affect the metabolism and levels of 25(OH)D. Nevertheless, this article does not investigate the relationship between medication use and 25(OH)D levels, which constitutes an important limitation of the study.[1]

Firstly, glucocorticoids can impact the metabolism of 25(OH)D. Glucocorticoids are commonly used in the treatment of SLE, especially during periods of disease activity. However, glucocorticoids may affect the metabolism of 25(OH)D through various mechanisms. Skversky and colleagues conducted a cross-sectional analysis using 2001 to 2006 National Health and Nutrition Examination Survey data to assess the relationship between serum 25(OH)D deficiency, defined as a 25(OH)D level <10 ng/mL, and the use of oral steroids, with the study results showing that the use of steroids is independently associated with 25(OH)D deficiency.[2] In a cohort study of 124 female patients with SLE, Toloza et al demonstrated by multivariable logistic regression that cumulative glucocorticoid exposure was significantly associated with low levels of 25(OH)D (P = 0.03) when adjusting for ethnicity, season, and serum creatinine level.[3] Animal experiments by Akeno et al confirmed that dexamethasone increased renal expression of vitamin D-24-hydroxylase, which in turn degrades vitamin D metabolites such as 25(OH)D and 1,25-dihydroxyvitamin D.[4] Dhawan and Christakos's research confirmed that glucocorticoids can directly enhance the transcription of the 24-hydroxylase through a new mechanism involving functional cooperation between the glucocorticoid receptor, CCAAT/enhancer binding protein β, and the vitamin D receptor. In summary, steroids can enhance inactivation of 25(OH)D by up-regulating 24-hydroxylase activity.[5]

Secondly, immunosuppressive agents can have a potential impact on 25(OH)D levels. To date, the interaction between immunosuppressive drugs and vitamin D has been extensively studied. For instance, the relationship between calcineurin inhibitors (CNIs) and 25(OH)D metabolism has been reported in numerous studies. The research by Filipov and colleagues suggests that the intake of CNIs is associated with lower 25(OH)D concentrations.[6] The study by Lee et al confirmed that CNIs induce vitamin D resistance.[7] Eyal et al found that tacrolimus and other immunosuppressive drugs have a negative effect on 25(OH)D in renal transplant recipients. A possible explanation for these findings may be the fact that liver CYP3A4 has 25-hydroxylase activity, which is suppressed by cyclosporine and tacrolimus, resulting in lower 25(OH)D levels.[8]

In conclusion, before these issues are clarified, this study's findings should be interpreted cautiously.

We would like to thank the members and staff of the Department of Rheumatology and Immunology of The Second Affiliated Hospital of Soochow University who contributed to this article.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25545.

Gang Wang, MD
Zhichun Liu, MD
zcliurheu@suda.edu.cn
The Second Affiliated Hospital of Soochow University
Suzhou, China

1. Madanchi N, Fava A, Goldman DW, et al. Association between 25-hydroxyvitamin D levels and adverse pregnancy outcomes in systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2025;77(4):432–439.
2. Skversky AL, Kumar J, Abramowitz MK, et al. Association of glucocorticoid use and low 25-hydroxyvitamin D levels: results from the National Health and Nutrition Examination Survey (NHANES): 2001-2006. J Clin Endocrinol Metab 2011;96(12):3838–3845.
3. Toloza SM, Cole DE, Gladman DD, et al. Vitamin D insufficiency in a large female SLE cohort. Lupus 2010;19(1):13–19.
4. Akeno N, Matsunuma A, Maeda T, et al. Regulation of vitamin D-1alpha-hydroxylase and -24-hydroxylase expression by dexamethasone in mouse kidney. J Endocrinol 2000;164(3):339–348.
5. Dhawan P, Christakos S. Novel regulation of 25-hydroxyvitamin D_3 24-hydroxylase (24(OH)ase) transcription by glucocorticoids: cooperative effects of the glucocorticoid receptor, C/EBPβ, and the vitamin D

receptor in 24(OH)ase transcription. J Cell Biochem 2010;110(6):1314–1323.
6. Filipov JJ, Zlatkov BK, Dimitrov EP, et al. Relationship between vitamin D status and immunosuppressive therapy in kidney transplant recipients. Biotechnol Biotechnol Equip 2015;29(2):331–335.
7. Lee CT, Ng HY, Lien YH, et al. Effects of cyclosporine, tacrolimus and rapamycin on renal calcium transport and vitamin D metabolism. Am J Nephrol 2011;34(1):87–94.
8. Eyal O, Aharon M, Safadi R, et al. Serum vitamin D levels in kidney transplant recipients: the importance of an immunosuppression regimen and sun exposure. Isr Med Assoc J 2013;15(10):628–633.

DOI 10.1002/acr.25546

Reply

To the Editor:

In our article,[1] we presented a proportional hazards model that indicated that women with systemic lupus erythematosus with very low 25-hydroxyvitamin D (25[OH]D) levels were at greater risk of having adverse pregnancy outcomes. This model included adjustments for disease activity, renal activity, and race. In their letter, Drs Wang and Liu suggested that the relationship between 25(OH)D and adverse pregnancy outcomes might be confounded by the use of prednisone or immunosuppressants. We analyzed our cohort to determine if this was so.

Of the 260 pregnant patients we analyzed in the cohort, 191 (73%) did not take prednisone during the pregnancy, whereas 17 (7%) took less than 5 mg/day, 29 (11%) took 5 to 9 mg/day, and 23 (9%) took 10 mg/day or more. Most of the patients did not take immunosuppressants during their pregnancy, whereas 70 (27%) took azathioprine, 14 (5%) took tacrolimus, and 1 each took mycophenolate, anakinra, and cyclosporin (we do not allow mycophenolate during pregnancy; exposure was accidental).

The mean 25(OH)D blood levels were not different in those not receiving immunosuppression (39.5 + 12.6 ng/mL), azathioprine (41.1 + 13.5 ng/mL) or tacrolimus (43.4 + 16.6 ng/mL). We took the multivariate model (original model) from our article[1] and compared the original model to the model adjusted for prednisone and immunosuppressive use (Table 1). We found that the association of 25(OH)D blood levels and adverse pregnancy outcomes remained essentially unchanged after adjustment. We thus conclude that the observed relationship between 25(OH)D blood levels and adverse pregnancy outcomes was not confounded by immunosuppressive drug or prednisone use.

Michelle Petri, MD, MPH
Andrea Fava, MD
Daniel W. Goldman, PhD
Nima Madanchi, MD
Johns Hopkins University School of Medicine
Baltimore, MD
Laurence S. Magder, PhD
University of Maryland
Baltimore, MD

1. Madanchi N, Fava A, Goldman DW, et al. Association between 25-hydroxyvitamin D levels and adverse pregnancy outcomes in systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2025;77(4):432–439.

DOI 10.1002/acr.25577

Pain sensitivity and chronic pain as a link between analgesic use and cardiovascular/gastrointestinal risk: comment on the article by Kaur et al

To the Editor:

I read with great interest the article by Kaur et al[1] in *Arthritis Care & Research*, which examined the association between acetaminophen use and the risk of cardiovascular and gastrointestinal complications in older adults. Although the authors conducted rigorous adjustments for osteoarthritis and other comorbidities, it remains difficult to fully account for individual differences in pain sensitivity or the presence of chronic pain, which may influence both analgesic use and disease risk.

Recent studies have demonstrated that individuals with high pain sensitivity or low pain tolerance are at increased risk of cardiovascular disease. For instance, the Tromsø Study showed that participants with low pain tolerance, measured by a cold pressor test, had significantly higher risks of coronary artery disease and all-cause mortality.[2] Furthermore, chronic pain itself has been identified as an independent risk factor for gastrointestinal

Table 1. Adjusted HRs for vitamin D based on three different multivariate models*

Comparison	Original model, HR (95% CI)	Model adjusted for prednisone, HR (95% CI)	Model adjusted for immunosuppressive drugs, HR (95% CI)
<20 vs 40–60 ng/mL	3.35 (1.64–6.80)	3.45 (1.67–7.10)	3.41 (1.67–6.98)
20–30 vs 40–60 ng/mL	1.53 (0.84–2.80)	1.50 (0.82–2.74)	1.56 (0.85–2.86)
30–40 vs 40–60 ng/mL	1.51 (0.91–2.51)	1.53 (0.92–2.54)	1.52 (0.92–2.51)
≥60 vs 40–60 ng/mL	1.34 (0.69–2.62)	1.32 (0.67–2.57)	1.32 (0.68–2.59)

* CI, confidence interval; HR, hazard ratio.

receptor in 24(OH)ase transcription. J Cell Biochem 2010;110(6):1314–1323.
6. Filipov JJ, Zlatkov BK, Dimitrov EP, et al. Relationship between vitamin D status and immunosuppressive therapy in kidney transplant recipients. Biotechnol Biotechnol Equip 2015;29(2):331–335.
7. Lee CT, Ng HY, Lien YH, et al. Effects of cyclosporine, tacrolimus and rapamycin on renal calcium transport and vitamin D metabolism. Am J Nephrol 2011;34(1):87–94.
8. Eyal O, Aharon M, Safadi R, et al. Serum vitamin D levels in kidney transplant recipients: the importance of an immunosuppression regimen and sun exposure. Isr Med Assoc J 2013;15(10):628–633.

DOI 10.1002/acr.25546

Reply

To the Editor:

In our article,[1] we presented a proportional hazards model that indicated that women with systemic lupus erythematosus with very low 25-hydroxyvitamin D (25[OH]D) levels were at greater risk of having adverse pregnancy outcomes. This model included adjustments for disease activity, renal activity, and race. In their letter, Drs Wang and Liu suggested that the relationship between 25(OH)D and adverse pregnancy outcomes might be confounded by the use of prednisone or immunosuppressants. We analyzed our cohort to determine if this was so.

Of the 260 pregnant patients we analyzed in the cohort, 191 (73%) did not take prednisone during the pregnancy, whereas 17 (7%) took less than 5 mg/day, 29 (11%) took 5 to 9 mg/day, and 23 (9%) took 10 mg/day or more. Most of the patients did not take immunosuppressants during their pregnancy, whereas 70 (27%) took azathioprine, 14 (5%) took tacrolimus, and 1 each took mycophenolate, anakinra, and cyclosporin (we do not allow mycophenolate during pregnancy; exposure was accidental).

The mean 25(OH)D blood levels were not different in those not receiving immunosuppression (39.5 + 12.6 ng/mL), azathioprine (41.1 + 13.5 ng/mL) or tacrolimus (43.4 + 16.6 ng/mL). We took the multivariate model (original model) from our article[1] and compared the original model to the model adjusted for prednisone and immunosuppressive use (Table 1). We found that the association of 25(OH)D blood levels and adverse pregnancy outcomes remained essentially unchanged after adjustment. We thus conclude that the observed relationship between 25(OH)D blood levels and adverse pregnancy outcomes was not confounded by immunosuppressive drug or prednisone use.

Michelle Petri, MD, MPH
Andrea Fava, MD
Daniel W. Goldman, PhD
Nima Madanchi, MD
Johns Hopkins University School of Medicine
Baltimore, MD
Laurence S. Magder, PhD
University of Maryland
Baltimore, MD

1. Madanchi N, Fava A, Goldman DW, et al. Association between 25-hydroxyvitamin D levels and adverse pregnancy outcomes in systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2025;77(4):432–439.

DOI 10.1002/acr.25577

Pain sensitivity and chronic pain as a link between analgesic use and cardiovascular/gastrointestinal risk: comment on the article by Kaur et al

To the Editor:

I read with great interest the article by Kaur et al[1] in *Arthritis Care & Research*, which examined the association between acetaminophen use and the risk of cardiovascular and gastrointestinal complications in older adults. Although the authors conducted rigorous adjustments for osteoarthritis and other comorbidities, it remains difficult to fully account for individual differences in pain sensitivity or the presence of chronic pain, which may influence both analgesic use and disease risk.

Recent studies have demonstrated that individuals with high pain sensitivity or low pain tolerance are at increased risk of cardiovascular disease. For instance, the Tromsø Study showed that participants with low pain tolerance, measured by a cold pressor test, had significantly higher risks of coronary artery disease and all-cause mortality.[2] Furthermore, chronic pain itself has been identified as an independent risk factor for gastrointestinal

Table 1. Adjusted HRs for vitamin D based on three different multivariate models*

Comparison	Original model, HR (95% CI)	Model adjusted for prednisone, HR (95% CI)	Model adjusted for immunosuppressive drugs, HR (95% CI)
<20 vs 40–60 ng/mL	3.35 (1.64–6.80)	3.45 (1.67–7.10)	3.41 (1.67–6.98)
20–30 vs 40–60 ng/mL	1.53 (0.84–2.80)	1.50 (0.82–2.74)	1.56 (0.85–2.86)
30–40 vs 40–60 ng/mL	1.51 (0.91–2.51)	1.53 (0.92–2.54)	1.52 (0.92–2.51)
≥60 vs 40–60 ng/mL	1.34 (0.69–2.62)	1.32 (0.67–2.57)	1.32 (0.68–2.59)

* CI, confidence interval; HR, hazard ratio.

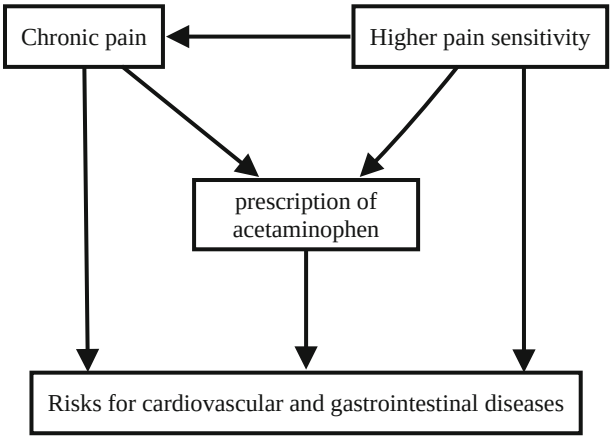

Figure 1. Hypothetical causal diagram of pain sensitivity, analgesic use, and disease risk.

diseases, particularly peptic ulcer disease (PUD). Wang et al reported that patients with fibromyalgia, a chronic pain condition, had a 40% higher risk of developing PUD compared to matched controls, even after adjusting for nonsteroidal anti-inflammatory drug use and other potential confounders.[3]

Taken together, these findings suggest that pain sensitivity and chronic pain burden may influence both the likelihood of analgesic use and the risk of cardiovascular and gastrointestinal diseases. I propose a hypothetical causal framework (Figure 1) in which acetaminophen use may not directly cause adverse outcomes but rather acts as a mediator reflecting the underlying high pain sensitivity and chronic pain burden, both of which are themselves associated with increased disease risk. Future studies should consider incorporating pain sensitivity or chronic pain burden into the causal framework to better elucidate the relationship between analgesic use and cardiovascular or gastrointestinal risk.

I would like to thank Associate Professor Noboru Hagino (Division of Rheumatology, Teikyo University Medical Center in Chiba) for sharing the original article on social media and for the insightful discussions, which provided us the opportunity to read and think deeply about this important study. This work was supported by part of the All-Osaka U Research in "The Nippon Foundation—The University of Osaka Project for Infectious Disease Prevention."

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25577.

Ryuichi Minoda Sada, MD
sadao@cider.osaka-u.ac.jp
Department of Transformative Protection to Infectious Disease, Graduate School of Medicine
The University of Osaka
Suita, Osaka, Japan
Department of Infection Control, Graduate School of Medicine
The University of Osaka
Suita, Osaka, Japan
Center for Infectious Disease Education and Research (CiDER)
The University of Osaka
Suita, Osaka, Japan

1. Kaur J, Nakafero G, Abhishek A, et al. Incidence of side effects associated with acetaminophen in people aged 65 years or more: a prospective cohort study using data from the clinical practice research datalink. Arthritis Care Res (Hoboken) 2025;77(5):666–675.
2. Fladseth K, Lindekleiv H, Nielsen C, et al. Low pain tolerance is associated with coronary angiography, coronary artery disease, and mortality: the Tromsø Study. J Am Heart Assoc 2021;10(22):e021291.
3. Wang KA, Wang JC, Lin CL, et al. Association between fibromyalgia syndrome and peptic ulcer disease development. PLoS One 2017; 12(4):e0175370.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 77, No. 12, December 2025, pp 1496–1498
DOI 10.1002/acr.70021

2025 AC&R Reviewers

I would like to thank the following individuals for their time and effort in reviewing articles for *Arthritis Care & Research* in the last year. The continued high quality of the journal depends on the dedicated service of these individuals.

Kelli D. Allen, PhD, Editor-in-Chief

Ackerman, Ilana
Ahedi, Harbeer
Ahmed, Salahuddin
Aitken, Dawn
Ajeganova, Sofia
Alarcón, Graciela
Albayda, Jemima
Allen, Kelli
Almeida, Gustavo
Alten, Rieke
Alvarez, Carolina
Amital, Howard
Anandarajah, Allen
Anastasiou, Christine
Andersson, Maria
Andonian, Brian
Andrade, Danieli
Andrews, James S.
Antwi, Henry Asante
Aoyagi, Kosaku
Appenzeller, Simone
Arbeeva, Liubov
Askanase, Anca
Assassi, Shervin
Bacon, Kathryn
Bade, Mike
Baek, Jonggyu
Bajaj, Puneet
Bajpai, Ram
Baker, Joshua F.
Baker, Matthew
Ballengee, Lindsay
Bankole, Adegbenga
Bannuru, Raveendhara
Bansback, Nick
Barbour, Kamil
Barnabe, Cheryl
Barnado, April
Baron, Murray
Bartels, Christie
Bartlett, Susan
Barton, Jennifer
Bass, Anne
Behrens, Edward
Beretta, Lorenzo
Bermas, Bonnie
Bertsias, Giorgos
Bindawas, Saad
Bingham III, Clifton
Blanco, Irene
Bloom, Jessica
Bogdanos, Dimitrios
Bolster, Marcy
Booysen, Nadine
Bove, Allyn
Boyer, Katherine
Brady, Teresa
Braverman, Genna
Broder, Anna
Brown, Matthew
Bruera, Sebastian
Bruni, Cosimo
Brunner, Hermine
Buehring, Bjoern
Bujor, Andreea
Bullock, Danielle
Bunker, Stephen
Button, Kate
Byram, Kevin
Cafaro, Giacomo
Cai, Guoqi
Callahan, Leigh
Callen, Jeffrey
Campar, Ana
Canizares, Mayilee
Cannella, Amy
Cannon, Grant W.
Cappelli, Laura
Carlesso, Lisa
Carmona, Loreto
Chang, Alison
Chang, Joyce
Chang, Joyce C.
Checko, Erica
Chen, Yen
Chighizola, Cecilia
Chimenti, Maria Sole
Christensen, Robin
Chrysidis, Stavros
Chung, Cecilia
Chung, Lorinda
Clauw, Daniel
Cleveland, Rebecca
Clowse, Megan
Coates, Laura
Coburn, Brian
Coffman, Cynthia
Colbert, Robert
Coleshill, Mathew
Collins, Jamie
Collins, Kelsey
Consolaro, Alessandro
Cook, Chad
Cooper, Jennifer
Cootes, Timothy
Corbitt, Kelly
Correll, Colleen
Costello, Anna
Courvoisier, Delphine
Crofford, Leslie
Cron, Randy
Crowson, Cynthia
Culvenor, Adam
Cutolo, Maurizio
Dampier, Carlton
Danve, Abhijeet
Daoud, Ansaam
Davis III, John
Day, Richard
de Boysson, Hubert
Del Galdo, Francesco
Denton, Christopher
DeQuattro, Kimberly
Desmarais, Julianna
DeVellis, Robert
Di Carlo, Marco
Diaz-Gonzalez, Federico
Ding, Huihua
Dixon, William
Dobson, Fiona
Dobson, Kathleen
Domsic, Robyn
Doria, Andrea
Drenkard, Cristina
Driban, Jeffrey
Dua, Anisha
Duarte-Garcia, Ali
Dubreuil, Maureen
Dufour, Alyssa
Duong, Vicky
Eckstein, Felix
Edens, Cuoghi
Eder, Lihi
Edwards, Christopher
El-Zaatari, Helal M.
England, Bryant
Ermann, Joerg
Eudy, Amanda
Eyles, Jillian
Falasinnu, Titilola
Fanouriakis, Antonis
Feghali-Bostwick, Carol
Feldman, Brian
Feldman, Madelaine
Fernandez, David
Ferreira, Gilda
Ferucci, Elizabeth
Figueroa-Parra, Gabriel
Fitzcharles, Mary-Ann
FitzGerald, John

Fixen, Danielle
Ford, Julia
Frey Law, Laura
Fu, Kai
Furst, Daniel
Gabriel, Kelley
Gaffo, Angelo
Gagne, Samuel
Gao, Xiaoming
Garg, Shivani
Gasteiger, Chiara
George, Michael
Georgopoulos, Vasileios
Giles, Jon
Ginzler, Ellen
Gladman, Dafna
Glintborg, Bente
Gmuca, Sabrina
Godziuk, K.
Gold, Laura
Goldbach-Mansky, Raphaela
Golightly, Yvonne
Gomez, Alvaro
Good, Samuel
Gordon, Jessica
Gossec, Laure
Grainger, Rebecca
Green, Karlyn
Gupta, Latika
Haber, Travis
Hadlandsmyth, Katherine
Hajra, Adrija
Hakim, Alan
Hall, Michelle
Hamijoyo, Laniyati
Hannan, Marian
Haro, Sara
Harris, Julia
Harrison, Mark
Hawker, Gillian
He, Jade
Heerey, Joshua
Helfgott, Simon
Helliwell, Toby
Henderson, Lauren
Hicks, Gregory
Hinchcliff, Monique
Hinman, Rana
Hinze, Alicia
Holladay, Emily
Hollick, Rosemary
Holmqvist, Marie
Horn, Kailyn
Horton, Daniel
Howren, Alyssa
Hsiao, Betty
Hudson, Marie
Hummers, Laura
Hwang, Mark
Isenberg, David
Izadi, Zara
Izmirly, Peter
Jackson, Lesley
Jacobs, Deborah
Jafarzadeh, S. Reza
Janiak, Elizabeth
Jansen, Mylène
Jaros, Brian
Jarraya, Mohammed
Jin, Yuxuan
Joerns, Elena K.
Johnson, Sindhu
Johnson, Tate
Johr, Chadwick
Jonas, Beth
Jones, Bethan
Joos, Rik
Jorge, April
Jöud, Anna
June, Rayford
Kalunian, Kenneth
Kandane-Rathnayake, Rangi
Kang, Eun Ha
Karacabeyli, Derin
Kashikar-Zuck, Susmita
Katz, Guy
Katz, Patricia
Katz, Steven
Kayser, Cristiane
Kerr, Gail
Khanna, Puja
Khoja, Samannaaz
Khosroshahi, Arezou
Kiltz, Uta
King, Lauren
Klareskog, Lars
Ko, Kichul
Kojima, Toshihisa
Kortekaas, Marion
Kosik, Kyle
Krustev, Eugene
Kuchinad, Kamini
Kumar, Deepak
Kumthekar, Anand
Kuriya, Bindee
Kuwana, Masataka
Landewé, Robert
Lane, Chris
LaValley, Michael
Lazaro, Deana
Lee, Jiha
Lee, Yvonne
Legge, Alexandra
Lentz, Trevor
Leong, Khai Pang
Lerman, Melissa
Lewis, Myles
Lewis, Tene
Li, Guoxing
Li, Linda
Li, Yongzhe
Lieber, Sarah
Liew, Jean
Liles, Sydney
Littlejohn, Geoffrey
Liu, Shao-Hsien
Lo, Grace
Lockwood, Megan
Lonowski, Sarah
Lovalekar, Mita
Lu, Bing
Luedders, Brent
Lund, Merete Hetland
Lundberg, Ingrid E.
MacDermid, Joy
Machado, Pedro
MacKay, Crystal
Madsen, Norman
Mahajan, Tina
Mahmoudian, Armaghan
Makris, Una
Mammen, Andrew
Mannion, Melissa
Marriott, Kendal
Marston, Bethany
Martinez-Taboada, Victor
Massardo, Loreto
Master, Hiral
Mattumpuram, Jishanth
Matza, Mark
Mauck, Matthew
Mayes, Maureen
McCormick, Natalie
McCoy, Sara
McDermott, Catherine
McGeoch, Paul
McKenna, Sean
McMahan, Zsuzsanna
McWilliams, Daniel
Meade, Tanya
Mease, Philip
Mecoli, Christopher
Medlin, Jennifer
Mehta, Bella
Menon, Aravind
Menz, Hylton
Merriman, Tony
Michaud, Kaleb
Mielenz, Thelma
Mikuls, Ted
Milatz, Florian
Miller, Frederick W.
Miller, Heather
Miller, John
Miller, Jordan
Mills, Kathryn
Minhas, Deeba
Misra, Devyani
Miyawaki, Yoshia
Mohan, Vishnu
Mokbel, Abir
Molto, Anna
Mongin, Denis
Moore, Duncan
Morgan, Esi
Mosca, Marta
Mukherjee, Monica
Mulrooney, Elisabeth
Myasoedova, Elena
Nagaraja, Vivek
Nagy, György
Nash, Peter
Neelapala, Raghava
Neelapala, YV
Nelson, Amanda
Niewold, Timothy
Ogbu, Ekemini
Ogdie, Alexis
Onel, Karen
Orbai, Ana-Maria
Ozen, Gulsen
Paley, Michael
Park, Elizabeth

Parsons, Camille
Pascart, Tristan
Patel, Naomi
Paudel, Misti
Pedro, Sofia
Pellegrini, Christine
Peoples, Christine
Peredo-Wende, Ruben
Perruccio, Anthony
Perry, Thomas
Peschken, Christine
Pierce, Jennifer
Pillinger, Michael
Pincus, Theodore
Poddubnyy, Denis
Pons-Estel, Guillermo
Pope, Janet
Proft, Fabian
Quinn, Kaitlin
Rai, Sharan
Rajan, Rajesh
Ramirez, Michelle
Ramos, Paula
Ramsey-Goldman, Rosalind
Raza, Shaina
Razmjou, Amir
Remalante-Rayco, Patricia
Reveille, John
Reynolds, Richard J.
Rezaei, Elham
Richette, Pascal
Riley, Thomas
Roberts, Eric
Robinson, Katie
Robson, Joanna
Rodeo, Scott
Rodriguez-Laso, Angel
Romich, Ellen
Roul, Punyasha
Ruderman, Eric
Rudwaleit, Martin
Sada, Ken-ei
Salsbury, Stacie
Sammaritano, Lisa
Santacroce, Leah
Santiago, Tânja
Sato, Emilia
Saygin, Didem
Schäfer, Valentin
Schmajuk, Gabriela
Schmidt, Nancyanne
Schmukler, Juan
Schrepf, Andrew
Scofield, Robert Hal
Scott, Ian
Sen, Rouhin
Shah, Ankoor
Shapiro, Abigail
Siegel, Caroline
Siegel, Patricia
Simard, Julia
Sims, Catherine
Singh, Namrata
Singh, Ram
Sivaraman, Vidya
Smith, Andy
Smith, Judith
Smith, Vanessa
Snow, Marcus
Soulsby, William
Spaetgens, Bart
Sparks, Jeffrey
Spitznagle, Jacob
Spoorenberg, Anneke
Stamp, Lisa
Steen, Virginia
Stefanik, Joshua
Stern, Brocha
Stoll, Matthew
Stovall, Rachael
Sulak, Dustin
Tamang, Suzanne
Tan, Bryan
Tang, Bowen
Tatangelo, Mark
Taxter, Alysha
Taylor, Steven
Tedeschi, Sara
Teng, Yoe Kie Onno
Terhorst, Lauren
Terkeltaub, Robert
Thabah, Molly Mary
Thoma, Louise
Thomas, Donald
Thorlund, Jonas
Toprak, Murat
Toprover, Michael
Tore, Gizem
Touma, Zahi
Treister-Goltzman, Yulia
Twilt, Marinka
Ugarte-Gil, Manuel
Valenzuela, Antonia
van der Heijde, Désirée
Van der Horst-Bruinsma, Irene
Van Herwaarden, Noortje
Vasquez-Canizares, Natalia
Venerito, Vincenzo
Venhoff, Nils
Verstegen, Ruud
Vina, Ernest
Vlad, Violeta
Vogel, Tiphanie
Volkmann, Elizabeth
Voskuyl, Alexandre
Wallace, Beth
Walsh, Jessica
Wang, Christine
Warrington, Kenneth J.
Wei, Jie
Wen, Chunyi
Werner, David
Werth, Victoria
Wheeler, Austin
Whibley, Daniel
White, Daniel
Widdifield, Jessica
Wilson, Timothy M.
Wright, Daniel FB
Wu, Jiayue
Xu, Shiwen
Yajima, Nobuyuki
Yazdany, Jinoos
Yazdanyar, Ali
Yelin, Edward
Yoshimi, Ryusuke
Yue, Wei-hua
Zamudio-Cuevas, Yessica
Zen, Margherita
Zeng, Chao
Zhao, Hui
Zhu, Guohao
Zickuhr, Lisa
Zimba, Olena

VOLUME INDEX

VOLUME 77 • January 2025 • NO. 1

Special Theme Issue: Environmental and Geographic Factors and Rheumatic Disease

VOLUME 77 • February 2025 • NO. 2

VOLUME 77 • March 2025 • NO. 3

VOLUME 77 • April 2025 • NO. 4

VOLUME 77 • May 2025 • NO. 5

VOLUME 77 • June 2025 • NO. 6

VOLUME 77 • July 2025 • NO. 7

VOLUME 77 • August 2025 • NO. 8

VOLUME 77 • September 2025 • NO. 9

VOLUME 77 • October 2025 • NO. 10

VOLUME 77 • November 2025 • NO. 11

VOLUME 77 • December 2025 • NO. 12