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. 3 March 2025

Arthritis Care & Research

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

Arthritis Care & Research

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

VOLUME 77 • MARCH 2025 • NO. 3

Special Article

Review Article: Rheum for Improvement? Delayed Diagnosis of Juvenile Idiopathic Arthritis: A Narrative Review
Anna Costello, Irit Rasooly, and Pamela Weiss 283

Clinicopathologic Conference: The Matter at Hand: A Case of Difficult-to-Treat Arthritis
Ana L. Altaffer, Lindsay C. Burrage, Ankur Kamdar, Tiphanie P. Vogel, and Maria Pereira 291

Editorial: Is Collaborative Care Better Care?
James T. Rosenbaum, Nicole Fett, Daniela Ghetie, and Julianna Desmarais. 297

Pediatrics

Assessing Methotrexate Adherence in Juvenile Idiopathic Arthritis Using Electronic Health Record–Linked Pharmacy Dispensing Data
Dori Abel, David Anderson, Michael J. Kallan, Levon Utidjian, Jon M. Burnham, Joyce C. Chang, Chén C. Kenyon, and Sabrina Gmuca 300

Exploring Pain Adaptation in Youth With Juvenile Idiopathic Arthritis: Identifying Youth and Parent Resilience Resources and Mechanisms
Yvonne N. Brandelli, Sean P. Mackinnon, Christine T. Chambers, Jennifer A. Parker, Adam M. Huber, Jennifer N. Stinson, Shannon A. Johnson, and Jennifer P. Wilson 309

Rheumatology Education & Practice

Measurable Outcomes of an Ophthalmology and Rheumatology Coordinated Care Clinic
Catherine Lavallee, Monica Ahrens, Stefanie L. Davidson, Haseeb Goheer, Allison Shuster, and Melissa A. Lerman 322

Brief Report: Bringing Reproductive Health Guidelines Into Fellowship Training: A National Survey of Adult and Pediatric Rheumatology Fellows and Program Directors
Selene Rubino, Michael J. Battistone, Stacy P. Ardoin, Elise D. Berlan, Kristine Carandang, Kate Chiseri, Arthur Kavanaugh, Whitney White, Kelly A. Wise, Andrew L. Wong, Bethany Marston, and Megan E. B. Clowse 330

Rheumatoid Arthritis

Multimorbidity Patterns and Rheumatoid Arthritis Disease Outcomes: Findings From a Multicenter, Prospective Cohort
Sarah Dutt, Punyasha Roul, Yangyuna Yang, Tate M. Johnson, Kaleb Michaud, Brian Sauer, Grant W. Cannon, Joshua F. Baker, Jeffrey R. Curtis, Ted R. Mikuls, and Bryant R. England 337

Preferences for Tapering Biologic Disease-Modifying Antirheumatic Drugs Among People With Rheumatoid Arthritis: A Discrete Choice Experiment
Suz Jack Chan, Lisa K. Stamp, Gareth J. Treharne, Janet M. Y. Cheung, Nicola Dalbeth, Rebecca Grainger, Simon Stebbings, and Carlo A. Marra 349

Osteoarthritis

Radiographic Changes Five Years After Treatment of Meniscal Tear and Osteoarthritic Changes
Jeffrey N. Katz, Jamie E. Collins, Robert H. Brophy, Brian J. Cole, Charles L. Cox, Ali Guermazi, Morgan H. Jones, Bruce A. Levy, Lindsey A. MacFarlane, Lisa A. Mandl, Robert G. Marx, Faith Selzer, Kurt P. Spindler, Rick W. Wright, Elena Losina, and Yuchiao Chang 359

Association of Foot Symptoms With Decreased Time to All-Cause Mortality: The Johnston County Osteoarthritis Project
Skylar Harmon, Carolina Alvarez, Marian T. Hannan, Leigh F. Callahan, Lucy S. Gates, Catherine J. Bowen, Hylton B. Menz, Amanda E. Nelson, and Yvonne M. Golightly 366

Spondyloarthritis

Low Socioeconomic Status and Female Sex are Associated With Worse Functional Status in Axial Spondyloarthritis
Rachael Stovall, Jing Li, Jessica Fitzpatrick, Eric Roberts, Andriko Palmowski, Christine Anastasiou, Zara Izadi, Janna Friedly, Namrata Singh, Lianne S. Gensler, Gabriela Schmajuk, and Jinoos Yazdany 376

Global Functioning in Axial Spondyloarthritis is Stronger Associated With Disease Activity and Function Than With Mobility and Radiographic Damage
David Kiefer, Jürgen Braun, Uta Kiltz, Varvara Chatzistefanidi, Daniela Adolf, Ilka Schwarze, Maria Kabelitz, Uwe Lange, Jan Brandt-Jürgens, Edgar Stemmler, Sabine Sartingen, and Xenofon Baraliakos 385

Systemic Sclerosis

Differential Item Functioning on the Cochin Hand Function Scale Among People With Systemic Sclerosis by Language, Sex, and Disease Subtype: A Scleroderma Patient-Centered Intervention Network (SPIN) Cohort Study
Mingyao Xu, Daphna Harel, Marie-Eve Carrier, Linda Kwakkenbos, Susan J. Bartlett, Karen Gottesman, Geneviève Guillot, Laura Hummers, Vanessa L. Malcarne, Michelle Richard, and Brett D. Thombs, on behalf of the SPIN Investigators 393

Other and Mixed Rheumatic Diseases

Patterns of Imaging Requests By General Practitioners for People With Musculoskeletal Complaints: An Analysis From a Primary Care Database
Romi Haas, Alexandra Gorelik, Denise A. O'Connor, Christopher Pearce, Danielle Mazza, and Rachelle Buchbinder 402

Brief Report: Urban and Rural Patterns of Health Care Utilization Among People With Rheumatoid Arthritis and Osteoarthritis in a Large US Patient Registry
Luke W. Desilet, Sofia Pedro, Patricia Katz, and Kaleb Michaud 412

Erratum

Correction to Benefits of Early Versus Late Initiation of Intravenous Immunoglobulin in the Treatment of Patients With Anti–3-Hydroxy-3-Methylglutaryl-Coenzyme A Reductase Immune-Mediated Necrotizing Myopathy. 419

Cover image: The image on the cover (from Altaffer et al; pages 291−296) shows magnetic resonance imaging of the right hand (coronal T1 fat saturated post-gadolinium) demonstrating small size and abnormal configuration of the carpal bones.

Arthritis Care & Research
Vol. 77, No. 3, March 2025, pp 283–290
DOI 10.1002/acr.25438

American College of Rheumatology
Empowering Rheumatology Professionals

REVIEW ARTICLE

Rheum for Improvement? Delayed Diagnosis of Juvenile Idiopathic Arthritis: A Narrative Review

Anna Costello,[1] Irit Rasooly,[2] and Pamela Weiss[2]

Juvenile idiopathic arthritis (JIA) is the most common rheumatic disease of childhood and a disease for which we have safe and effective therapies. Early diagnosis of JIA enables timely initiation of therapy and improves long-term disease outcomes. However, many patients with JIA experience prolonged diagnostic delays and have a turbulent course to diagnosis. In this narrative review, we explore the importance of early diagnosis in JIA, what is known about time to diagnosis and diagnostic trajectory, and factors that contribute to delayed diagnosis. We also discuss next steps to improve time to diagnosis for these vulnerable patients.

Introduction

A 12-year male presents to rheumatology clinic for progressively worsening joint swelling and morning stiffness. Three months prior to presentation he developed right ankle swelling, initially evaluated by an orthopedist. Ankle x-rays were normal. Given concern for an occult fracture, he was immobilized with a walking boot. He then developed bilateral knee pain and swelling and returned to care with the orthopedist. His symptoms were attributed to gait abnormalities while in a walking boot, and the boot was removed. However, his ankle remained swollen and the stiffness and pain in his knees continued to worsen. His pediatrician ordered laboratory tests, which were notable for elevated markers of inflammation and normocytic anemia, and magnetic resonance imaging (MRI) of the knee, which showed a large effusion. His pediatrician then referred to a pediatric rheumatologist for further evaluation.

On review of systems, parents share that 7 months ago their child was diagnosed with anterior uveitis of the left eye and was treated with prednisolone forte drops. He had a good response to therapy, and basic laboratory work up did not identify a cause of his uveitis. His uveitis recurred around the time of his right ankle swelling and was again controlled with prednisolone forte drops. Since initial referral, he also developed elbow and finger pain and stiffness. At his initial pediatric rheumatology visit, he is found to have arthritis in five joints and is diagnosed with juvenile idiopathic arthritis (JIA) about 7 months after his initial episode of uveitis. The patient was started on a short course of oral steroids given the severity of his symptoms in addition to adalimumab and methotrexate. He has had excellent response to therapy.

This patient's diagnostic journey is common among children with JIA. Although best practice guidelines suggest diagnosis within 10 weeks of symptom onset,[1] pediatric rheumatologists encounter patients who have been symptomatic for months to years before diagnosis. Many, like the patient described above, are evaluated for their symptoms by multiple providers in primary care, emergency medicine, infectious disease, and orthopedics before being referred to rheumatology. They undergo significant diagnostic workups, sometimes including sedated imaging or invasive procedures. Some receive incorrect diagnoses of traumatic injuries and are immobilized and restricted from sports and other activities. Many experience prolonged periods of pain and limited physical function. Patients often present to their initial rheumatology visit with physical examination findings suggestive of longstanding disease, including irreversible ocular or joint damage. In this narrative review, we will explore what is known about diagnostic delays in JIA, including the clinical importance of early diagnosis and pertinent best practice guidelines. We will synthesize what is known about time to diagnosis, before diagnosis health care utilization, and the

Dr Costello's work was supported by grants from the Childhood Arthritis and Rheumatology Research Alliance/the Arthritis Foundation and from AcademyHealth/the Betty and Gordon Moore Foundation. Dr Rasooly's work was supported by NIH/AHRQ (grant K08-HS028682). Dr Weiss's work was supported by PCORI and the NIH/National Institute of Arthritis and Musculoskeletal and Skin Diseases (grants K24-AR078950).

[1]Anna Costello, MD: Children's Hospital of Philadelphia, Philadelphia, Pennsylvania; [2]Irit Rasooly, MD, MSCE, Pamela Weiss, MD, MSCE: Children's Hospital of Philadelphia and University of Pennsylvania, Philadelphia.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25438.

Address correspondence via email to Anna Costello, MD, at Costelloa@chop.edu.

Submitted for publication September 10, 2024; accepted in revised form September 12, 2024.

family experience of the diagnostic process. Finally, we will hypothesize about why JIA is such a diagnostically complex condition and what the next steps may be to improve time to diagnosis for these vulnerable patients.

Why is early diagnosis of JIA important?

JIA is the most common rheumatic disease of childhood, affecting approximately 1 in 1,000 children.[2] Without early and aggressive treatment of JIA, patients can suffer numerous complications including destructive joint damage, blindness from uveitis, and long-lasting pain and disability.[3,4] New disease-modifying therapies have altered the course of and outcomes for JIA, and disease guidelines now suggest that we should strive for inactive disease in every patient.[5] Early, aggressive treatment of JIA results in better short-term and long-term outcomes.[3,4,6] For instance, a multicenter observational cohort study that followed patients into adulthood associated initiation of a biologic within 2 years of symptom onset with a higher likelihood of drug-free remission and improved function in adulthood.[3] Shorter disease duration at diagnosis has also been associated with a higher likelihood of and a longer duration of clinically inactive disease.[6] A single-center cross-sectional study of children with enthesitis related arthritis (ERA) found a 20% higher risk of long-term poor functional outcomes for each year of diagnostic delay.[7] These findings are in keeping with the "window of opportunity" theory, that early in its course. arthritis is more susceptible to treatments, which is well supported in the adult rheumatology literature for rheumatoid arthritis[8] and psoriatic arthritis (PsA).[9] Ensuring early diagnosis for our patients, and thereby enabling early initiation of therapy, is an important step in improving JIA outcomes.

More broadly, diagnostic excellence, establishing a timely, accurate explanation of a patient's health problem, is increasingly recognized as a fundamental aspect of delivering safe, high quality health care,[10–12] and diagnostic errors are a leading source of patient harm.[13–15] Characterizing trajectories of JIA diagnosis and the opportunities to improve JIA diagnosis are priorities in pediatric rheumatology and will allow for targeted interventions to improve diagnosis of our most common condition.

What is the target timeframe for diagnosis?

Because early diagnosis and expedient initiation of therapy are clearly important, pediatric rheumatology national organizations have adopted best practices guidelines for the timeline of rheumatology evaluation, JIA diagnosis, and treatment initiation. For example, the British Society for Pediatric and Adolescent Rheumatology published guidelines setting a benchmark that patients with JIA should be seen by a rheumatologist within 10 weeks of symptom onset and within 4 weeks of referral.[1] A proposed set of quality measures for the care of JIA in the United States guided by a survey to American health care professionals set a goal for a first visit to rheumatology within 60 days of referral for patients with symptoms concerning for JIA,[16] whereas guidelines in Australia recommend referral to pediatric rheumatology if children have symptoms consistent with JIA that persist for more than 4 weeks.[17]

What is known about time to diagnosis in JIA?

For the purposes of this review, we evaluated 23 primary articles and 1 abstract from 2007 to 2024 that discuss time to JIA diagnosis and health care utilization proceeding diagnosis (Table 1). These articles contain data from 11 countries and 1 multinational study. All studies used clinical data from single-center or multicenter cohorts, which were collected via chart review, national pediatric rheumatology registries, or surveys of families, and there were no studies that used administrative claims data. Our literature search was conducted in PubMed and Google Scholar using the following keywords "Juvenile Idiopathic Arthritis," "diagnostic delay," "delayed diagnosis," "time to diagnosis," "time to referral," "referral pathways," and "health care utilization." Reference lists of the included articles were also manually reviewed to identify further studies.

Overall, the existing literature on time to diagnosis of JIA suggests that patients experience notable delays in diagnosis and often require multiple visits to health care professionals before they are appropriately referred to a pediatric rheumatologist. A 2020 systematic review article by Chausset et al reviewed 15 articles published between 1994 and 2020 from six countries and one multinational study. The pooled time to access pediatric rheumatology care internationally was 23 weeks,[18] which is longer than the 10 weeks recommended by the British Society of Paediatric and Adolescent Rheumatology guidelines.[1] The distribution for time to diagnosis was wide (4–656 weeks) and suggests that a large fraction of patients is symptomatic for many months or even several years before diagnosis. For instance, 16% of patients in a Canadian cohort[19] and 16% of patients in a US cohort[20] had symptoms for more than a year before their first rheumatology assessment. In a retrospective study using data from the Childhood Arthritis Prospective Study cohort, a national JIA registry in the UK, Abid et al found that 21% of children had symptoms for more than a year before diagnosis. They also demonstrated an association between longer time to rheumatology evaluation and higher active joint count at time of diagnosis.[21]

When considering the overall time to diagnosis, there are multiple time periods or intervals that can contribute to delayed diagnosis, which are demonstrated in Figure 1.[22] Understanding which of these intervals is the largest driver of diagnostic delay will help to determine the types of interventions that might improve time to diagnosis. Applying this framework, three European studies (combined n = 202) evaluated the "careseeking interval" and found the median time between symptom onset and first presentation to any health care setting ranged from 10 to 14 days.[23–25]

Table 1. Diagnostic time intervals by study*

Reference	Country	n	Type of study	Careseeking interval, days[a]	Referral interval, weeks[a]	Scheduling interval weeks[a]	Time to PRh, weeks[a]	Time to diagnosis, weeks[a]
Schiff et al 2009[23]	Canada	35	Single-center questionnaire	14 (3–560)	8 (0–708)	5 (0–24)	24 (4–713)[b]	—
Schiff et al 2010[19]	Canada	319	National cohort (ReACCh Out)	—	—	—	16 [6–31]	—
Barber et al 2020[26]	Canada	164	Single-center cohort	—	—	3 [1–6]	—	—
Boiu et al 2012[47]	France	95	Single-center cohort	—	—	—	—	16 (0–27)
Aoust et al 2017[31]	France	67	Single-center cohort	—	—	—	—	12 (1–324)
Freychet et al 2019[32]	France	111	Two center cohort	—	—	—	13.2 [5.2–42.8]	—
Chausset et al 2023[36]	France and Switzerland	250	Two center cohort	0 [0–2.8]			9.6 [5.2–27.6]	
Chausset et al 2024[30]	France	19	Qualitative interviews	—	—	—	—	—
Tzaribachev et al 2009[24]	Germany	132	Single-center cohort	10 (0–1,610)	—	—	13 (0–308)	—
Agarwal et al 2020[33]	India	520	Single-center cohort	—	—	—	16 [7–62][c]	13 [5–40][c]
Frenkel et al 2023[34]	Israel	201	Multicenter cohort	—	—	—	—	8 (0–350)
Marino et al 2024[25]	Italy	40	Single-center cohort	14.0 [3–43.25]	8 [2–22]	—	—	16 [8–36]
Khawaja et al 2017[35]	UAE	66	Two center cohort	—	—	—	40 (4–192)	—
Foster et al 2007[27]	UK	152	Single-center cohort	—	—	1 (0–4)	20 (0–416)	—
Adib et al 2008[21]	UK	507	Multicenter cohort, CAPS	—	13 [6–32][d]	3 [1–7]	18 [9–38]	—
Hyrich et al 2009[48]	UK	740	Multicenter cohort, CAPS	—	—	—	21 [10–42]	—
Kavirayani et al 2013[49]	UK	428	Multicenter cohort	—	—	—	—	—
Verstappen et al 2015[39]	UK	923	Multicenter cohort, CAPS	—	—	—	22 [11–48]	—
McErlane et al 2016[28]	UK	1,066	Multicenter cohort, CAPS	—	—	4 [1.3–8]	24 [12.3–50.4]	—
Shoop-Worrall et al 2022[37]	UK	1,577	Multicenter cohort, CAPS	—	—	—	—	<5: 14 [7–25], 5–11: 20 [9–36], >11: 26 [13–54][e]
Rapley et al 2021[29]	UK	36	Qualitative interviews	—	—	—	22 (4–364)	—
Balmuri et al 2021[38]	US	1,684	Multicenter cohort, CARRA	—	—	—	12 [4–24]	—
Ong et al 2020[20]	US	1,195	Multicenter cohort, CARRA	—	—	—	—	12
Consolaro et al 2019[50]	49 nations	9,081	Multicenter, multinational cohort	—	—	—	NE 14 [5–38], WE 21 [10–52], SE 15 [5–47], EE 15 [5–52], NA 15 [5–42], LA 21 [10–52], A&ME 21 [10–78], SA 31 [10–104][f]	—

* A&ME, Africa and Middle East; CAPS, Childhood Arthritis Prospective Study; CARRA, Childhood Arthritis and Rheumatology Research Alliance; EE, Eastern Europe; IQR, interquartile range; JIA, juvenile idiopathic arthritis; LA, Latin America; NA, North America; NE, Northern Europe; PRh, Pediatric Rheumatologist; ReACCH Out, Research in Canadian Children Emphasizing Outcomes; SA, Southeast Asia; SE, Southern Europe; UAE, United Arab Emirates; UK, United Kingdom; US, United States; WE, Western Europe.
[a] Median (range) [IQR].
[b] This study was not limited to patients with JIA but instead included pooled data of all PRh new patient visits.
[c] Time to diagnosis is shorter than time to PRh in approximately 25% of patients because some patients seen, diagnosed, and treated by adult rheumatologist.
[d] Time between symptoms onset and referral.
[e] Median times by age groups of <5, 5–11, and >11years.
[f] Median times by region: NE, WE, SE, EE, NA, LA, A&ME, and SA.

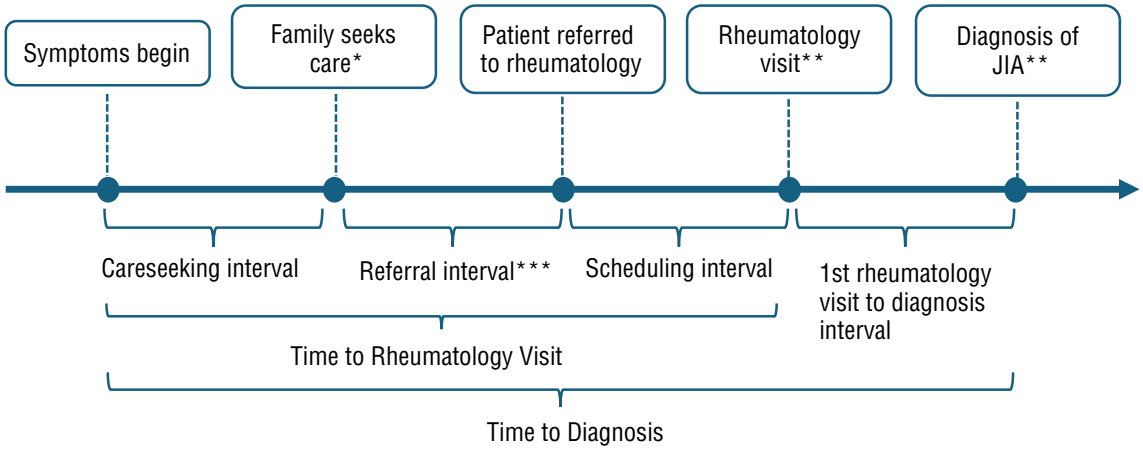


Figure 1. Time intervals in the diagnostic process for patients with JIA. *Family may seek care from multiple providers or from the same provider(s) multiple times before referral. **Patient may receive a diagnosis of JIA at their initial rheumatology visit. ***Some studies may define the referral interval as the time from symptom onset to referral, whereas others define it as the time from the first health care visit until referral is made. JIA, juvenile idiopathic arthritis.

The "referral interval," time from first visit to a health care provider until referral is made to pediatric rheumatology, was examined in three studies, which identified a median range to referral of 9 to 13 weeks with a large interquartile range of 2 to 32.[21,23,25] The "scheduling interval," time between referral and first rheumatology appointment, was evaluated in five studies and varied by country with the median time from referral to first available visit typically between 1 and 4 weeks.[21,23,26–28] In summary, the referral interval, the time in which patients are seeing providers for their symptoms but have not yet been referred to a rheumatologist, seems to be the largest contributor to prolonged time to diagnosis.

How do patients and families experience the journey to diagnosis?

These data align with the qualitative literature characterizing the patient and family experience of JIA diagnosis. In two qualitative studies, researchers performed semistructured interviews with the families of children with JIA. Rapley et al interviewed 51 family members of 31 children with JIA in the United Kingdom,[29] and Chausset et al interviewed 19 families of children with JIA in France.[30] Families characterized their diagnostic journey as "turbulent," with multiple visits (range 2–17), referrals, and incorrect diagnoses before appropriate referral to pediatric rheumatology was made.[29] Parents detailed the ways in which symptoms were initially attributed to activity-related injuries (eg, sports), normal development (toddler stumbling and falls), environmental factors (such as shoes not fitting), or behavior (such as attempting to avoid school or another activity).[29] Persistence or worsening of these symptoms was critical in triggering the family to present for care. When families sought care, they reported that they were perceived as overly anxious and worried by clinicians and that their own parental persistence in seeking an explanation of symptoms was necessary for diagnosis.[29,30] Families often turned to friends and the internet for more guidance about next steps because they were frustrated by the provider's uncertainty and often had to pursue or advocate for a second opinion to find a diagnosis.[30] Some found that that new symptoms were reflexively attributed to coexisting conditions, such as Trisomy 21 and psoriasis, which is especially alarming as these conditions are associated with a higher likelihood of arthritis.[29] Families expressed frustration with both the prolonged period of diagnostic uncertainty and at having concerns dismissed by providers.[29,30] Adolescents shared that the process made them sad and angry, and some found it affected their ability to trust doctors. Families shared that they were looking for empathy and open communication with their providers throughout the process. They appreciated when doctors were open about their uncertainty but had a positive outlook about working toward a diagnosis.[30]

What is the path to diagnosis for patients with JIA?

The family experience of a "turbulent" course to diagnosis is consistent with what has been demonstrated by a series of chart reviews and JIA registry studies about the diagnostic course for these patients.[23,25,27,31–34] They are often seen by multiple provider types and have many interactions with the health care system before diagnosis. Unsurprisingly, referral patterns are highly variable between countries given the inherent difference in health care systems. However, the major themes regarding health care utilization are summarized here to create a framework for understanding the often prolonged diagnostic journeys of our patients.

Patients saw a median of three health care providers (range 1–11) before the pediatric rheumatologist.[23,27,31–33] They sometimes saw one or more of these providers multiple times, with one study that used surveys and chart review to understand

diagnostic course finding that patients were seen for an average of five visits (range 1–59) before referral.[23] General practitioners, pediatricians, orthopedic surgeons, and emergency room providers were most likely to see patients with JIA prior to their diagnosis; less frequently, referrals were made by other subspecialists (ophthalmologists, plastic surgeons, neurologists, and infectious disease specialists) or physical therapists.[28] Adult rheumatologists played a large role in the diagnostic process in some countries, such as United Arab Emirates[35] and India,[33] and sometimes initiated therapy for these patients before referral to a pediatric rheumatologist. In Israel, hospitalization is quite common with about half of patients having a hospital stay before diagnosis.[34]

Orthopedists account for a large percentage of the referrals to pediatric rheumatology for JIA in many countries, (eg, 18.9% in Israel,[34] 24% in the United Kingdom,[27] and 34% in France).[32] It is common for patients to be referred by general practitioners, emergency room physicians, or pediatricians to orthopedists who then referred to rheumatology.[32,34] Several studies found that patients seen by orthopedics during the diagnostic process have more prolonged times to diagnosis.[24,27,32,36] However, it remains unclear if this is because they fail to recognize the condition or whether it is simply because any additional referral contributes to delay in referral to rheumatology.

Only three studies assessed whether patients were seen by ophthalmology for uveitis screening before their first visit with pediatric rheumatology,[23,25,27] and there is high variance in the frequence of screening with a range of 0% to 22%, likely signaling highly variable practice patterns between nations and health care systems.

Proceeding evaluation and treatment. Most children have some variety of imaging such as x-ray, ultrasound, or MRI before being referred to rheumatology.[25,27,31] Between 7% and 23.9% have invasive diagnostic procedures such as joint fluid aspiration (most common), synovial biopsy, or synovectomy before rheumatology evaluation.[32,34] More than 10% of children spent time immobilized in a cast or splint, limiting their participation in activities.[32]

Proceeding diagnoses. The formal diagnoses and treatments received prior to JIA diagnosis are not well characterized. In a retrospective chart review of 76 patients, Aoust et al reported reactive arthritis and septic arthritis were the most common proceeding diagnoses,[31] but the other reviewed studies did not comment on proceeding diagnoses.

Which clinical characteristics are associated with prolonged time to diagnosis?

Several of the reviewed studies used univariate or multivariate modeling to try to understand whether demographic or clinical factors were associated with more prolonged time to diagnosis. Universally, patients with systemic JIA had a lower median time to diagnosis seemingly secondary to the more overt symptoms at presentation, such as persistent fevers and elevated markers of inflammation.[21,28,32–34] Conversely, multiple studies found a more prolonged time to diagnosis for patients with ERA.[19,25,31,32,34] In France, for example, the median time to diagnosis for the full JIA cohort was 13.2 weeks (interquartile range [IQR] 5.2–42.8), meanwhile for patients with ERA the median was 44.8 weeks (IQR 16–96). Older age at diagnosis[25,32,37] and male sex[25,34] were also found to be associated with more prolonged diagnostic delays. However, as ERA is much more common in males and older patients and, because ERA is the subtype most commonly associated with prolonged time to diagnosis, it is possible that subtype was a confounding factor.

Interestingly, patients with an elevated erythrocyte sedimentation rate were found to have a shorter time to diagnosis,[21,25,32,34] likely because providers use elevated markers of inflammation to screen for autoimmune conditions and are, therefore, more likely to refer these children to rheumatology. Similarly, a positive antinuclear antibody (ANA) was found to be associated with a shorter time to diagnosis,[32] as was a family history of autoimmunity.[33] We hypothesize that referring providers inappropriately associate a positive ANA with a high likelihood of arthritis, and so this lowers the threshold for referral for some, whereas prolonging time to referral for ANA negative patients with JIA. A family history of autoimmunity is likely another red flag that triggers providers to refer and may also encourage parents to request referral or other workup.

Several of the reviewed studies evaluated the degree to which sociodemographic factors were related to time to diagnosis. Shiff et al found that higher levels of parental education were associated with more rapid time to diagnosis,[19] and Balamuri et al used data from the Childhood Arthritis and Rheumatology Research Alliance, an American JIA registry, to demonstrate a weak association between community poverty level and delayed time to diagnosis.[38] Meanwhile, Verstappen et al did not find any difference in time to diagnosis based on socioeconomic status.[39] One study from India[33] and another from Germany[24] demonstrated an association between delayed diagnosis and more prolonged travel to reach a pediatric rheumatologist.

Why is JIA diagnostically complex?

Diagnostic delays in JIA are likely multifactorial (Figure 2). JIA is a heterogeneous condition; there are seven different categories of disease under the International League Against Rheumatism,[40] each of which has unique features with which generalists may not be familiar. For example, PsA can present with nail pitting and arthralgias but very subtle joint effusions, whereas ERA may present with axial symptoms and enthesitis in the absence of peripheral arthritis. Systemic arthritis, on the other hand, presents with

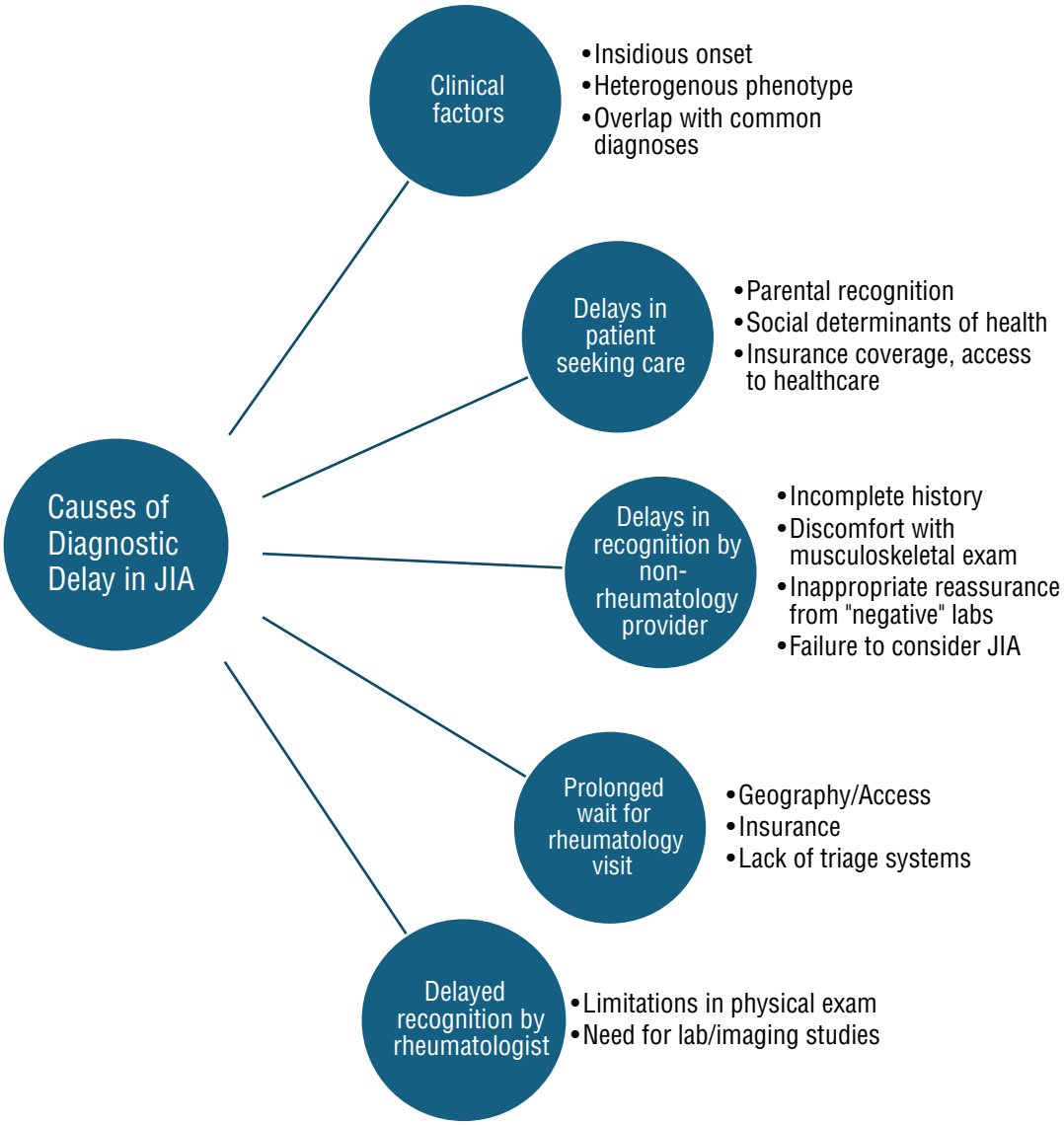


Figure 2. Conceptual model of factors contributing to diagnostic delay in JIA. JIA, juvenile idiopathic arthritis. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25438/abstract.

prolonged daily fevers and a subtle rash, and the development of arthritis may come weeks into the disease course.

Additionally, in many categories of JIA, the clinical presentation have insidious onset and up to 25% of children have little to no pain but instead describe morning stiffness or limp.[41] Parents may not notice mild findings especially in young children who do not have verbal skills to describe or localize their symptoms, and nonrheumatology providers may not ask about symptoms like morning stiffness as part of their routine history. In addition, many of the examination findings seen in JIA, such as joint effusions and enthesitis, are difficult to detect by those not accustomed to performing joint examinations, especially early in the disease course. Primary care providers may struggle to identify these findings especially because they have self-reported low confidence in the musculoskeletal examination.[42] Meanwhile, orthopedic clinicians, who have more expertise in the musculoskeletal examination, have the longest interval to referral to pediatric rheumatology.[24,32,36] Furthermore, clinicians may be falsely reassured by normal markers of inflammation and negative serologic testing when, in fact, most children with JIA have normal laboratory evaluations. In the previously mentioned qualitative study by Rapley et al, they performed semistructured interview with 10 clinicians (including orthopedists, pediatricians, pediatric immunologist, a general practitioner, and a nurse) initially involved in diagnosis. They reported consideration of JIA but being falsely reassured by many of the factors mentioned above including normal laboratory tests, well appearance of the child, or the lack of significant pain. Providers were more likely to consider a diagnosis of JIA if they had experience caring for a patient with JIA, rotated with Pediatric Rheumatology, or recollection of a lecture about the condition.[29]

Even once providers appropriately recognize JIA, extremely limited access to pediatric rheumatology in some areas of the world[43,44] may further delay diagnosis if waitlists are prolonged

or patients must travel a long distance for evaluation. Importantly, even once a patient establishes care with a pediatric rheumatologist, it may take some time for the rheumatologist to establish the final diagnosis. Even after a diagnosis is established, there may be further delays in initiating therapy, as pediatric patients often require sedation for joint injections and prior authorizations for biologic medications. Pitfalls at each step of this process can contribute to delayed diagnosis and treatment and, ultimately, impacts the outcomes of our patients.

What gaps in our knowledge remain?

The existing literature clearly demonstrates opportunities to improve timely and accurate diagnosis of JIA. Although the diagnostic course for these patients varies by country and health care system, it is often tumultuous, requiring many visits to physicians and parental persistence before a diagnosis is made.

Our review of the literature identified several gaps in our knowledge about the diagnostic trajectory for patients with JIA that limit our ability to start designing interventions to improve time to diagnosis. Although some interventions might be applicable across countries and health care systems, many of the recommended interventions may be specific to local systems. As such, it is notable that there is very little research on time to diagnosis in privatized health care systems in which referral patterns and access issues are likely different from those in countries with nationalized health care.

Conclusions

There is very little information about what initial and/or inaccurate diagnoses patients receive before their diagnosis with JIA and no information about what diagnoses are made by specific provider types. This information will be crucial in understanding the confounding diagnoses that prevent recognition of JIA and developing interventions to improve the diagnostic process. In addition, only one small qualitative study has attempted to understand the clinical reasoning of the providers who manage patients before they are referred to rheumatology and no studies used a safety two framework,[45,46] the framework dedicated to understanding when and how existing systems succeed, to understand the factors that promote timely diagnosis. It is our hope that further research in this area will elucidate the diagnostic trajectory for patients with JIA across health care settings, identify barriers and facilitators to diagnostic excellence, and allow targeted interventions to improve time to diagnosis and thereby improve disease outcomes.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published.

REFERENCES

1. Davies K, Cleary G, Foster H, et al; British Society of Paediatric and Adolescent Rheumatology. BSPAR Standards of Care for children and young people with juvenile idiopathic arthritis. Rheumatology (Oxford) 2010;49:1406–1408.
2. Petty RE, Laxer RM, Lindsley CB, et al. Textbook of Pediatric Rheumatology. 8th ed. Elsevier, Inc; 2020.
3. Minden K, Horneff G, Niewerth M, et al. Time of disease-modifying antirheumatic drug start in juvenile idiopathic arthritis and the likelihood of a drug-free remission in young adulthood. Arthritis Care Res (Hoboken) 2019;71:471–481.
4. Wallace CA, Giannini EH, Spalding SJ, et al. Trial of early aggressive therapy in polyarticular juvenile idiopathic arthritis. Arthritis Rheum 2012;64:2012–2021.
5. Consolaro A, Negro G, Lanni S, et al. Toward a treat-to-target approach in the management of juvenile idiopathic arthritis. Clin Exp Rheumatol 2012;30(4 suppl 73):S157–S162.
6. Wallace CA, Giannini EH, Spalding SJ, et al. Clinically inactive disease in a cohort of children with new-onset polyarticular juvenile idiopathic arthritis treated with early aggressive therapy: time to achievement, total duration, and predictors. J Rheumatol 2014;41:1163–1170.
7. Ravichandran N, Guleria S, Mohindra N, et al. Predictors of long-term functional outcomes of juvenile idiopathic arthritis–enthesitis-related arthritis: a single centre experience. Rheumatology (Oxford) 2023;62:3110–3116.
8. Burgers LE, Raza K, van der Helm-van Mil AH. Window of opportunity in rheumatoid arthritis - definitions and supporting evidence: from old to new perspectives. RMD Open 2019;5:e000870.
9. Haroon M, Gallagher P, FitzGerald O. Diagnostic delay of more than 6 months contributes to poor radiographic and functional outcome in psoriatic arthritis. Ann Rheum Dis 2015;74:1045–1050.
10. Yang D, Fineberg HV, Cosby K. Diagnostic excellence. JAMA 2021;326:1905–1906.
11. Committee on Diagnostic Error in Health Care, Board on Health Care Services, Institute of Medicine, The National Academies of Sciences, Engineering, and Medicine. Balogh EP, Miller BT, Ball JR, eds. Improving Diagnosis in Health Care. National Academies Press; 2015.
12. Hoffman JM, Keeling NJ, Forrest CB, et al. Priorities for pediatric patient safety research. Pediatrics 2019;143:e20180496.
13. Giardina TD, Hunte H, Hill MA, et al. Defining diagnostic error: A scoping review to assess the impact of the National Academies' Report Improving Diagnosis in Health Care. J Patient Saf 2022;18:770–778.
14. Singh H, Graber ML. Improving Diagnosis in Health Care–the next imperative for patient safety. N Engl J Med 2015;373:2493–2495.
15. Graber ML. The incidence of diagnostic error in medicine. BMJ Qual Saf 2013;22 Suppl 2(suppl 2):ii21–ii27.
16. Lovell DJ, Passo MH, Beukelman T, et al. Measuring process of arthritis care. A proposed set of quality measures for the process of care in juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2011;63:10–16.
17. The Royal Australian College of General Practitioners. Clinical guideline for the diagnosis and management of juvenile idiopathic arthritis. Published online 2009. Accessed 4/19/2024. https://www.racgp.org.au/FSDEDEV/media/documents/Clinical%20Resources/Guidelines/Joint%20replacement/Clinical-guideline-for-the-diagnosis-and-management-of-juvenile-idiopathic-arthritis.pdf
18. Chausset A, Pereira B, Echaubard S, et al. Access to paediatric rheumatology care in juvenile idiopathic arthritis: what do we know? A systematic review. Rheumatol Oxf Engl 2020;59:3633–3644.
19. Shiff NJ, Tucker LB, Guzman J, et al. Factors associated with a longer time to access pediatric rheumatologists in Canadian children with juvenile idiopathic arthritis. J Rheumatol 2010;37:2415–2421.

20. Ong MS, Schiff, G, Natter M. Incidence, contributing factors, and impact of diagnostic delay in juvenile idiopathic arthritis: analysis of the Childhood Arthritis and Rheumatology Research Alliance (CARRA) Registry. In: Arthritis Rheumatol 2020;72 suppl 4. https://acrabstracts.org/abstract/incidence-contributing-factors-and-impact-of-diagnostic-delay-in-juvenile-idiopathic-arthritis-analysis-of-the-childhood-arthritis-and-rheumatology-research-alliance-carra-registry/
21. Adib N, Hyrich K, Thornton J, et al. Association between duration of symptoms and severity of disease at first presentation to paediatric rheumatology: results from the Childhood Arthritis Prospective Study. Rheumatology (Oxford) 2008;47:991–995.
22. Villeneuve E, Nam JL, Bell MJ, et al. A systematic literature review of strategies promoting early referral and reducing delays in the diagnosis and management of inflammatory arthritis. Ann Rheum Dis 2013;72:13–22.
23. Shiff NJ, Abdwani R, Cabral DA, et al. Access to pediatric rheumatology subspecialty care in British Columbia, Canada. J Rheumatol 2009;36:410–415.
24. Tzaribachev N, Benseler SM, Tyrrell PN, et al. Predictors of delayed referral to a pediatric rheumatology center. Arthritis Rheum 2009;61:1367–1372.
25. Marino A, Baldassarre P, Ferrigno C, et al. Pre-rheumatology referral consultation and investigation pattern in children with joint complaints: focus on juvenile idiopathic arthritis. Children (Basel) 2024;11:600.
26. Barber CEH, Barnabe C, Benseler S, et al. Patient factors associated with waiting time to pediatric rheumatologist consultation for patients with juvenile idiopathic arthritis. Pediatr Rheumatol Online J 2020;18:22.
27. Foster HE, Eltringham MS, Kay LJ, et al. Delay in access to appropriate care for children presenting with musculoskeletal symptoms and ultimately diagnosed with juvenile idiopathic arthritis. Arthritis Rheum 2007;57:921–927.
28. McErlane F, Foster HE, Carrasco R, et al. Trends in paediatric rheumatology referral times and disease activity indices over a ten-year period among children and young people with juvenile idiopathic arthritis: results from the Childhood Arthritis Prospective Study. Rheumatology (Oxford) 2016;55:1225–1234.
29. Rapley T, May C, Smith N, et al. 'Snakes & Ladders': factors influencing access to appropriate care for children and young people with suspected juvenile idiopathic arthritis – a qualitative study. Pediatr Rheumatol Online J 2021;19:43.
30. Chausset A, Freychet C, Lohse A, et al. Diagnosis journey for children with juvenile idiopathic arthritis: a qualitative study. Arch Dis Child Published online August 22, 2024. doi: https://doi.org/10.1136/archdischild-2024-327426
31. Aoust L, Rossi-Semerano L, Koné-Paut I, et al. Time to diagnosis in juvenile idiopathic arthritis: a French perspective. Orphanet J Rare Dis 2017;12:43.
32. Freychet C, Lambert C, Pereira B, et al. Medical pathways of children with juvenile idiopathic arthritis before referral to pediatric rheumatology centers. Joint Bone Spine 2019;86:739–745.
33. Agarwal M, Freychet C, Jain S, et al. Factors impacting referral of JIA patients to a tertiary level pediatric rheumatology center in North India: a retrospective cohort study. Pediatr Rheumatol Online J 2020;18:21.
34. Frenkel Y, Kraushar I, Saied MH, et al. Obstacles in early diagnosis of children with juvenile idiopathic arthritis: a nationwide Israeli retrospective study. J Rheumatol 2023;50:799–803.
35. Khawaja K, Al-Maini M. Access to pediatric rheumatology care for juvenile idiopathic arthritis in the United Arab Emirates. Pediatr Rheumatol 2017;15:41.
36. Chausset A, Lambert C, Belot A, et al. Individual and environmental determinants associated with longer times to access pediatric rheumatology centers for patients with juvenile idiopathic arthritis, a JIR cohort study. Pediatr Rheumatol Online J 2023;21:24.
37. Shoop-Worrall SJW, Moull L, McDonagh JE, et al. The role of age in delays to rheumatological care in juvenile idiopathic arthritis. J Rheumatol 2022;49:1037–1041.
38. Balmuri N, Soulsby WD, Cooley V, et al; CARRA Registry Investigators. Community poverty level influences time to first pediatric rheumatology appointment in polyarticular juvenile idiopathic arthritis. Pediatr Rheumatol Online J 2021;19:122.
39. Verstappen SMM, Cobb J, Foster HE, et al. The association between low socioeconomic status with high physical limitations and low illness self-perception in patients with juvenile idiopathic arthritis: results from the Childhood Arthritis Prospective Study. Arthritis Care Res (Hoboken) 2015;67:382–389.
40. Petty RE, Southwood TR, Manners P, et al; International League of Associations for Rheumatology. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. J Rheumatol 2004;31:390–392.
41. Sherry DD, Bohnsack J, Salmonson K, et al. Painless juvenile rheumatoid arthritis. J Pediatr 1990;116:921–923.
42. Jandial S, Myers A, Wise E, et al. Doctors likely to encounter children with musculoskeletal complaints have low confidence in their clinical skills. J Pediatr 2009;154:267–271.
43. Correll CK, Ditmyer MM, Mehta J, et al. 2015 American College of Rheumatology Workforce Study and Demand Projections of Pediatric Rheumatology Workforce, 2015–2030. Arthritis Care Res (Hoboken) 2022;74:340–348.
44. Vilaiyuk S, Hadef D, Hamdi W, et al. The inequity of global healthcare in pediatric rheumatology. Best Pract Res Clin Rheumatol 2024;38:101983.
45. Choi JJ. What is diagnostic safety? A review of safety science paradigms and rethinking paths to improving diagnosis. Diagn Berl Ger Published online May 27, 2024. doi:https://doi.org/10.1515/dx-2024-0008
46. Anderson JE, Watt AJ. Using Safety-II and resilient healthcare principles to learn from Never Events. Int J Qual Health Care 2020;32:196–203.
47. Boiu S, Marniga E, Bader-Meunier B, Mouy R, Compeyrot-Lacassagne S, Quartier P, Wouters CH, Functional status in severe juvenile idiopathic arthritis in the biologic treatment era: an assessment in a French paediatric rheumatology referral centre. Rheumatology 2012;51(7):1285–1292.
48. Hyrich KL, Lal SD, Foster HE, Thornton J, Adib N, Baildam E, Gardner-Medwin J, Wedderburn LR, Chieng A, Davidson J, Thomson W. Disease activity and disability in children with juvenile idiopathic arthritis one year following presentation to paediatric rheumatology. Results from the Childhood Arthritis Prospective Study. Rheumatology 2009;49(1):116–122.
49. Kavirayani A, Foster HE. Paediatric rheumatology practice in the UK benchmarked against the British Society for Paediatric and Adolescent Rheumatology/Arthritis and Musculoskeletal Alliance Standards of Care for juvenile idiopathic arthritis. Rheumatology 2013;52(12):2203–2207.
50. Consolaro A, Giancane G, Alongi A, et al. Phenotypic variability and disparities in treatment and outcomes of childhood arthritis throughout the world: an observational cohort study. Lancet Child Adolesce 2019;3(4):255–263.

Arthritis Care & Research
Vol. 77, No. 3, March 2025, pp 291–296
DOI 10.1002/acr.25488

American College of Rheumatology
Empowering Rheumatology Professionals

CLINICOPATHOLOGIC CONFERENCE

The Matter at Hand: A Case of Difficult-to-Treat Arthritis

Ana L. Altaffer,[1] Lindsay C. Burrage,[1] Ankur Kamdar,[2] Tiphanie P. Vogel,[1] and Maria Pereira[1]

CASE PRESENTATION

History of the present illness

A 12-year-old boy with a history of longstanding oligoarticular juvenile idiopathic arthritis (JIA) presented to a new pediatric rheumatology clinic to establish care following a 2-year gap in access.

He originally presented with right wrist pain and stiffness without associated swelling when he was age 4 years, and he was diagnosed with antinuclear antibody (ANA)–negative oligoarticular JIA. Shortly after his initial diagnosis, he underwent an intra-articular corticosteroid injection without subsequent improvement to his pain. He failed trials of several systemic therapies, including oral prednisolone, methotrexate, and, subsequently, adalimumab. Despite these therapies, he had no symptomatic improvement and developed progressive loss of range of motion of the right wrist. No additional joints were involved. He did not have evidence of uveitis on screening eye examinations.

Over time, he and his mother noted that his right upper extremity appeared smaller than his left upper extremity. He was evaluated for brachial neuritis by both a neurologist and a physical medicine and rehabilitation specialist. However, an electromyography test was unremarkable, and a magnetic resonance imaging (MRI) scan of his brachial plexus demonstrated no abnormalities.

Upon re-establishing care, he reported resolution of his right wrist pain despite discontinuing all medications over the prior 2 years. However, he had developed worsening decreased range of motion of the right wrist, with associated weakness and difficulty with manual activities, such as opening water bottles. He denied joint swelling or morning stiffness. Consent for the use of patient data and images was obtained from the patient's parent. This case report was exempt from institutional review board approval.

Past medical history

The patient and his mother denied any previous medical diagnoses other than JIA. He had never had any surgeries. He had never had any fractures or serious injuries.

Social and family history

The patient was in the sixth grade and performing well in school. There was no family history of rheumatic diseases, including in two older siblings. Specifically, there was no history of JIA or other inflammatory arthropathies, known genetic disorders, or individuals with similar symptoms as the proband.

Review of systems

The patient and his mother denied any history of recurrent or persistent unexplained fevers, fatigue, or weight loss. He did not have any numbness or tingling. They denied any visual disturbances, eye redness or pain, or photophobia. They denied any rashes or skin changes. They denied any gross hematuria or peripheral edema. He had met all developmental milestones appropriately. A complete review of systems was otherwise unremarkable.

Physical examination

The patient's height was at the 93rd percentile, and his weight was at the 92nd percentile. He had upper extremity asymmetry, with muscle atrophy and shorter length of his right hand and wrist. The range of motion of his right wrist was severely restricted in both flexion and extension (Figure 1). Notably, he did not have overt swelling or tenderness to palpation of the right wrist. His comprehensive joint examination was otherwise normal, with no synovitis, effusions, joint tenderness, or reduced range of motion in any other joints.

[1]Ana L. Altaffer, MD, Lindsay C. Burrage, MD, PhD, Tiphanie P. Vogel, MD, PhD, Maria Pereira, MD: Baylor College of Medicine and Texas Children's Hospital, Houston; [2]Ankur Kamdar, MD: University of Texas Health Science Center at Houston and Children's Memorial Hermann Hospital.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25488.

Address correspondence via email to Maria Pereira, MD, at maria.pereira2@bcm.edu.

Submitted for publication September 25, 2024; accepted in revised form December 17, 2024.

Figure 1. Physical examination findings. (A) Limb length discrepancy and right upper extremity muscle atrophy. (B) Decreased size of right hand compared with left, with reduced range of motion of the right wrist. The patient was unable to lay his right hand flat on the examination table because of limited wrist extension.

Laboratory evaluation

A basic laboratory work-up was unremarkable, including negative rheumatoid factor, negative anticyclic citrullinated peptide, and normal markers inflammation (Table 1). An x-ray of his right wrist was notable for irregularity with collapse and sclerotic changes of his carpal bones, with marked joint-space narrowing. His metacarpal and phalangeal bones were normal in appearance. Similarly, a right wrist MRI scan demonstrated extensive erosive changes, decreased size, and abnormal configuration of the carpal bones with synovitis (Figure 2).

CASE SUMMARY

The patient is a 12-year-old boy with a longstanding history of ANA-negative oligoarticular JIA refractory to multiple therapies presenting to re-establish rheumatology care. He was found to have severely restricted range of motion of his right wrist, with extensive collapse, sclerotic changes, and abnormal configuration of his right carpal bones noted on imaging studies but without other joint involvement.

Table 1. Laboratory evaluation

Test	Reference range	Patient's results
White blood cell count, $\times 10^3/\mu L$	3.7–10.5	8.2
Hemoglobin, g/dL	11.7–15.7	14.6
Platelet count, $\times 10^3/\mu L$	150–450	320
Antinuclear antibody screen	Negative	Negative
Rheumatoid factor, IU/mL	0.0–13.9	<10.0
Anticyclic citrullinated peptide, units	<20	<20
Erythrocyte sedimentation rate, mm/h	0–15	7
C-reactive protein, mg/L	0–7	1

DIFFERENTIAL DIAGNOSIS

Persistent oligoarticular JIA. This patient presented with a reported history of persistent oligoarticular JIA. According to the International League of Associations for Rheumatology JIA classification, oligoarticular JIA involves four or fewer joints in the first 6 months. Patients are considered to have persistent oligoarticular JIA if the disease does not extend to involve more than four joints after the first 6 months.[1] Much like our patient, most patients with oligoarticular JIA present before age 6 years with persistent joint swelling.[2] Patients who are inadequately treated can develop destructive joint disease over time. The patient's carpal bone abnormalities could, therefore, be the sequelae of chronic, untreated arthritis.

However, several elements of his presentation and response to therapy were atypical for JIA. Monoarticular wrist involvement is not common.[3,4] In fact, wrist involvement on presentation is often considered a risk factor for the development of polyarticular disease.[5] The intensity of his disease, with destructive monoarticular wrist involvement, was also unusual. His imaging findings, including the near-total collapse of his carpal bones, were striking in

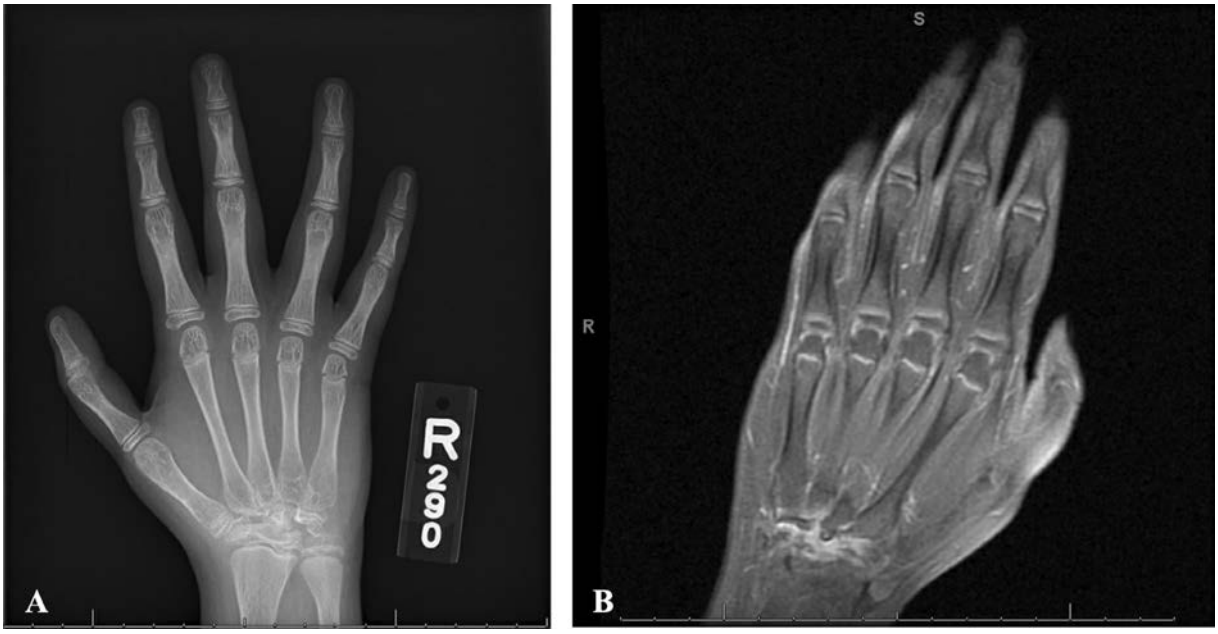


Figure 2. Imaging studies. (A) Plain radiograph of the right hand (posterior-anterior view) demonstrating collapse and sclerotic change of the carpal bones as well as joint-space narrowing of the carpal rows. (B) Magnetic resonance imaging scan of the right hand (coronal T1 fat-saturated postgadolinium) demonstrating small size and abnormal configuration of the carpal bones.

their severity. Finally, his lack of response to multiple anti-inflammatory agents, including intra-articular and systemic glucocorticoids, raises consideration of noninflammatory conditions affecting the joint. Given that his chronic disease course made infections and malignancies less likely, genetic and metabolic JIA mimics were considered next on the differential diagnosis.

Multicentric carpotarsal osteolysis. Multicentric carpotarsal osteolysis (MCTO) is a rare skeletal dysplasia caused by heterozygous pathogenic variants in *MAFB* and characterized by osteolysis of the carpal and/or tarsal bones. Less often, the bones in the shoulders, elbows, or knees can also be affected. Osteolysis is typically observed in more than one location (ie, it is multicentric) and is usually symmetric. Subtle craniofacial differences, such as triangular face, micrognathia, and exophthalmos, have also been described but are not always present.[6] Ocular and renal disease have been reported.[6,7] Patients with MCTO often present with joint swelling and stiffness in early childhood, and they can be misdiagnosed with JIA.[6,7] Although this patient did not have multifocal disease on clinical assessment, at least one case of unifocal disease had been reported.[8]

Farber disease. Farber disease (FD) is an autosomal recessive lysosomal storage disorder caused by biallelic pathogenic variants in the gene-encoding acid ceramidase, *ASAH1*.[9] Acid ceramidase deficiency leads to accumulation of ceramide in tissues. There are multiple distinct subtypes of FD with different patterns of organ involvement ranging from progressive neurologic deterioration to interstitial pneumonia. Type 1 FD is considered a genetic mimic of JIA. The classic triad of manifestations of Type 1 FD includes joint involvement, subcutaneous skin nodules near joints or over points of mechanical pressure, and progressive hoarseness owing to laryngeal involvement. Patients often first present in infancy and develop joint contractures and severe joint deformities over time.[10]

This patient did not have polyarticular involvement, as would have been expected with FD. He also lacked other classic manifestations, particularly nodules. Although patients can sometimes present initially with joint disease alone and develop other manifestations over time, it would be unusual to have no other manifestations by age 12 years.

Progressive pseudorheumatoid dysplasia. Progressive pseudorheumatoid dysplasia (PPRD) is an autosomal recessive skeletal dysplasia caused by biallelic pathogenic variants in *CCN6* that lead to altered cartilage homeostasis. Patients with PPRD are typically healthy at birth. They first present in childhood with enlarged interphalangeal joints, gait abnormalities, and joint stiffness, most often of the interphalangeal joints, knees, and hips. They can therefore be misdiagnosed with JIA. This skeletal disease is progressive, and, by adulthood, most patients have kyphoscoliosis, short stature with a short trunk, diffuse joint contractures with restricted range of motion of most joints, and camptodactyly.[11]

Although most joints end up affected with progressive disease in PPRD, wrist involvement at onset, as seen in this patient, is not typical.[11] Although this patient had restricted range of

motion of his right wrist, he had normal range of motion of all other joints, normal gait, and normal growth.

Type II collagen disorders. Type II collagen disorders are a spectrum of diseases caused by heterozygous pathogenic variants in *COL2A1*, which encodes the alpha-1 chain of type II collagen. These disorders are generally characterized by skeletal dysplasia, eye disease, and hearing impairment. However, their manifestations are diverse, ranging from severe disorders that are often lethal in the perinatal period to milder disorders that present with early-onset osteoarthritis in adolescence or adulthood.[12]

As type II collagen is a component of cartilage, it is perhaps not surprising that type II collagen disorders can present with joint dysfunction. Patients with Stickler syndrome type 1, which can present with a range of manifestations and severities, can develop early-onset arthropathy. This patient is less likely to have Stickler syndrome without associated hearing loss, cleft palate, or ocular manifestations, such as myopia, retinal detachment, and early-onset cataracts.[12]

PATIENT'S COURSE

Following his initial evaluation, the patient was started on tocilizumab for management of presumed uncontrolled JIA, as he had not previously been tried on interleukin (IL)-6 inhibition. He was also referred to hand surgery for assistance with management of his restricted range of motion, which interfered with his activities of daily living. He underwent right forearm tendon lengthening followed by serial casting and occupational therapy (OT). He reported functional improvement following his procedure and course of OT. Finally, he was referred to genetics to be evaluated for genetic and metabolic mimics of JIA.

A commercially available skeletal dysplasia genetic panel identified a heterozygous *MAFB* pathogenic variant (NM_005461.4: c.187C>T, p.Pro63Ser), diagnostic of MCTO. Parental testing confirmed that this variant was de novo.

Following his genetic diagnosis, a skeletal survey was performed and did not demonstrate any additional skeletal abnormalities. His tocilizumab was discontinued, and he did not have any changes in his clinical status following drug discontinuation. He was started on vitamin D supplementation to optimize his bone health. Screening for additional organ involvement, including an ophthalmologic examination and a renal evaluation, was unremarkable. He will continue to have annual follow-up visits for surveillance. He was referred to physical therapy to help prevent additional muscle atrophy from disuse.

DISCUSSION

We have presented the case of a 12-year-old boy with a presumed history of longstanding ANA-negative oligoarticular JIA refractory to multiple therapies who was found instead to have a genetic mimic of JIA, MCTO.

Patients with MCTO typically present in early childhood with joint pain, swelling, and stiffness, symptoms that often lead to a misdiagnosis of JIA. They develop progressive osteolysis of the carpal and/or tarsal bones. Other joints, such as the shoulders, elbows, and knees, are variably affected.[6,7] Over time, focal joint destruction can lead to impaired growth and functional impairment of the affected area. This patient's presentation was atypical for MCTO in that he did not have multicentric disease. Nevertheless, he has a molecular confirmation, and unifocal disease has been previously reported.[8]

There is significant clinical heterogeneity in the cases of MCTO described in the literature, with different degrees of skeletal involvement and functional impairment.[13] In most cases, progressive skeletal deformities occur during childhood and adolescence, but case reports describing long-term follow-up into adulthood are limited.

MCTO results from heterozygous pathogenic variants in the *MAFB* gene,[7,14] but the pathogenesis is not completely understood. The gene encodes a transcription factor, MafB, that has multiple roles, including negative regulation of osteoclast differentiation.[15] It is still unclear, however, why the osteolysis of MCTO is most prominent in the carpal and tarsal bones. A component of altered bone formation at these sites, in addition to osteolysis, has been hypothesized.[7] Indeed, differential expression of the transcription factor MafB in wrist bones during endochondral ossification has been demonstrated in healthy mice, supporting the notion that distinct mechanisms of bone formation in the wrist may be responsible for the distribution of bone involvement in MCTO.[16] The role of inflammation in this disease is also unclear. However, the fact that MCTO mimics JIA and that MRI findings of joint inflammation have been documented[17,18] suggests that it may play a role. Interestingly, this patient's wrist MRI scan demonstrated synovitis.

Once this patient's genetic diagnosis was made, it was necessary to decide whether to continue his tocilizumab, which had been prescribed for presumed JIA. Just as a role for inflammation in the pathogenesis of MCTO has not yet been elucidated, the role of anti-inflammatory agents in the management of MCTO remains unclear. Although some patients have reported symptomatic improvement with anti–tumor necrosis factor and anti–IL-6 agents, evidence that it deters further osteolysis is lacking.[17,18] Given his lack of response to therapy, the lack of strong evidence to support benefit in the context of immune suppression, and the need for ongoing safety monitoring, his tocilizumab was discontinued.

Although he was started on vitamin D supplementation to optimize bone health, our patient was not started on any additional systemic therapies following his diagnosis. There is currently no standard of care for MCTO. Antiresorptive agents, such as bisphosphonates and denosumab, have been attempted for MCTO management. However, as with anti-inflammatory agents,

Table 2. Mimics of juvenile idiopathic arthritis*

Diagnosis	Gene	Inheritance pattern	Clinical manifestations	Ref no.
Multicentric carpotarsal osteolysis	*MAFB*	AD	Osteolysis of carpal and tarsal bones and ocular and renal disease	6
Farber disease	*ASAH1*	AR	Joint contractures, subcutaneous nodules, and hoarse voice	9,10
Progressive pseudorheumatoid dysplasia	*CCN6*	AR	Joint contractures, camptodactyly, kyphoscoliosis, and short stature	11
Type II collagen disorders	*COL2A1*	AD	Arthrosis, eye disease, and hearing impairment	12
Nodulosis, arthropathy, and osteolysis syndrome	*MMP2* and *MMP14*	AR	Osteolysis of the hands and feet and palmar and plantar nodular lesions	24
Camptodactyly-arthropathy-coxa vara-pericarditis syndrome	*PRG4*	AR	Camptodactyly of fifth finger, joint swelling, coxa vara, and acetabular cysts	25
Pachydermodactyly	Unknown		Asymptomatic fusiform swelling of the soft tissues over the lateral aspect of the proximal interphalangeal joints	26
Mucopolysaccharidosis	Various (inborn lysosomal storage disorders resulting in improper processing of complex carbohydrates)	AR/XL	Joint contractures, dysplasia, and dysostosis multiplex	27
Holt-Oram syndrome	*TBX5*	AD	Upper limb defects, abnormal carpal bones, congenital heart malformation, and cardiac conduction disease	28

* AD, autosomal dominant; AR, autosomal recessive; Ref, reference; XL, X-linked.

they have had little success in preventing further focal bone destruction.[8,19,20] They may play a role in addressing generalized osteopenia, which can be observed in patients with MCTO.[8,19] Surgical interventions can be considered in some situations and may offer both symptomatic relief and improved function.[21]

A multidisciplinary approach to care for patients with MCTO is recommended, as they require surveillance for disease complications. Up to two-thirds of patients with MCTO develop kidney disease (focal segmental glomerulosclerosis), and approximately half of those can progress to kidney failure.[6] This may be because the transcription factor MafB regulates podocyte survival in the kidneys.[22,23] Ocular complications, such as corneal opacities, have also been reported.[6] There are no current guidelines for renal and ophthalmologic screening of patients with MCTO. This patient has had normal renal and ophthalmologic evaluations. Nevertheless, he continues to follow serially with his pediatric rheumatologist, nephrologist, ophthalmologist, and clinical geneticist. He undergoes yearly screening urinalyses and chemistry panels.

This case illustrates features of MCTO, a skeletal dysplasia that mimics JIA. It highlights that a diagnosis of MCTO should be considered in patients with localized destruction of the carpal and/or tarsal bones, especially, although not exclusively, when in association with kidney disease and/or craniofacial differences. More broadly, this case also supports the need to consider genetic and metabolic JIA mimics (Table 2) when evaluating patients with treatment-refractory disease. Genetic testing should be considered early in patients with atypical joint disease that does not respond to conventional therapies, as timely identification of genetic diagnoses can guide therapy, surveillance for complications, and genetic counseling for the patient and family.

FINAL DIAGNOSIS

Multicentric carpotarsal osteolysis syndrome

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 Pereira 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. Petty RE, Southwood TR, Manners P, et al; International League of Associations for Rheumatology. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. J Rheumatol 2004;31(2):390–392.
2. Martini A, Lovell DJ, Albani S, et al. Juvenile idiopathic arthritis. Nat Rev Dis Primers 2022;8(1):5.
3. Stoll ML, Nigrovic PA, Gotte AC, et al. Clinical comparison of early-onset psoriatic and non-psoriatic oligoarticular juvenile idiopathic arthritis. Clin Exp Rheumatol 2011;29(3):582–588.
4. Huemer C, Malleson PN, Cabral DA, et al. Patterns of joint involvement at onset differentiate oligoarticular juvenile psoriatic arthritis from pauciarticular juvenile rheumatoid arthritis. J Rheumatol 2002;29(7): 1531–1535.
5. Felici E, Novarini C, Magni-Manzoni S, et al. Course of joint disease in patients with antinuclear antibody-positive juvenile idiopathic arthritis. J Rheumatol 2005;32(9):1805–1810.
6. Wu J, Wang L, Xu Y, et al. Multicentric carpo-tarsal osteolysis syndrome mimicking juvenile idiopathic arthritis: two case reports and review of the literature. Front Pediatr 2021;9:745812.
7. Ma NS, Mumm S, Takahashi S, et al. Multicentric carpotarsal osteolysis: a contemporary perspective on the unique skeletal phenotype. Curr Osteoporos Rep 2023;21(1):85–94.
8. Regev R, Sochett EB, Elia Y, et al. Multicentric carpotarsal osteolysis syndrome (MCTO) with generalized high bone turnover and high serum RANKL: response to denosumab. Bone Rep 2021;14:100747.
9. Alves MQ, Le Trionnaire E, Ribeiro I, et al. Molecular basis of acid ceramidase deficiency in a neonatal form of Farber disease: identification of the first large deletion in ASAH1 gene. Mol Genet Metab 2013; 109(3):276–281.
10. Ehlert K, Frosch M, Fehse N, et al. Farber disease: clinical presentation, pathogenesis and a new approach to treatment. Pediatr Rheumatol Online J 2007;5(1):15.
11. Torreggiani S, Torcoletti M, Campos-Xavier B, et al. Progressive pseudorheumatoid dysplasia: a rare childhood disease. Rheumatol Int 2019;39(3):441–452.
12. Barat-Houari M, Sarrabay G, Gatinois V, et al. Mutation update for COL2A1 gene variants associated with type II collagenopathies. Hum Mutat 2016;37(1):7–15.
13. Mehawej C, Courcet JB, Baujat G, et al. The identification of MAFB mutations in eight patients with multicentric carpo-tarsal osteolysis supports genetic homogeneity but clinical variability. Am J Med Genet A 2013;161A(12):3023–3029.
14. Zankl A, Duncan EL, Leo PJ, et al. Multicentric carpotarsal osteolysis is caused by mutations clustering in the amino-terminal transcriptional activation domain of MAFB. Am J Hum Genet 2012;90(3):494–501.
15. Kim K, Kim JH, Lee J, et al. MafB negatively regulates RANKL-mediated osteoclast differentiation. Blood 2007;109(8):3253–3259.
16. Lazarus S, Tseng HW, Lawrence F, et al. Characterization of normal murine carpal bone development prompts re-evaluation of pathologic osteolysis as the cause of human carpal-tarsal osteolysis disorders. Am J Pathol 2017;187(9):1923–1934.
17. Nishikomori R, Kawai T, Toshiyuki K, et al. Remarkable improvement of articular pain by biologics in a multicentric carpotarsal osteolysis patient with a mutation of MAFB gene. Pediatr Rheumatol 2015; 13(S1):P152.
18. Upadia J, Gomes A, Weiser P, et al. A familial case of multicentric carpotarsal osteolysis syndrome and treatment outcome. J Pediatr Genet 2018;7(4):174–179.
19. Park PG, Kim KH, Hyun HS, et al. Three cases of multicentric carpotarsal osteolysis syndrome: a case series. BMC Med Genet 2018; 19(1):164.
20. Lerman MA, Francavilla M, Waqar-Cowles L, et al. Denosumab treatment does not halt progression of bone lesions in multicentric carpotarsal osteolysis syndrome. JBMR Plus 2023;7(5):e10729.
21. Goldfarb CA, Steffen JA, Whyte MP. Idiopathic multicentric osteolysis: upper extremity manifestations and surgical considerations during childhood. J Hand Surg Am 2012;37(8):1677–1683.
22. Usui T, Morito N, Shawki HH, et al. Transcription factor MafB in podocytes protects against the development of focal segmental glomerulosclerosis. Kidney Int 2020;98(2):391–403.
23. Sadl V, Jin F, Yu J, et al. The mouse Kreisler (Krml1/MafB) segmentation gene is required for differentiation of glomerular visceral epithelial cells. Dev Biol 2002;249(1):16–29.
24. Al-Mayouf SM, Majeed M, Hugosson C, et al. New form of idiopathic osteolysis: nodulosis, arthropathy and osteolysis (NAO) syndrome. Am J Med Genet 2000;93(1):5–10.
25. Shayan K, Ho M, Edwards V, et al. Synovial pathology in camptodactyly-arthropathy-coxa vara-pericarditis syndrome. Pediatr Dev Pathol 2005;8(1):26–33.
26. Dallos T, Oppl B, Kovács L, et al. Pachydermodactyly: a review. Curr Rheumatol Rep 2014;16(9):442.
27. Morishita K, Petty RE. Musculoskeletal manifestations of mucopolysaccharidoses. Rheumatology (Oxford) 2011;50(suppl 5):v19–v25.
28. Vanlerberghe C, Jourdain AS, Ghoumid J, et al. Holt-Oram syndrome: clinical and molecular description of 78 patients with TBX5 variants. Eur J Hum Genet 2019;27(3):360–368.

Arthritis Care & Research
Vol. 77, No. 3, March 2025, pp 297–299
DOI 10.1002/acr.25434

American College of Rheumatology
Empowering Rheumatology Professionals

EDITORIAL

Is Collaborative Care Better Care?

James T. Rosenbaum,[1] Nicole Fett,[2] Daniela Ghetie,[2] and Julianna Desmarais[2]

In offering care to patients, providers render advice; unfortunately, and despite the noblest of intentions, sometimes that advice is wrong. For women, taking estrogen during or after menopause has obvious health benefits. But we have learned that estrogen can increase the risk of heart disease, stroke, and certain cancers.[1] Additionally, how to apply or interpret the implications of this risk remains in flux.[2] Is it not intuitively obvious that suppressing premature ventricular contractions during a heart attack would prolong life by preventing fatal arrhythmias? Except what is obvious in this instance is also wrong.[3] Beta blockers have a negative inotropic effect, so prescribing one to a patient with heart failure in the setting of an acute myocardial infarction would certainly be harmful. Except the data indicate that a beta blocker increases survival in such a patient.[4] What is intuitively obvious is not always correct.

In this issue of *Arthritis Care & Research*, Lavallee and colleagues from the Children's Hospital of Philadelphia report something that seems equally apparent regarding the care of a patient with a multisystem disease: when two experts collaborate by sharing different but mutually relevant areas of expertise, health care improves.[5] Lavallee et al reach this conclusion by retrospectively studying patient outcomes in a clinic for children with juvenile idiopathic arthritis. All the children in this clinic had uveitis, a synonym for intraocular inflammation. Thus, all the children were under the care of an ophthalmologist. However, a rheumatologist could also contribute to patient management, such as by providing greater knowledge or experience with immunosuppressive therapy. About a quarter of the patients were seen in a collaborative setting to optimize the communication between the two specialists. As judged by objective findings, including reduction of inflammatory cells in the anterior chamber of the eye, reduced incidence of complications such as cataract formation, and a 64% reduction in physician visits, care improved when the ophthalmologist and rheumatologist combined their collective wisdom. Although other publications have described interdisciplinary clinics,[6–14] the article by Lavallee and colleagues[5] stands out by virtue of studying a rather uniform patient population and relying on a variety of objective outcome measures.

A retrospective study is fraught with problems. For example, the reasons that some families elected care in two separate locations might have been associated with some sort of selection bias. Frequently, in retrospective studies, critical outcome measures have not been recorded. The duration of follow-up varies. Although recommendations for collaborative care are sometimes incorporated into guidelines, there is a dearth of data to support the conclusion that interdisciplinary care is better care. Obstacles to a prospective study on interdisciplinary care include the ethics of potentially randomizing to care, which many believe to be inferior, and identifying end points that are likely to separate the two different approaches. If the study comes from a single center, there is the pitfall that the findings might be determined primarily by a specific physician and thus do not represent a definitive test of the hypothesis.

Creating an interdisciplinary clinic is simple in concept and extremely challenging in actualization. How do you achieve time management for two or more specialists to evaluate patients efficiently? Space management to ensure that rooms are not over- or underutilized? Understanding so that the patient appreciates the benefit of two physicians collaborating and is not surprised by two separate bills? And, potentially most challenging, consensus in advice such that the personalities of the practitioners complement each other and their different skillsets create optimum care, not antagonistic care?

The study by Lavallee and colleagues has flaws but still fills a major unmet need for data on the utility of interdisciplinary clinics. A separate question is whether observations on collaboration between a rheumatologist and ophthalmologist can be extrapolated to similar interactions with dermatologists, nephrologists, pulmonologists, neurologists, cardiologists, and other subspecialists who might provide insight helpful to a patient. The vast majority of publications on interdisciplinary care[6–14] describe dermatology–rheumatology collaboration. For example, Argobi and colleagues[10] noted that this paradigm resulted in 95% of

[1]James T. Rosenbaum, MD: Legacy Devers Eye Institute, Portland, Oregon; [2]Nicole Fett, MD, Daniela Ghetie, MD, Julianna Desmarais, MD: Oregon Health & Science University, Portland.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25434.

Address correspondence via email to James T. Rosenbaum, MD, at rosenbaumjames04@gmail.com.

Submitted for publication September 9, 2024; accepted in revised form September 11, 2024.

attendees having a more complete examination and consequently "improved management." Foti et al[11] described a clinic for patients with psoriasis and/or psoriatic arthritis. They felt that the collaboration resulted in improved metrics and more frequent use of biologics or disease-modifying antirheumatic drugs. Brazzelli et al[6] concluded that they could establish a "diagnosis of minimal psoriasis based on Caspar criteria in over 70% of the patients" with undifferentiated arthritis, whereas Samycia and colleagues[8] reported on 320 patients and established a dermatologic diagnosis "often unrelated to the underlying rheumatologic diagnosis." Our experience with multidisciplinary care convinces us that such a clinic improves patient care, facilitates teaching, and potentially creates a wealth of observations, enhanced understanding of the disease process, and collaborative research opportunities.

Some examples from our personal experience include the following:

1. A young male patient presented in cardiac arrest and was found later to have diffuse, infiltrative cardiac sarcoidosis. A clinic with cardiologists, rheumatologists, and pulmonologists helped to optimize his complex care and resulted in successful amelioration of his underlying disease.
2. A woman in her sixties had severe progressive myositis requiring a feeding tube. Studies for autoantibodies were negative. She was treated with four days of intravenous methylprednisolone, and a muscle biopsy was arranged. The biopsy did not show inflammatory disease, and the illness remained a mystery. She developed erythroderma during treatment with oral glucocorticoids, intravenous immunoglobulin, and mycophenolate mofetil. A skin biopsy was consistent with a connective tissue disease. Given her presentation, her diagnosis was likely dermatomyositis. The results of the skin biopsy prompted a repeat electromyogram with a guided muscle biopsy that led to confirmation of the dermatomyositis diagnosis, ensuring that the patient received correct therapy and screening. An interdisciplinary clinic with dermatology facilitated the necessary communication and allowed the diagnosis to be determined expeditiously.
3. A woman in her fifties with long-standing but well-controlled granulomatosis with polyangiitis (GPA) developed several large, painful, rapidly progressive ulcers of the lower extremities. Dermatology evaluation was consistent with pyoderma gangrenosum–like ulcerations, which occur in GPA as a harbinger of flare. The skin diagnosis prompted the rheumatologist to confidently and successfully escalate GPA therapy for the underlying vasculitis. By meeting together, the dermatologist and rheumatologist shared knowledge so that the relationship between the skin and vasculitis was understood quickly and the management could proceed appropriately.

What attracts a trainee to choose rheumatology as a career? Some enjoy the personal relationship and integral role that a rheumatologist can play in helping a patient cope with a chronic illness. Some are attracted by the improving outcomes that are being achieved in treating diseases, which seemed for generations to defy the best efforts of care. Some are fascinated by the science of immunology and how that science relates to many rheumatic diseases. Many enjoy the challenge of a multisystem illness because the practitioner is forced to maintain a broad knowledge of medicine and pathophysiology. Because diseases such as systemic lupus, rheumatoid arthritis, progressive systemic sclerosis, vasculitis, sarcoidosis, inflammatory myositis, and spondyloarthropathy frequently involve multiple organ systems, rheumatologists collaborate with other subspecialists, as well as with allied health personnel such as a physical therapist or a social worker, to optimize care. The practice of medicine requires lifelong learning; peers are ideal contributors to this learning. Better outcomes lead to more satisfying care and career.

We each aspire to practice evidence-based medicine. Lavallee and colleagues provide support for a seemingly obvious conclusion: Two heads are better than one. Obvious conclusions are not always correct conclusions, and proving an obvious conclusion is not nearly as simple as it may seem. It is reassuring to read evidence that collaboration is good. Enjoying the rewards of collaboration is a major reason to practice rheumatology.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published.

REFERENCES

1. Banks E, Canfell K. Invited commentary: hormone therapy risks and benefits—the Women's Health Initiative findings and the postmenopausal estrogen timing hypothesis. Am J Epidemiol 2009;170(1):24–28.
2. Vigneswaran K, Hamoda H. Hormone replacement therapy – current recommendations. Best Pract Res Clin Obstet Gynaecol 2022;81:8–21.
3. Echt DS, Liebson PR, Mitchell LB, et al. Mortality and morbidity in patients receiving encainide, flecainide, or placebo. The Cardiac Arrhythmia Suppression Trial. N Engl J Med 1991;324(12):781–788.
4. Gheorghiade M, Goldstein S. Beta-blockers in the post-myocardial infarction patient. Circulation 2002;106:394–398.
5. Lavallee C, Ahrens M, Davidson S, et al. Measurable outcomes of an ophthalmology and rheumatology coordinated care clinic. Arthritis Care Res (Hoboken) 2025;77(3):322–329.
6. Brazzelli V, Bobbio Pallavicini F, Maggi P, et al. A multidisciplinary dermatology-gastroenterology-rheumatology (DER.RE.GA) unit for the care of patients with immune-mediated inflammatory diseases: analysis of the first 5 years from the dermatologist's perspective. Front Med (Lausanne) 2023;10:1290018.

7. Campagna M, Castillo R, Mattessich S, et al. Impact of a combined dermatology-rheumatology clinic on management of autoimmune connective tissue disorders. Clin Exp Dermatol 2019;44(2):e24–e25.
8. Samycia M, McCourt C, Shojania K, et al. Experiences from a combined dermatology and rheumatology clinic: a retrospective review. J Cutan Med Surg 2016;20(5):486–489.
9. Theodorakopoulou E, Dalamaga M, Katsimbri P, et al. How does the joint dermatology-rheumatology clinic benefit both patients and dermatologists? Dermatol Ther 2020;33(3):e13283.
10. Argobi Y, Smith GP. Analysis of the first 5 years of an interdisciplinary rheumatology-dermatology clinic. Clin Exp Dermatol 2021;46(2):270–275.
11. Foti R, Giuffrida G, Ramondetta A, et al. Multidisciplinary rheumaderma clinics: 5 years of experience at the San Marco Hospital in Catania. Dermatol Ther (Heidelb) 2022;12(12):2829–2837.
12. Mendes-Bastos P, Nero P, Ferreira P, et al. A multidisciplinary approach in psoriatic disease: the different models of dermatology-rheumatology collaborations in Portugal. Acta Reumatol Port 2021;46(4):333–341.
13. Goedvolk CA, von Rosenstiel IA, Bos AP. Immune complex associated complications in the subacute phase of meningococcal disease: incidence and literature review. Arch Dis Child 2003; 88(10):927–930.
14. Melo AT, Dourado E, Campanilho-Marques R, et al. Myositis multidisciplinary clinic in a tertiary referral center. J Multidiscip Healthc 2023; 16:1127–1139.

Arthritis Care & Research
Vol. 77, No. 3, March 2025, pp 300–308
DOI 10.1002/acr.25441

American College of Rheumatology
Empowering Rheumatology Professionals

Assessing Methotrexate Adherence in Juvenile Idiopathic Arthritis Using Electronic Health Record–Linked Pharmacy Dispensing Data

Dori Abel,[1] David Anderson,[1] Michael J. Kallan,[2] Levon Utidjian,[1] Jon M. Burnham,[1] Joyce C. Chang,[3] Chén C. Kenyon,[1] and Sabrina Gmuca[1]

Objective. We linked pharmacy dispensing data to clinical data in the electronic health record (EHR) to (1) identify characteristics associated with adherence to methotrexate (MTX) and (2) determine the association between adherence and disease activity in patients with juvenile idiopathic arthritis (JIA).

Methods. We conducted a single-center retrospective cohort study of incident MTX recipients with JIA treated between January 2016 and September 2023 for ≥12 months. Using pharmacy dispensing data, complemented by EHR data, we estimated adherence using medication possession ratios (MPRs) over the first 365 days of treatment. We used Fisher's exact and Wilcoxon rank-sum tests to compare patient characteristics between adherent (MPR ≥80%) and nonadherent (MPR <80%) groups and multivariable linear regression to investigate associations between MPR and active joint count.

Results. Among 224 patients, 81 (36.2%) were classified as nonadherent. In bivariate analysis, patients of younger age, of Black race, and from areas with lower child opportunity index were more likely to be classified as nonadherent. In multivariable analysis, active joint count changed from baseline to 12-month follow-up by −0.38 joints in the adherent compared to nonadherent group (95% confidence interval [CI] −0.74 to −0.01) and by −1.18 joints in patients with polyarticular course (95% CI −2.23 to −0.13).

Conclusion. Linking dispense data to clinical EHR data offers a novel, objective method for evaluating adherence to chronic medications. We identified demographic and area-level determinants of adherence, along with small but statistically significant differences in JIA disease activity measures by adherence status. Future work is needed to evaluate adherence as a potential mediator of known outcome disparities for socially disadvantaged populations.

INTRODUCTION

Children with juvenile idiopathic arthritis (JIA) frequently require long-term immunosuppressive therapies to preserve joint function and prevent physical disability.[1] Although biologic and nonbiologic disease-modifying antirheumatic drugs (DMARDs) are highly effective in controlling disease and preventing physical disability,[2–5] many children experience persistently active disease, with only about half of patients attaining remission off medications within five years.[3] Even in patients who do attain inactive disease status, maintenance of remission off medication is uncommon.[6] These outcomes are worse for patients who are Black, Hispanic, Medicaid-insured, or of lower socioeconomic status.[7,8] Although differences in adherence to chronic medications by demographic characteristics and social determinants of health (SDOH) have been reported in other chronic conditions in childhood,[9,10] this has not been studied in JIA, highlighting the need for further investigations into health disparities in JIA to direct interventions aimed

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

Supported by a pilot grant from Clinical Futures, a Research Center of Emphasis at The Children's Hospital of Philadelphia. Drs Abel, Chang, Kenyon, and Gmuca's work was supported by the NIH (grants 5-T32-HS-026116, K23-HL-148539, K23-HL-136842, and K23-AR-081409, respectively).

[1]Dori Abel, MD, MSHP, David Anderson, MD, Levon Utidjian, MD, MBI, Jon M. Burnham, MD, MSCE, Chén C. Kenyon, MD, MSHP, Sabrina Gmuca, MD, MSCE: The Children's Hospital of Philadelphia and The Perelman School of Medicine, University of Pennsylvania, Philadelphia; [2]Michael J. Kallan, MS: The Perelman School of Medicine, University of Pennsylvania, Philadelphia; [3]Joyce C. Chang, MD, MSCE: The Perelman School of Medicine, University of Pennsylvania, Philadelphia, and Boston Children's Hospital, Harvard Medical School, Boston, Massachusetts.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25441).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25441.

Address correspondence via email to Dori Abel, MD, MSHP, at abeld@chop.edu.

Submitted for publication March 15, 2024; accepted in revised form September 19, 2024.

SIGNIFICANCE & INNOVATIONS

- Using pharmacy dispense data offers a novel, clinically relevant, and objective tool for assessing patients' adherence to methotrexate (MTX) by calculating metrics such as the medication possession ratio (MPR).
- We successfully linked aggregated pharmacy dispense data to clinical data available within the electronic health record, allowing us to investigate associations between MTX adherence, patient and area-level characteristics, and disease activity measures.
- We demonstrated significant associations between individual demographic and area-level factors and nonadherence (MPR <80%) and significant differences between adherence status and change in active joint count.
- Given that medication adherence has significant implications across a wide range of diseases other than juvenile idiopathic arthritis, this methodology using electronic health record–enabled pharmacy dispense data to estimate adherence can be generalized to other populations.

at mitigating these disparities. Given the complexity of SDOH, comprehensive indices have been developed to study health disparities more meaningfully, including the Childhood Opportunity Index (COI), a multidimensional surveillance tool based on census-level data that provides a composite measure of physical and social conditions hypothesized to contribute to a range of pediatric health outcomes.[11] Neither the COI nor similar neighborhood indices have been used in any previous studies of medication adherence in JIA to date.

Estimates of medication adherence among children with JIA vary across studies,[12–16] which is likely related to the challenge of measuring adherence. Many methods rely on the patient or caregivers' report, which is known to overestimate adherence.[17–20] Objective methods, including electronic monitoring, pill counts, and drug level measurements, are often burdensome, inaccessible, and/or costly. Pharmacy dispensing data, which has been shown to more reliably identify nonadherent patients compared to self-reported adherence measures[21] and is accessible within some electronic health records (EHRs), offers an objective means of assessing medication adherence through metrics such as the medication possession ratio (MPR),[22] which measures the proportion of time a patient has access to a medication.

Although MPR has been used to assess medication adherence in adults with rheumatoid arthritis and youth with systemic lupus erythematosus,[23,24] nearly all studies assessing adherence in JIA have relied on the patient or caregivers' self-report, through surveys or questionnaires. The few studies that did measure MPR in patients with JIA were small[13] or obtained data from a pharmacy benefit management firm, which included only a small subset of patients with public insurance and lacked diagnostic codes and patient-level data.[16] Additionally, neither of these studies paired MPR with disease activity outcome measures. By linking medication adherence estimates to clinical data available through the EHR, we aimed to (1) identify patient and area-level characteristics associated with adherence to methotrexate (MTX) in JIA and (2) determine the association between adherence to MTX and disease activity in JIA. By focusing solely on MTX, the most commonly prescribed DMARD for JIA, this study also serves as an initial inquiry into the ease and practicality of using EHR-linked prescription fulfillment data to calculate MPR, with the intention of generalizing this methodology to other medications.

PATIENTS AND METHODS

Study design. This was a retrospective cohort study of patients with a physician diagnosis of JIA who were treated with MTX and observed at our pediatric rheumatology clinic. An exemption for secondary use of clinical data and waiver of informed consent was granted by the Children's Hospital of Philadelphia institutional review board (no. 23-020728).

Study population. All incident recipients of MTX, treated between January 2016 and September 2023 for ≥12 months, who had two or more outpatient visits during their first year of treatment, with one of the visits occurring between 10 and 14 months after the first MTX dispense date (Figure 1), were included in the sample (Figure 2). Patients without rheumatology clinic visits within 2 months of the first MTX order ("baseline visit") and 10 to 14 months after their first prescription fill date ("follow-up visit") were excluded, as the clinical outcomes data were extracted from these specific visits. Patients who filled prescriptions at pharmacies that did not participate in electronic prescription capture were also excluded.

Data source. Prescription fill data, accessible within our EHR, was provided by Surescripts, an information technology company that supports electronic prescriptions and provides these data to subscribing institutions. These data are available for all prescriptions filled at our center's retail and specialty pharmacies, and for at least 94% of all other pharmacies, as per Surescripts' 2021 national annual report.[25] We used Surescripts prescription fill data to calculate the MPR for each included patient. We excluded patients with no prescription fill data or with incomplete data (Supplementary Appendix S1 details the methodology used to identify these patients). We reviewed the pharmacy-reported days' supply for fills in which the reported supply was less than 28 days, not a multiple of 7 (because MTX is dosed once weekly among providers at our institution), or an unusual multiple of 7 for a pharmacy to dispense (35 days, 42 days, or 70 days) for a medication that is dosed once a week.

Figure 1. Study design timeline. *Visits from which disease activity data were obtained. †Patients without a visit between –2 and +2 months from the first MTX order were excluded. ‡Patients without a visit between the 10- and 14-month window were excluded. JIA, juvenile idiopathic arthritis; MPR, medication possession ratio; MTX, methotrexate.

Based on discussions with two clinical pharmacists and several retail pharmacists, we created a standard set of rules (Supplementary Appendix S2) for correcting days' supply entries that we deemed inaccurate.

We extracted clinical data, including patient-level covariates and JIA outcome measures, from our EHR-enabled JIA research registry from January 2016 to September 2023.

Study measures. We calculated the MPR for each patient by total days' supply over a fixed 365-day interval following the index MTX prescription fill date (total days supplied by dispensed prescriptions / 365 days × 100). Excess days' supply beyond the 365-day interval was truncated at 100% for perfect adherence.[26] We categorized MPR into two levels according to cutoff values used in other studies of adherence in patients with chronic disease.[27,28] Patients with adherence rates ≥80% were considered adherent, and patients with adherence rates <80% were considered nonadherent. The adherence category was used as the exposure for our regression models investigating the association between adherence and JIA disease activity.

The primary outcome measure was change in number of active joints from the baseline visit to the 12-month follow-up visit. This was chosen as the primary outcome given the known completeness of this variable in our institution's clinical documentation. We also investigated the percent change in joint count from baseline to follow-up as a secondary analysis. Secondary outcomes included absolute change in physician global assessment (PhGA) of disease activity score (0–10), patient or parent global assessment (PtGA) of overall well-being score (0–10), pain intensity score (0–10), and clinical Juvenile Arthritis Disease Activity Score (cJADAS-10; 0–30), a three-variable measure that is calculated from the active joint count, PhGA, and PtGA.

Covariates included sociodemographic characteristics (age at the index MTX fill, race or ethnicity [as reported in the medical record based on self-report during registration process], biologic sex, public vs commercial insurance, nationally normed COI 2.0 [a multidimensional measure of 29 neighborhood resources and conditions that impact children's healthy development based on census tract data]) and clinical factors (route of MTX prescribed, concomitant treatment with biologics, comorbid uveitis diagnosis, comorbid mental health diagnosis).

Statistical analysis. Comparisons of categorical risk factor variables between the nonadherent and adherent MTX groups were performed using Fisher's exact test. For continuous variables, comparisons were performed using the Wilcoxon rank-sum test. Association between adherence to MTX and JIA disease activity (continuous) outcome measures were determined through the use of linear regression, adjusting for the following covariates: baseline visit value of the specified outcome, age category (<6, 6–12, ≥13 years) at the index MTX prescription date, biologic sex, insurance type (public or commercial), nationally normed COI, route of MTX prescribed, concomitant treatment with biologics, comorbid uveitis diagnosis, and comorbid mental health diagnosis. We considered including race and ethnicity as a proxy for interpersonal forms of racism in addition to COI, which may capture consequences of structural racism.[29] However, including individual-level race and ethnicity (in addition to COI) had minimal effect on our results, so they were not retained in the final model. The main outcome (absolute change in active joint count) was looked at both overall and then stratified by subtype at presentation (oligoarticular and polyarticular respectively). Complete case analysis was conducted for all of the secondary outcome measures with missing data under the assumption that data were missing at random. All statistical analyses were performed using SAS, version 9.4.

RESULTS

A total of 224 patients (Figure 2) were included in the analysis. There were no statistically significant differences in demographic characteristics or JIA subtypes between patients included in the study and patients with JIA who were prescribed MTX during the same period but excluded from the analysis

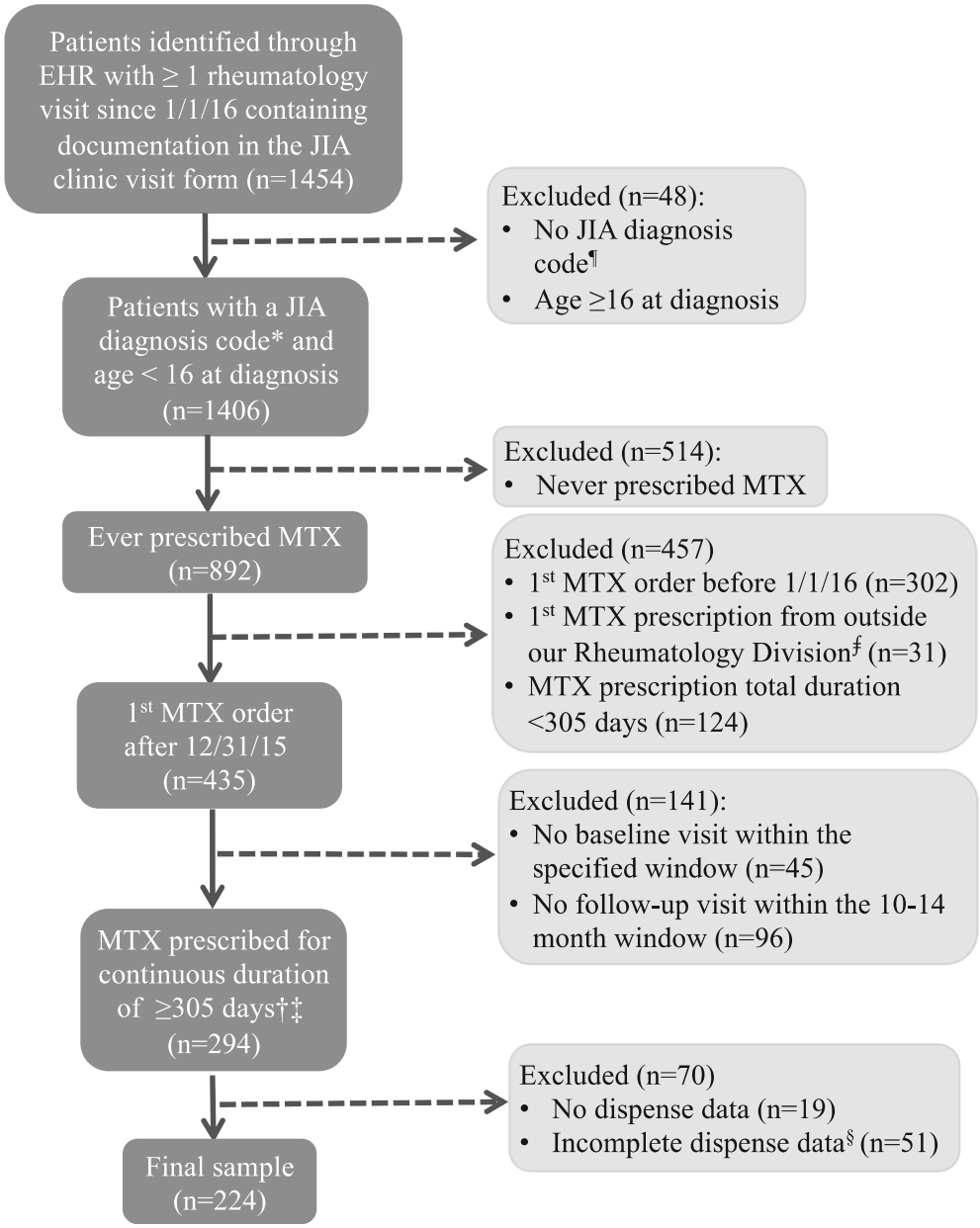


Figure 2. Flow diagram of patient selection into the study sample. *Any diagnosis code within M08, M45, or L40.5. †From prescription orders in the EHR. ‡Allowing for 60 days of flexibility to account for follow-up visits that occurred at the 10-month mark. ¶Patients with documentation within the JIA visit form but with diagnosis codes of inflammatory bowel disease–associated arthritis, reactive arthritis, transient effusion, or Lyme arthritis. ≠Outside institution or different specialty (gastroenterology, dermatology). §Determined by chart review of patients with a (1) >30-day lag time between a patient's first MTX order and the first dispense/claim, (2) gap in the middle of the dispense data >3 times the previous supply duration, or (3) gap at the end of the dispense data >3 times the previous supply duration. EHR, electronic health record; JIA, juvenile idiopathic arthritis; MTX, methotrexate.

(Supplementary Table 1). There was a mean of 10.2 dispenses per patient (SD 3.3) within the first year. The mean MPR was 83.7% (SD 19.2), and the majority (63.8%) of patients were classified as adherent.

Patient-level characteristics associated with adherence to MTX. Bivariate comparisons of baseline characteristics and risk factors between adherent and nonadherent groups are reported in Table 1. Patients of younger age at MTX start, of Black race, and living in areas with lower COI were more likely to be classified as nonadherent, and there was a greater proportion of patients with public insurance in the nonadherent group compared to the adherent group, although this was not statistically significant (P = 0.06). There was substantial overlap between Black race, public insurance, and very low COI (of the 18 patients living in areas with very low COI, 9 reported Black race

Table 1. Comparison of baseline characteristics and risk factors between adherent and nonadherent groups*

Variable	Mean MPR ± SD	Nonadherent (MPR <80%) (n = 81)	Adherent (MPR ≥80%) (n = 143)	*P* value[a]
Age at start of MTX, y				**0.002**
<6	78.2 ± 21.2	37 (45.7)	36 (25.2)	
6–12	84.7 ± 18.2	26 (32.1)	47 (32.9)	
≥13	87.9 ± 16.8	18 (22.2)	60 (42.0)	
Race and ethnicity				**0.022**
Asian	76.8 ± 20.7	5 (6.2)	4 (2.8)	
Black or African American	72.3 ± 25.2	10 (12.3)	5 (3.5)	
Hispanic or Latino/a	78.9 ± 23.2	12 (14.8)	12 (8.4)	
Multiracial/other[b]	84.7 ± 18.7	7 (8.6)	10 (7.0)	
White	85.6 ± 17.4	46 (56.8)	108 (75.5)	
Unknown/refused/missing	89.5 ± 20.2	1 (1.2)	4 (2.8)	
Biologic sex				0.88
Female	84.0 ± 19.0	57 (70.4)	99 (69.3)	
Male	82.9 ± 19.7	24 (29.6)	44 (30.7)	
Insurance type				0.06
Commercial	85.5 ± 17.3	59 (72.8)	120 (83.9)	
Public	76.4 ± 24.1	22 (27.2)	23 (16.1)	
National Child Opportunity Index				**0.001**
Very low	64.0 ± 24.6	15 (18.5)	3 (2.1)	
Low	86.2 ± 14.3	5 (6.2)	10 (7.0)	
Moderate	83.4 ± 20.7	10 (12.3)	16 (11.2)	
High	83.9 ± 19.3	20 (24.7)	41 (28.7)	
Very high	86.6 ± 16.4	31 (38.3)	73 (51.0)	
Baseline disease activity measures,[c] mean ± SD; n				
Clinical Juvenile Arthritis Disease Activity Score	–	10.1 ± 6.1; 60	10.9 ± 6.7; 110	0.50
Joint count	–	4.2 ± 5.3; 81	4.6 ± 6.1; 143	0.83
Physician Global Disease Activity Score	–	3.2 ± 1.7; 73	3.4 ± 1.9; 125	0.53
Patient Global Disease Activity Score	–	3.6 ± 2.7; 64	3.8 ± 2.7; 115	0.56
Pain intensity score	–	4.0 ± 2.8; 64	4.1 ± 2.8; 116	0.72
JIA subtype				0.06
Oligoarticular persistent	82.9 ± 19.7	24 (29.6)	51 (35.7)	
Oligoarticular extended	71.8 ± 16.9	7 (8.6)	2 (1.4)	
Polyarticular, RF-positive	85.0 ± 19.2	7 (8.6)	14 (9.8)	
Polyarticular, RF-negative	85.6 ± 17.1	19 (23.5)	31 (21.7)	
Enthesitis-related arthritis	86.2 ± 18.6	10 (12.3)	27 (18.9)	
Systemic	92.5 ± 11.6	2 (2.5)	4 (2.8)	
Psoriatic	81.8 ± 27.6	4 (4.9)	10 (7.0)	
Undifferentiated	76.8 ± 16.9	8 (9.9)	4 (2.8)	
Route of MTX administration				0.21
Subcutaneous	82.7 ± 18.8	53 (65.4)	81 (56.6)	
Oral	85.1 ± 19.7	28 (34.6)	62 (43.4)	
MTX + biologic/small molecule combination[d]				0.48
Yes	84.8 ± 18.1	48 (59.3)	92 (64.3)	
No (MTX monotherapy)	81.8 ± 20.8	33 (40.7)	51 (35.7)	
Uveitis diagnosis				0.13
Yes	74.0 ± 25.6	13 (16.0)	13 (9.1)	
No	84.9 ± 17.8	68 (84.0)	130 (90.9)	
Mental health diagnosis[e]				>.999
Yes	83.6 ± 16.8	9 (11.1)	15 (10.5)	
No	83.7 ± 19.5	72 (88.9)	128 (89.5)	

* Values listed as n (%) unless otherwise noted. Bold indicates significance. JIA, juvenile idiopathic arthritis; MPR, medication possession ratio; MTX, methotrexate, RF, rheumatoid factor.
[a] Fisher's exact test for categorical variables and Wilcoxon rank-sum test for continuous variables.
[b] n = 6 for multiracial category, n = 11 for other category.
[c] Missing data due to incomplete documentation of global disease activity scores and pain scores by both providers and patients/caregivers.
[d] Any of the following: adalimumab, etanercept, infliximab, golimumab, secukinumab, ixekinumab, leflunomide, abatacept, canakinumab, anakinra, tofacitinib, ruxolitinib, baricitinib, or upadacitinib.
[e] Any of the following conditions listed in the patient's diagnosis or problem list before MTX initiation: manic episode; bipolar disorder; depressive disorders (n = 1); mood affective disorders (n = 1); schizophrenia, schizotypal, delusional, and other nonmood psychotic disorders; anxiety disorders (n = 10); obsessive compulsive disorder (n = 1); reaction to severe stress and adjustment disorders (n = 3); dissociative and conversion disorders; somatoform disorders; attention-deficit hyperactivity disorder (n = 11); and conduct disorders (n = 1). Four patients had two mental health diagnoses.

Table 2. Association between adherence to MTX and JIA disease activity outcome measures*

Model	Outcome	β, Unadjusted (95% CI) (*P* value)	β, Adjusted[a] (95% CI) (*P* value)
A	Absolute change in joint count	−0.36 (−0.70 to −0.02) (**0.039**)	−0.38 (−0.74 to −0.01) (**0.043**)
B	Absolute change in joint count in patients with oligoarticular presentation	−0.07 (−0.29 to 0.14) (0.50)	0.02 (−0.20 to 0.25) (0.84)
C	Absolute change in joint count in patients with polyarticular presentation	−0.88 (−1.83 to 0.07) (0.07)	−1.18 (−2.23 to −0.13) (**0.028**)
D	Percent change in joint count	−8.7 (−23.0 to 5.6) (0.23)	−7.5 (−21.4 to 6.3) (0.29)
E	Absolute change in cJADAS-10 (n = 149)	−0.66 (−1.80 to 0.47) (0.25)	−0.72 (−1.88 to 0.44) (0.22)
F	Absolute change in PhGA (n = 187)	−0.23 (−0.50 to 0.05) (0.10)	−0.22 (−0.51 to 0.06) (0.13)
G	Absolute change in PtGA (n = 166)	−0.13 (−0.77 to 0.51) (0.70)	−0.14 (−0.79 to 0.52) (0.68)
H	Absolute change in pain intensity score (n = 167)	0.11 (−0.57 to 0.78) (0.76)	0.13 (−0.57 to 0.82) (0.72)

* Bold indicates significance. CI, confidence interval; cJADAS-10, clinical Juvenile Arthritis Disease Activity Score 10; JIA, juvenile idiopathic arthritis; MTX, methotrexate; PhGA, physician global assessment of disease activity score; PtGA, patient or parent global assessment of overall well-being score.
[a] All models were adjusted for the baseline visit value of the specified outcome, as well as the following covariates: age category (<6, 6–12, or ≥13 years) at start of MTX, biologic sex, insurance type (public or commercial), nationally normed Childhood Opportunity Index, route of MTX prescribed, concomitant treatment with biologics, comorbid uveitis diagnosis, and comorbid mental health diagnosis.

and 15 had public insurance). No statistically significant differences in adherence were observed based on disease activity at baseline, JIA subtype, route of MTX, concomitant treatment with biologics, or presence of uveitis or a mental health diagnosis.

Association between adherence to MTX and disease activity in JIA. For the whole cohort (n = 224), the unadjusted mean change in joint count from baseline to follow-up was –3.99 joints. Table 2 shows the adjusted estimates for associations between adherence to MTX and JIA disease activity outcomes measures. The adjusted number of active joints changed from baseline to 12-month follow-up by –0.38 more joints in the adherent group compared to the nonadherent group (95% confidence interval [CI] –0.74 to –0.01; $P = 0.043$). When stratified by polyarticular (≥5 joints at baseline) versus oligoarticular (<5 joints at baseline) presentation, there was an adjusted average absolute change from baseline to follow-up of –1.18 more joints for patients with polyarticular course in the adherent group compared to those in the nonadherent group (95% CI –2.23 to –0.13; $P = 0.028$). For our secondary outcomes (cJADAS-10, PhGA, PtGA, pain intensity), there were numerically greater decreases in cJADAS-10, PhGA, and PtGA from baseline to follow-up for patients in the adherent group compared to the nonadherent group, although these trends were not statistically significant, and we were limited by missing data (data not shown).

DISCUSSION

In this study designed to assess the associations between a clinically available measure of medication adherence, patient characteristics, and JIA disease activity, we found overall good adherence to MTX in our cohort, significant associations between nonadherence and SDOH, and significant differences between adherence group and change in active joint count.

In our study, the mean MPR was 83.7%, and 63.8% of patients classified as adherent. This is higher than a previous study that used pharmacy claims data from CVS Caremark to assess MTX adherence in children with rheumatic disease, which found mean MPRs of 47% and 58% for injectable MTX and oral MTX, respectively, with only 15% and 30% of patients in these respective groups having MPRs ≥80%.[15] However, our 63.8% adherence rate is lower than that reported in another study that found an overall 82% adherence rate using a self-report questionnaire to measure adherence to MTX across two centers.[14] Although these variations may be partially due to differences in patient characteristics across cohorts, each method for estimating adherence is subject to different biases. The study using self-reported measures likely resulted in an overestimation of adherence due to the biases associated with self-report,[19,20] and the study using pharmacy claims data may have underestimated adherence, as it included children with at least one claim for MTX prescribed by a rheumatologist in the CVS Caremark database from 2009 to 2010 but did not account for pharmacy dispenses that may have occurred outside the CVS Caremark database.

Despite overall good MTX adherence rates in our cohort, we identified statistically significant bivariate associations between both individual demographic and area-level factors (COI) and nonadherence (MPR <80%) (Table 1). Previous literature has demonstrated the importance of considering COI or other similar area-level determinants of health when examining disparities.[30–33] 33 As a composite measure incorporating 29 neighborhood

attributes that span three domains, COI measures structural determinants, which are different aspects of social disadvantage than race and ethnicity, insurance status, or other individual factors alone. Structural racism strongly influences where people live, what resources they have access to, and in turn their opportunity for upward mobility.[29] Mechanistically, if reliable transportation is less available, if neighborhood walkability and safety are poor, or if well-stocked pharmacies are farther from disadvantaged populations, then obtaining medication from a pharmacy is more challenging. It is also conceivable that living in neighborhoods with severe poverty, high crime, and poor community cohesion creates high stress and unstable conditions that may limit one's ability to adhere to medication regimens. It should be noted, however, that although COI and other similar area-based socioeconomic deprivation indices can be used to identify patients who may be at higher risk for nonadherence, these socioecological variables are not substitutes for measuring individual SDOH, as not everyone who resides in a neighborhood with low COI has the same individual risk factors for nonadherence.

Along with the association between COI and nonadherence, we also found significant differences between adherence group and change in active joint count, our primary outcome. As shown in Supplementary Table 2, we found that the very low COI group had a somewhat greater unadjusted absolute change in joint count (0.73 vs the adjusted 0.53) and was statistically significant when unadjusted (P = 0.047 vs P = 0.17). Larger studies are needed to formally assess whether differences in adherence (driven by logistical and behavioral factors)[34] mediate the known differences in clinical outcome disparities for patients who are Black, Medicaid-insured, or of a lower social economic position.[7,8] Although individual factors (medication beliefs, concerns, motivations, priorities, and comorbidities, among others) contribute to medication adherence, perhaps more impactful are the area-level factors (access to transportation, health care facilities, healthy food, quality education, safe water, insurance, among others), each requiring different approaches to effect change. Our research team is conducting qualitative work exploring patient perspectives on the mechanisms that influence medication adherence to better understand factors contributing to adherence differences, which will inform future adherence-focused interventions.

The significant association that we found between younger age at initiation of MTX and lower adherence was unexpected and differed from previous studies reporting better adherence in younger compared to older patients.[14,15] There are a few possible explanations for this unexpected finding. We used strict cutoffs (Supplementary Appendix S2) as recommended by pharmacists and nurses for discarding MTX vials after 28 days from first treatment, although for younger patients who require smaller doses, there is presumably more liquid remaining in the vial after the first four doses. We designed this study and developed rules for correcting presumed inaccurate pharmacy-reported days' supply with the best practice in mind, although we recognize that this may not represent a real-world setting, as some families may continue to use the same MTX vial beyond the recommended 28 days from time of initial puncture. It is also possible that younger children were more frequently sick with viral illnesses and were advised during these instances to hold MTX doses. The lack of association between comorbid mental health diagnoses and medication adherence is also inconsistent with what has been reported in the literature.[35,36] This may be due to several reasons, including the young age group studied, the difficulty assessing mental health retrospectively, and unstandardized documentation of mental health diagnoses in this cohort.

Our ability to link MTX dispense data to clinical data in the EHR illustrates a novel method for studying adherence to chronic medications. In settings where complete EHRs and pharmacy dispensing information exist, this offers an objective way of ascertaining adherence. Considering the inherent biases (social desirability and recall biases) that impact the reliability of self-reported adherence measures,[19,20] along with the challenges impacting implementation of many objective adherence instruments, identifying an objective and accessible method for measuring adherence is critical. Given that medication adherence has significant implications across a wide range of diseases, this methodology using Surescripts data to calculate the MPR as an adherence measure for children with JIA can be generalized to other conditions.

We acknowledge several study limitations. Although statistically significant, the clinical significance of the 0.38 greater decrease in active joint count among patients in the adherent compared to the nonadherent group is uncertain, as there is no reported minimal clinically important difference for this measure. However, we feel that this difference is worth noting, as any improvement in active joints can be meaningful to patients. By only including patients who have been on MTX for at least 12 months, it is likely that survival bias was at play, in which the sample favored patients who did not have significant MTX side effects and for whom it was efficacious. Our aim, however, was not to accurately measure adherence among our entire cohort of patients ever treated at our center, but rather to investigate how adherence to MTX based on pharmacy dispense data is associated with disease activity outcome measures and to assess for differences in adherence based on patient and area-level characteristics among patients for whom MTX is continuously prescribed and tolerated. Similarly, by nature of the study design, we excluded patients who were lost to follow-up, presumably the most nonadherent patients, as we prioritized obtaining outcomes data from the 12-month follow-up visit. It is therefore even more meaningful that even among the most adherent patients (those who return for follow-up), we still demonstrated associations with outcomes, as well as disparities related to individual and area-level characteristics. Despite excluding patients without

baseline or follow-up rheumatology visits within the specified two-month windows, we nonetheless encountered significant data missingness. Although cJADAS-10 is a more comprehensive measure of disease activity than joint count, the 33% of patients with missing cJADAS-10 limited our use of this as an outcome measure; we therefore opted for joint count (with 100% complete data) as our primary outcome, although we recognize that other variables beyond active joints alone contribute to overall disease activity in JIA.

There are also limitations with respect to pharmacy dispense data. First, although these data are available within many EHRs in high-income countries, we recognize that EHRs and electronic access to pharmacy dispensing information are not universal. Among settings with access to this information, data coming across the Surescripts interface may be incomplete in a way that we cannot quantify with absolute accuracy, especially across different drug formulations (inaccurate days' supply reported, inaccurate dispense unit reported such as two "each" referring to 2 mL vs two vials [Supplementary Appendix S2], or some pharmacies not reporting data, for example). We addressed this limitation to our best ability by performing chart review of cases in which there were gaps in the dispense data report over the specified 365-day interval and excluding patients with dispense data that appeared incomplete, by calling some pharmacies to clarify, and by manually correcting days' supply entries that we deemed inaccurate based on our standardized rules (Supplementary Appendix S2) created with input from pharmacists and nurses. It should also be noted that some missed doses are due to physician orders, such as when a patient is ill or undergoing surgery, which is not reflected in pharmacy dispense data. Finally, although pharmacy dispense data more reliably identify nonadherent patients compared to self-reported adherence measures,[21] they remain an indirect estimate of adherence, as we cannot know that a medication was truly taken.

In summary, the use of pharmacy dispense data to assess patients' adherence to medications offers unique advantages compared to other adherence measurements in that it is an objective, inexpensive tool that allows for large sample sizes and geographic coverage, without relying on self-report. We successfully linked aggregated pharmacy dispense data to clinical data available within the EHR, allowing us to investigate associations between adherence, patient characteristics, and disease activity measures. We also created rules to clean the data (Supplementary Appendix S2) so that this methodology can be implemented more easily in the future. This work can therefore be expanded to study other medications in JIA and even to medications for other chronic conditions. Given the complexity of factors impacting prescription filling and ultimate medication administration behaviors, a deeper understanding of the barriers and facilitators influencing medication adherence is needed to develop targeted interventions aimed at improving overall adherence and narrowing the disparity in outcomes for children with JIA. The methodology reported in this manuscript offers a clinically relevant tool as well as a robust way to assess changes in adherence in response to future interventions targeting adherence.

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 Abel 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/Helsinki Declaration requirements.

REFERENCES

1. Chhabra A, Robinson C, Houghton K, et al. Long-term outcomes and disease course of children with juvenile idiopathic arthritis in the ReACCh-Out cohort: a two-centre experience. Rheumatology (Oxford) 2020;59(12):3727–3730.
2. Ruperto N, Martini A. Current medical treatments for juvenile idiopathic arthritis. Front Pharmacol 2011;2:60.
3. Guzman J, Oen K, Tucker LB, et al; ReACCh-Out investigators. The outcomes of juvenile idiopathic arthritis in children managed with contemporary treatments: results from the ReACCh-Out cohort. Ann Rheum Dis 2015;74(10):1854–1860.
4. Lovell DJ, Reiff A, Jones OY, et al; Pediatric Rheumatology Collaborative Study Group. Long-term safety and efficacy of etanercept in children with polyarticular-course juvenile rheumatoid arthritis. Arthritis Rheum 2006;54(6):1987–1994.
5. Lovell DJ, Ruperto N, Goodman S, et al; Pediatric Rheumatology Collaborative Study Group; Pediatric Rheumatology International Trials Organisation. Adalimumab with or without methotrexate in juvenile rheumatoid arthritis. N Engl J Med 2008;359(8):810–820.
6. Guzman J, Oen K, Huber AM, et al; ReACCh-Out investigators. The risk and nature of flares in juvenile idiopathic arthritis: results from the ReACCh-Out cohort. Ann Rheum Dis 2016;75(6):1092–1098.
7. Chang JC, Xiao R, Burnham JM, et al. Longitudinal assessment of racial disparities in juvenile idiopathic arthritis disease activity in a treat-to-target intervention. Pediatr Rheumatol Online J 2020;18(1):88.
8. Akinsete AM, Woo JMP, Rubinstein TB. Disparities in pediatric rheumatic diseases. Rheum Dis Clin North Am 2022;48(1):183–198.
9. Snodgrass SR, Vedanarayanan VV, Parker CC, et al. Pediatric patients with undetectable anticonvulsant blood levels: comparison with compliant patients. J Child Neurol 2001;16(3):164–168.
10. Crocker D, Brown C, Moolenaar R, et al. Racial and ethnic disparities in asthma medication usage and health-care utilization: data from the National Asthma Survey. Chest 2009;136(4):1063–1071.
11. Acevedo-Garcia D, Noelke C, McArdle N, et al. Racial and ethnic inequities in children's neighborhoods: evidence from the new child opportunity index 2.0. Health Aff (Millwood) 2020;39(10):1693–1701.
12. Adriano LS, de França Fonteles MM, de Fátima Menezes Azevedo M, et al. Medication adherence in patients with juvenile idiopathic arthritis. Rev Bras Reumatol Engl Ed 2017;57(1):23–29.
13. Nieto-González JC, Trives-Folguera L, Melgarejo-Ortuño A, et al. Persistence and adherence to biologic therapies in juvenile idiopathic arthritis. Sci Rep 2021;11(1):16195.

14. Pelajo CF, Sgarlat CM, Lopez-Benitez JM, et al. Adherence to methotrexate in juvenile idiopathic arthritis. Rheumatol Int 2012;32(2):497–500.
15. Ringold S, Grant S, Girdish C, et al. Methotrexate and injectable tumor necrosis factor-α inhibitor adherence and persistence in children with rheumatic diseases. J Rheumatol 2013;40(1):80–86.
16. Lee WJ, Briars L, Lee TA, et al. Use of tumor necrosis factor-alpha inhibitors in children and young adults with juvenile idiopathic arthritis or rheumatoid arthritis. Pharmacotherapy 2016;36(12):1201–1209.
17. Waterhouse DM, Calzone KA, Mele C, et al. Adherence to oral tamoxifen: a comparison of patient self-report, pill counts, and microelectronic monitoring. J Clin Oncol 1993;11(6):1189–1197.
18. Daniels T, Goodacre L, Sutton C, et al. Accurate assessment of adherence: self-report and clinician report vs electronic monitoring of nebulizers. Chest 2011;140(2):425–432.
19. Osterberg L, Blaschke T. Adherence to medication. N Engl J Med 2005;353(5):487–497.
20. Rapoff MA, Duncan C, Karlson C. Assessing adherence and methodological considerations. In: Adherence to Pediatric Medical Regimens. Springer; 2023:115–159.
21. Mehat P, Atiquzzaman M, Esdaile JM, et al. Medication nonadherence in systemic lupus erythematosus: a systematic review. Arthritis Care Res (Hoboken) 2017;69(11):1706–1713.
22. Hess LM, Raebel MA, Conner DA, et al. Measurement of adherence in pharmacy administrative databases: a proposal for standard definitions and preferred measures. Ann Pharmacother 2006;40(7-8):1280–1288.
23. Grijalva CG, Chung CP, Arbogast PG, et al. Assessment of adherence to and persistence on disease-modifying antirheumatic drugs (DMARDs) in patients with rheumatoid arthritis. Med Care 2007;45-(10 Suppl 2):S66–S76.
24. Hsu CY, Lin YS, Cheng TT, et al. Adherence to hydroxychloroquine improves long-term survival of patients with systemic lupus erythematosus. Rheumatology (Oxford) 2018;57(10):1743–1751.
25. Surescripts. 2021 National Progress Report. February 2022. Accessed October 2, 2022. https://surescripts.com/docs/default-source/national-progress-reports/2021-national-progress-report.pdf?sfvrsn=71fcbe15_14
26. Sperber CM, Samarasinghe SR, Lomax GP. An upper and lower bound of the medication possession ratio. Patient Prefer Adherence 2017;11:1469–1478.
27. Karve S, Cleves MA, Helm M, et al. Good and poor adherence: optimal cut-point for adherence measures using administrative claims data. Curr Med Res Opin 2009;25(9):2303–2310.
28. Kang JH, Choi SE, Xu H, et al. Non-adherence to methotrexate was associated with high disease activity and poor health-related outcomes during a 4-year follow-up of rheumatoid arthritis patients. Clin Exp Rheumatol 2022;40(9):1744–1753.
29. Blatt LR, Sadler RC, Jones EJ, et al. Historical structural racism in the built environment and contemporary children's opportunities. Pediatrics 2024;153(2):e2023063230.
30. Hu J, Kind AJH, Nerenz D. Area deprivation index predicts readmission risk at an urban teaching hospital. Am J Med Qual 2018;33(5):493–501.
31. Kind AJH, Jencks S, Brock J, et al. Neighborhood socioeconomic disadvantage and 30-day rehospitalizations: a retrospective cohort study. Ann Intern Med 2014;161(11):765–774. https://doi.org/10.7326/M13-2946
32. Trinidad S, Brokamp C, Mor Huertas A, et al. Use Of area-based socioeconomic deprivation indices: a scoping review and qualitative analysis. Health Aff (Millwood) 2022;41(12):1804–1811.
33. Bartels CM, Rosenthal A, Wang X, et al. Investigating lupus retention in care to inform interventions for disparities reduction: an observational cohort study. Arthritis Res Ther 2020;22(1):35.
34. Peh KQE, Kwan YH, Goh H, et al. An adaptable framework for factors contributing to medication adherence: results from a systematic review of 102 conceptual frameworks. J Gen Intern Med 2021;36(9):2784–2795.
35. Chang JC, Davis AM, Klein-Gitelman MS, et al. Impact of psychiatric diagnosis and treatment on medication adherence in youth with systemic lupus erythematosus. Arthritis Care Res (Hoboken) 2021;73(1):30–38.
36. Rapoff MA, Duncan C, Karlson C. Consequences of nonadherence and correlates of adherence. In: Adherence to Pediatric Medical Regimens. Springer; 2023:25–51.

Arthritis Care & Research
Vol. 77, No. 3, March 2025, pp 309–321
DOI 10.1002/acr.25439

American College of Rheumatology
Empowering Rheumatology Professionals

Exploring Pain Adaptation in Youth With Juvenile Idiopathic Arthritis: Identifying Youth and Parent Resilience Resources and Mechanisms

Yvonne N. Brandelli,[1] Sean P. Mackinnon,[2] Christine T. Chambers,[1] Jennifer A. Parker,[3] Adam M. Huber,[1] Jennifer N. Stinson,[4] Shannon A. Johnson,[2] and Jennifer P. Wilson[5]

Objective. Although juvenile idiopathic arthritis (JIA) is often associated with pain, this experience does not necessitate negative outcomes (eg, depression, functional impairment). Little research has explored youth and parent resilience resources (ie, stable traits) and mechanisms (ie, dynamic processes) in this context, and studies have focused on their contributions independently rather than collectively. This study, informed by the Ecological Resilience-Risk Model in Pediatric Chronic Pain, sought to (1) explore the relationships among youth and parent resilience resources and mechanisms and (2) identify the relative importance (RI; ie, independent contributions when entered simultaneously) of evidence-based youth and parent resources and mechanisms in contributing to youth-reported recovery, sustainability, and growth outcomes.

Methods. Youth (13–18 years) with JIA and their parents (156 dyads) completed a battery of online questionnaires assessing resilience resources (optimism, resilience), mechanisms (psychological flexibility, pain acceptance, self-efficacy), recovery and sustainability (pain intensity, functional disability, health-related quality of life), and growth (benefit finding) outcomes.

Results. Analyses demonstrated significant positive correlations across within-person resources and mechanisms and weaker correlations across within-dyad resources and mechanisms. Although the RI of predictors varied by outcome, youth pain acceptance was the most robust predictor across models (RI = 0.03–0.15). Some predictors (eg, parent psychological flexibility and pain acceptance) were generally categorized as "Not Important," whereas others (eg, youth resilience) had "Inconclusive" results, suggesting construct overlap.

Conclusion. Although additional research is needed to further understand resilience, results highlight the importance of fostering pain acceptance in youth and incorporating parents in psychosocial interventions to optimize living with JIA.

INTRODUCTION

The hallmark experience of juvenile idiopathic arthritis (JIA) is pain,[1] which has been identified as a top research priority for families.[2] To date, research has focused on negative outcomes associated with JIA pain (eg, internalizing symptoms, lower health-related quality of life [HRQoL], impaired social functioning)[3]; however, the experience of pain and the presence of risk factors do not guarantee that youth with JIA will endure these negative outcomes. There is individual variation in pain experiences,[4] which is likely due to the presence of promotive and/or protective factors (ie, factors that have a positive influence on outcomes regardless of risk factors, and factors that dampen risk factors, respectively).[5]

The funders did not play a role in the study design, collection, analysis, interpretation, or reporting of these data.

Salary support was provided to Ms. Brandelli by a Scotia Scholars award from Research Nova Scotia), a Nova Scotia Research and Innovation Graduate Scholarship, the Dr. Mabel E. Goudge Award from the Department of Psychology and Neuroscience at Dalhousie University, and a Graduate Studentship from IWK Health (to Ms Brandelli). Dr Chambers' work was supported by a Canada Research Chair, the Canadian Institutes of Health Research (FRN 167902), and the Canada Foundation for Innovation.

[1]Yvonne N. Brandelli, BA (Hons), Christine T. Chambers, PhD, Adam M. Huber, MD, MSc: Dalhousie University and IWK Health, Halifax, Nova Scotia, Canada; [2]Sean P. Mackinnon, PhD, Shannon A. Johnson, PhD: Dalhousie University, Halifax, Nova Scotia, Canada; [3]Jennifer A. Parker, PhD: IWK Health, Halifax, Nova Scotia, Canada; [4]Jennifer N. Stinson, PhD, RN-EC, CPNC: University of Toronto and Hospital for Sick Children, Toronto, Ontario, Canada; [5]Jennifer P. Wilson, BA: Cassie and Friends: A Society for Children with Juvenile Arthritis and other Rheumatic Diseases, Vancouver, British Columbia, Canada.

Drs Mackinnon and Chambers contributed equally to this work.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25439).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25439.

Address correspondence via email to Yvonne N. Brandelli, BA (Hons), at Yvonne.Brandelli@dal.ca.

Submitted for publication April 5, 2024; accepted in revised form September 11, 2024.

SIGNIFICANCE & INNOVATIONS

- This study assessed the relative importance of youth and parent resilience resources and mechanisms to advance knowledge as it pertains to juvenile idiopathic arthritis (JIA) pain.
- Most youth resilience resources and mechanisms were significantly related to one another, as were most parent resilience resources and mechanisms. Relationships between youth and parent resilience resources and mechanisms were less likely to be significant.
- Across surrogate markers of pain adaptation, youth pain acceptance was one of the most robust predictors. Parent contributions, such as optimism and psychosocial self-efficacy, also played an important role.
- To promote resilience in the context of JIA, results highlight the importance of fostering youth pain acceptance and incorporating parents in the psychosocial interventions provided.

The study of promotive and protective factors is encompassed within the resilience literature. Although it is a complex, systemic, and dynamic process without a universal definition, resilience can be conceptualized as the capacity of a dynamic system to adapt successfully to disturbances (such as a diagnosis of JIA) that threaten system function, viability, or development.[6] There is a growing need within the JIA and pediatric pain literature to further this field of study,[7] particularly because the aim of many treatments is to manage pain and prevent irreversible damage rather than "cure" the disease.[8] Thus, by shifting emphasis to understanding and promoting the conditions necessary for resilience, youth can be protected from unfavorable outcomes and learn to optimize living in the face of adversity.

Although there is no unified outcome of resilience, Sturgeon and Zautra[9] theorized that pain adaptation can be measured in terms of one's recovery (ie, resumed functioning; psychological, physical, or academic well-being), sustainability (ie, perseverance with valued activities), and growth (ie, new learning or a better understanding of one's capabilities). These are nevertheless surrogate markers of adaptation[10] because it is a process that depends on the individual and their context, and it is unclear which of the many outcomes are necessary or sufficient to determine successful adaptation. Cousins et al[7] tailored this model for pediatric populations, placing greater emphasis on the ecological system. Specifically, these outcomes are the result of an interaction between resilience resources and risk factors (ie, stable traits) and resilience and risk mechanisms (ie, dynamic processes) that occur within and between the individual, their family and/or social environment, and their culture and time.

There is preliminary support for components of the Ecological Resilience-Risk Model in Pediatric Chronic Pain in the broader literature. In terms of resources, trait optimism (ie, having favorable expectations for the future) predicts improved HRQoL directly in youth with abdominal pain[11] and through reduced fear and catastrophizing in youth with chronic pain.[12] Trait resilience (ie, a general disposition of bouncing back) is associated with reduced disease severity, pain, and disability; and greater HRQoL.[13] There is also preliminary support for other resources, including mindfulness,[14] positive affect,[15] and positive peer relationships.[16]

By way of mechanisms, psychological flexibility, or the ability to be present focused and engaged in values-based action, is associated with less daily activity avoidance in youth with chronic pain,[15,17] and in parents, it is positively associated with youth HRQoL in some[18] but not all[14] studies. Support also exists for pain acceptance. In pediatric pain rehabilitation programs, increases in acceptance are predictive of decreased depressive symptoms, catastrophizing, and functional disability.[19] More broadly, youth pain acceptance is positively associated with HRQoL[14,18] and negatively associated with pain intensity,[18] and parent pain acceptance is indirectly associated with decreases in pain interference and increases in mobility through youth pain acceptance.[20] Finally, self-efficacy, or one's belief in their ability to function effectively in the presence of pain or disease, also contributes to pain acceptance,[18] psychological flexibility,[18] HRQoL,[18] reduced pain intensity,[18] reduced disability,[21] and fewer depressive symptoms.[21]

Despite this literature, these constructs have only been minimally applied to the context of JIA. Hynes et al[22] systematically reviewed the risk and resilience resources and mechanisms in the JIA literature. Briefly, they found that across different outcomes of pain adaptation, family functioning[23] and child-perceived social support[24] are relevant resources; and child self-efficacy,[24–27] psychological flexibility,[28,29] pain acceptance,[28,29] and parent-reported child pain coping (specifically problem-solving)[26] are relevant mechanisms. See the study by Hynes et al[22] for a comprehensive review and summative figure.

Given this scant literature, numerous variables remain to be explored (eg, parent optimism, trait resilience).[7,22] Moreover, much of the literature has used small samples, relied on proxy reports, and emphasized outcomes of HRQoL.[22] Studies have largely focused on resources and mechanisms independently, neglecting to explore their relationships with one another and the broader sociocultural environment.[22,30] As such, there is a need to identify the resilience resources and mechanisms relevant to this population in a holistic manner to better understand what to emphasize to optimize living with JIA.[7] Based on the aforementioned JIA[3,22] and pain literatures, the following youth and parent resources and mechanisms were identified as potentially relevant: optimism, trait resilience, psychological flexibility, pain acceptance, and self-efficacy.

The aims of this study were to (1) explore the relevance of, and relationships between, youth and parent resilience resources and mechanisms that have been identified in the broader literature

(ie, optimism, trait resilience, psychological flexibility, pain acceptance, and self-efficacy) in the context of JIA pain and (2) explore their relative importance (RI; ie, their independent contributions while simultaneously accounting for other resources and mechanisms) in contributing to youth-reported recovery and sustainability (ie, pain intensity, functioning, HRQoL) and growth (ie, benefit finding) outcomes. A priori hypotheses were that (1) there would be significant positive relationships among the resilience resources and mechanisms and (2) resilience resources and mechanisms would predict positive adaptation in the presence of JIA pain; however, no a priori predictions were made regarding which constructs would emerge as most important in the analyses.

PATIENTS AND METHODS

Study design. The data used for the current study were part of a larger data set. Another study with a different research question, variables, and analyses has been submitted for publication elsewhere. Data and syntax for the present study are openly available through Open Science Framework (https://doi.org/10.17605/OSF.IO/8G29D). This cross-sectional internet-based study was approved by the IWK Research Ethics Board (#1026950) and complies with the Declaration of Helsinki.

Following best practice in patient engagement, a leader in the field codeveloped the patient partnership plan for this study. In addition to partnering with Cassie and Friends, a parent-led organization for families of children with rheumatic diseases (www.cassieandfriends.ca), two parents and one youth with JIA provided consultation, support, and feedback on this study from conceptualization through to dissemination. Partners were compensated following the Solutions for Kids in Pain guidelines (https://kidsinpain.ca/wp-content/uploads/2021/03/SKIP-Patient-Partner-Compensation-Guidelines-approved-Feb-10-2020-1.pdf).

Participants. Youth (13–18 years old) with a diagnosis of JIA and a parent or caregiver were recruited through online and social media platforms (eg, arthritis and pain communities, Facebook advertisements, blog posts), previous studies, posters at rheumatology and pain clinics, the IWK Health research registry, and industry partnerships. Recruitment took place between November 2021 and April 2023.

Of the 206 youth and parent dyads who consented online, 33 were ineligible given their diagnosis or age, and 17 stopped after providing consent. The final sample size was 156 unique dyads. Missing data were complex. Parents generally filled out the entire survey (n = 129, 82.7%), with a small number providing partial data or not completing the survey at all (n = 11 [7.0%] and n = 16 [10.3%], respectively). Most youth also filled out the entire survey (n = 122, 78.2%), with a minority providing partial or no data (n = 7 [4.5%] and n = 27 [17.3%], respectively). All data, including partial data, were analyzed when possible. A sensitivity analysis was conducted in G*power using the final sample of 156 dyads, an α of 0.05, and power of 0.80. With 12 predictors, there is sufficient power to detect an overall R^2 of 0.12. When considering power for individual predictors, there is sufficient power to detect an f^2 of 0.051 or $\Delta R^2 = 0.05$.

Measures and procedures. Participants self-selected into this study. After completing an eligibility screening questionnaire, youth and parents were emailed unique survey links that contained a consent form and a 45-minute battery of validated questionnaires through Qualtrics. Questions probing background information measured demographic (eg, age, sex, ethnicity via fixed categories and open-ended responses[31]) and medical variables (eg, diagnosis, disease activity). Resilience resources were assessed via measures of optimism[32,33] and trait resilience in youth.[34] Resilience mechanisms were assessed via measures of psychological flexibility,[35,36] pain acceptance,[37,38] and arthritis-specific self-efficacy.[25,27] Pain adaptation was assessed via the following youth-reported recovery, sustainability, and growth outcomes: usual pain intensity,[39] functioning,[40] generic and rheumatology-specific HRQoL,[41] and benefit finding.[42] Table 1 outlines the list of measures, including their definitions, scaling, reporter, and psychometric properties. Items were averaged to create total scores, with higher scores reflecting greater endorsement of the construct. Responses were mandatory; however, participants could select "prefer not to answer" (treated as missing data). On completion, participants received a $15 (Canadian) online gift card, and dyads were entered into a draw to win one of two pairs of $250 (Canadian) gift cards.

Analyses. To ensure data validity, in addition to screening participants during data collection (eg, screening questionnaire, passwords, the prevention of multiple submissions from the same internet protocol address),[43] data were also screened before analyses (eg, review of attention checks, "spam trap" questions, captchas).

Analyses were completed using the psych() and lavaan() packages in R (https://www.r-project.org/). Youth–parent dyads were paired, and total scores were calculated. Assumptions of normality were met. A full information maximum likelihood approach was used for missing data.

Descriptive statistics and bivariate correlations (aim 1) were used to describe measures of pain, resilience resources and mechanisms, and outcomes. To address aim 2, a series of five multiple regressions were tested through structural equation models using the 12 resources and mechanisms as predictors. RI was calculated with the Pratt index,[44] the product of the bivariate correlation and standardized regression coefficients. This method partitions the total R^2 across all variables to quantify the RI of each predictor variable in a way that sums to the total R^2 value (eg, if the total R^2 is 0.10 and a single RI value is 0.05, then

Table 1. Description, scoring, and psychometric support for study measures*

Domain	Concept/definition	Scale	Reporter	Scoring and psychometric support
Resilience resources	Dispositional optimism: a set of positive expectations regarding the future	Youth Life Orientation Test[32]	Youth	Twelve items were rated on a scale from 0 ("not true for me") to 3 ("true for me"). Six items measuring pessimistic expectations were reverse coded. This scale has good internal consistency, test–retest reliability, and convergent validity.[32]
		Life Orientation Test Revised[33]	Parent	Six items (and four fillers) were rated on a scale from 0 ("strongly disagree") to 4 ("strongly agree"). Half of the items were reverse scored. Adequate predictive and discriminant validity and good internal consistency have been observed.[33]
	Trait resilience: the ability to bounce back or recover from stress	Brief Resilience Scale[34]	Youth	Six items assessing resilience were rated on a scale from 1 ("strongly disagree") to 5 ("strongly agree"). Select items were reverse scored. Although not specifically developed for adolescents, it has been used in many adolescent populations and has demonstrated acceptable internal consistency and appropriate concurrent, discriminant, and criterion validity.[50]
Resilience mechanisms	Psychological flexibility: the capacity to stay present focused and engaged in values-based action	Avoidance and Fusion Questionnaire for Youth[35]	Youth	Eight items were rated on a scale from 0 ("not at all true") to 4 ("very true"). They were averaged to create a total score of psychological inflexibility. This scale has demonstrated excellent internal consistency and good convergent and construct validity.[35]
		Parental Psychological Flexibility Questionnaire[36]	Parent	Ten items were rated on a scale from 0 ("never true") to 6 ("always true"). Select items were reverse scored. This scale has good internal consistency and construct validity.[36]
	Pain acceptance: a willingness to permit pain to be present without trying to stop or reduce it (pain willingness) and a willingness to persist with important activities (activity engagement)	Chronic Pain Acceptance Questionnaire for Adolescents[37]	Youth	Eight items assessing pain willingness (reverse scored) and activity engagement were rated on a scale from 0 ("never true") to 4 ("always true") and averaged to create a total score. Good internal consistency, validity, and sensitivity to treatment have been seen with this abbreviated scale.[37]
		Parent Pain Acceptance Questionnaire[38]	Parent	Fifteen items assessing acceptance of pain-related thoughts and feelings (reverse scored) and activity engagement were rated from 0 ("never true") to 4 ("always true") and averaged to create a total score. This scale has strong internal consistency and construct validity.[38]
	Arthritis self-efficacy: the perceived ability to perform the courses of action that produce desired attainments in various domains of life with JIA (ie, activities, symptoms, and emotions in youth and symptoms and psychosocial domains in parents)	Children's Arthritis Self-Efficacy Scale[25]	Youth	Four items assessing activity, three items assessing symptom, and three items assessing emotion self-efficacy were rated on a scale from 1 ("not at all sure") to 5 ("very sure"). The scale has good to excellent internal consistencies and strong concurrent and construct validity.[25]
		Parent's Arthritis Self-Efficacy Scale[27]	Parent	Seven items assessing parent's belief in their ability to manage their child's symptoms and seven items assessing their certainty in managing psychosocial components were rated from 0 ("very uncertain") to 7 ("very certain"). This was not 1 to 7 because of an administration error.

(Continued)

Table 1. (*Cont'd*)

Domain	Concept/definition	Scale	Reporter	Scoring and psychometric support
				Excellent internal consistency and good concurrent and construct validity were seen in the validation study.[27]
Recovery and sustainability outcomes	Pain intensity: a sensory component of the pain experience assessing one's perceived intensity of their usual pain	11-point numeric rating scale	Youth	Youth rated their current and usual pain intensity on a scale from 0 ("no hurt") to 10 ("the worst hurt you could ever imagine"). A recent review demonstrated sufficient reliability and criterion validity for this scale and strongly recommended its use among adolescent populations.[39]
	Functional disability: an assessment of the impact of a disease on one's daily functioning	Functional Disability Inventory[40]	Youth	Fifteen items assessing difficulty engaging in activities were rated by youth on a scale from 0 ("no trouble") to 4 ("impossible"). In the validation study, there was strong support for the internal consistency and the construct, concurrent, and predictive validity of this measure.[40]
	HRQoL: a multidimensional assessment of one's physical, mental, and social functioning (generic core total score) and rheumatology-specific physical and psychosocial health (rheumatology total score)	The Pediatric Quality of Life Inventory 4.0 generic core scales and 3.0 rheumatology module[41]	Youth	Youth rated 23 items pertaining to their general HRQoL and 22 items pertaining to their rheumatology-specific HRQoL from 0 ("never") to 4 ("almost always"). Items were reverse scored, transformed to a 0–100 scale, and averaged to create two total scores. Good to excellent internal consistency has been seen for total scores, and both construct validity and responsiveness have been demonstrated.[41]
Growth outcomes	Benefit finding: an acknowledgment of positive changes or benefits in the presence of an event, such as an illness or trauma	Benefit Finding Scale for Children[42]	Youth	Ten items were rated from 0 ("not true for me") to 4 ("very true for me"). Instructions were modified to "having had my arthritis…" as opposed to "illness." This scale has excellent internal consistency and some evidence of construct validity.[42]

* HRQoL, health-related quality of life; JIA, juvenile idiopathic arthritis.

that predictor accounts for 5% of the variance in the outcome and 50% of the total R^2 value). Standardized correlation coefficients (r) and their *P* values, regression coefficients (β) and their *P* values and confidence intervals, the total variance predicted by each model, and the Pratt RI index are reported. Because our study has 80% power to detect ΔR^2 of 0.05 or larger, and because the effect size observed in much of the psychological literature[45] is 0.21 or $R^2 = 0.044$, RI values at or above 0.05 will be considered important. Given the large number of coefficients (r, β, RI), each of the 12 predictors will be classified into one of four categories for ease of exposition. Predictors coded as "Important" will have an RI ≥ 0.05, and each coefficient will be statistically significant. Predictors coded as "Potentially Important" will have an RI ≥ 0.05 and one statistically significant coefficient. Predictors coded as "Not Important" will have an RI between –0.05 and 0.05 and no statistically significant coefficients. Predictors coded as "Inconclusive" will incorporate all other cases (ie, RI < 0.05 with varying patterns of statistical significance). Definitive conclusions regarding "Inconclusive" predictors cannot be made because discrepancies may reflect the complexity of these constructs or a lack of statistical power.

Note that this coding may oversimplify the results in exchange for ease of interpretation. In comparing the r coefficient (the independent relationship between the predictor and outcome) and the β coefficient (the relationship of each predictor and the outcome while accounting for other predictors), more

Table 2. Descriptive and medical variables*

Demographics and medical variables	Parent (n = 140)	Youth (n = 129)
Participant demographics		
Age in years, mean ± SD (min, max)	45.24 ± 4.87 (33, 57)[a]	15.29 ± 1.62 (13, 18)[b]
Female sex, n (%)[b]	–	110 (70.5)
Gender, n (%)[b]		
Mother/girl	148 (94.9)	106 (67.9)
Father/boy	8 (5.1)	46 (29.5)
Other (transgender, nonbinary, gender fluid)	–	4 (2.5)
Race and ethnicity, n (%)[c]		
White	123 (78.8)	110 (70.5)
Aboriginal	7 (4.5)	9 (5.8)
South Asian	4 (2.6)	4 (2.6)
Black	3 (1.9)	3 (1.9)
East/Southeast Asian	3 (1.8)	3 (1.9)
Other (Jewish, West Asian, Latin American)	5 (3.2)	4 (2.6)
Prefer not to answer	2 (1.3)	1 (0.6)
Country of residence, n (%)		
Canada	96 (68.6)	95 (74.2)
United Kingdom	24 (17.1)	19 (14.8)
United States	16 (11.4)	11 (8.6)
Other (Ireland, South Africa, Australia)	4 (2.8)	3 (2.4)
Income (CAD$), n (%)		
<$50,000	16 (11.5)	–
$50,000–$99,999	42 (30.0)	–
$100,000–$149,999	29 (20.7)	–
>$150,000	36 (25.7)	–
Prefer not to answer	17 (12.1)	–
Youth medical characteristics		
Diagnosis, n (%)[b,d]		
Oligoarticular arthritis	–	26 (16.8)
Polyarticular arthritis	–	37 (23.9)
Enthesitis-related arthritis	–	27 (17.4)
Psoriatic arthritis	–	10 (6.5)
Systemic arthritis	–	17 (11.0)
Undifferentiated or unknown[e]	–	38 (24.5)
Age at diagnosis in years, mean ± SD (min, max)[b,f]		8.09 ± 4.72 (0, 16)
Current disease activity (active/flare), n (%)	87 (62.1)	75 (59.1)[g]

* Parent-reported data were used for medical characteristics. Percentage was calculated based on the number of participants who completed the question rather than the total N. CAD$, Canadian dollars.
[a] n = 139.
[b] Data from parents/youth were combined to achieve N = 156. If parent/youth data did not match, youth data were used before parent data for demographic information, and parent data were used before youth data for medical information.
[c] Participants could select more than one response.
[d] n = 155.
[e] Three participants indicated also having a diagnosis of autoimmune arthritis, dermatomyositis, and scleroderma.
[f] n = 152.
[g] n = 127.

complex patterns might be observed. These include (1) the r coefficient is significant but the β coefficient is not (suggestive of construct overlap), (2) the r coefficient is insignificant but the β coefficient is significant (suggesting that outcome-irrelevant variance has been removed by the other included predictors), or (3) a predictor has a positive r coefficient but a negative β coefficient (or vice versa), akin to a suppressor variable,[46] which may enhance the predictive ability of other predictors in the model by accounting for some of their outcome-irrelevant variance. Such results will be described in the text to reflect these nuances. Covariates (youth sex and age) were explored in Supplementary Figures S1 and S2 and Supplementary Table S1 on reviewer request (analyses in the Supplementary Material are not reproducible because demographic information was redacted from the open data set to preserve confidentiality).

RESULTS

Descriptive statistics. Participants were 156 dyads, including 129 adolescents with JIA and 140 parents. Adolescents were generally female (67.9%) and had a mean age of 15.29 years (SD 1.62 years). Similarly, 95% of parents were mothers and had a mean age of 45.24 years (SD 4.87 years). Youth were on average diagnosed at age 8.09 years (SD 4.72 years), most of whom had been diagnosed with polyarticular (23.9%) or enthesitis-related (17.4%) arthritis. More than half the sample was currently experiencing active disease by both youth and parent report. Internal consistencies of measures ranged from adequate to excellent (α = 0.76–0.95). See Tables 2 and 3 for demographics and study variables.

Associations between youth and parent resilience resources, mechanisms, and outcomes. Pearson's correlation coefficients are presented in Figure 1. As hypothesized, correlations between the 12 resilience resources and mechanisms were generally positive (except for youth psychological inflexibility, an inverse score) and significant. Correlations between within-youth resources and mechanisms were all significant (except for symptom self-efficacy and pain acceptance), ranging from weak (r = –0.28, $P < 0.05$) to strong ($r = 0.71$, $P < 0.001$). Correlations between within-parent resources and mechanisms were also positive and significant (except for psychological flexibility and symptom self-efficacy), ranging from weak ($r = 0.20$, $P < 0.05$) to strong ($r = 0.82$, $P < 0.001$). Nonsignificant results were more likely to occur across dyad members.

A similar pattern was observed between resources/mechanisms and outcomes, wherein every youth predictor was significantly related to the outcomes of usual pain, functional disability, and HRQoL (benefit finding was not significantly related to pain acceptance or activity self-efficacy); however, parent predictors were less strongly, if at all, related.

Relative importance of resilience resources and mechanisms in predicting outcomes. Results of the five multiple regression analyses with the relative contributions of the 12 resilience resources and mechanisms across recovery,

Table 3. Questionnaire data*

Questionnaire and theoretical range of scores	n	Mean	SD	Min, max	α
Optimism (Y): YLOT; 0–3	124	1.73	0.72	0.00, 3.00	0.92
Optimism (P): LOT-R; 0–4	131	2.39	0.71	0.83, 4.00	0.83
Resilience (Y): BRS; 1–5	122	3.16	0.78	1.00, 4.83	0.84
Psychological flexibility (Y): AFQ-Y8; 0–4	125	1.29	0.87	0.00, 4.00	0.87
Psychological flexibility (P): PPFQ-10; 0–6	137	4.02	0.94	1.30, 5.80	0.86
Pain acceptance (Y): CPAQ-A8; 0–4	124	2.48	0.67	0.88, 4.00	0.76
Pain acceptance (P): PPAQ; 0–4	137	2.29	0.66	0.00, 3.60	0.83
Self-efficacy (Y): CASE; 1–5					
Activity	120	2.89	1.14	1.00, 5.00	0.90
Symptom	119	2.76	1.00	1.00, 5.00	0.85
Emotion	120	3.05	1.22	1.00, 5.00	0.87
Self-efficacy (P): PASE; 1–7					
Symptom	125	3.05	1.42	0.00, 6.57	0.89
Psychosocial	126	4.43	1.51	0.40, 7.00	0.93
Pain intensity (Y): NRS-11; 0–10	125	4.96	2.23	0.00, 10.00	–
Functional disability (Y): FDI; 0–4	126	1.20	0.82	0.00, 3.40	0.94
HRQoL (Y): PedsQL generic 4.0; 0–100	122	60.01	21.21	16.30, 100.00	0.95
HRQoL (Y): PedsQL rheumatology 3.0; 0–100	122	63.06	20.39	9.09, 100.00	0.94
Benefit finding (Y): BFSC; 0–4	120	2.28	0.80	0.00, 4.00	0.89

* AFQ-Y, Avoidance and Fusion Questionnaire for Youth; BFSC, Benefit Finding Scale for Children; BRS, Brief Resilience Scale; CASE, Children's Arthritis Self-Efficacy Scale; CPAQ-A, Chronic Pain Acceptance Questionnaire for Adolescents; FDI, Functional Disability Inventory; HRQoL, health-related quality of life; LOT-R, Life Orientation Test Revised; NRS-11, 11-point numeric rating scale; P, parent; PASE, Parent's Arthritis Self-Efficacy Scale; PedsQL, Pediatric Quality of Life Inventory; PPAQ, Parent Pain Acceptance Questionnaire; PPFQ-10, Parental Psychological Flexibility Questionnaire; Y, youth; YLOT, Youth Life Orientation Test.

Figure 1. Bivariate correlations for study variables. Pearson's correlation coefficient is significant at *** $P = 0.001$; ** $P = 0.01$; * $P = 0.05$ level (two tailed). HRQoL, health-related quality of life; P, parent; Y, youth.

sustainability, and growth outcomes can be seen in Table 4. A summary of the findings and their coding is depicted in Figure 2.

In Model 1, 29% of the variance in usual pain intensity was accounted for by the included predictors. Youth pain acceptance (RI = 0.12) was the most robust contributor to reduced pain intensity. Potentially Important contributors included parent psychosocial self-efficacy (RI = 0.07) and parent and youth optimism (RI = 0.05 each). In the supplementary analyses, when age and sex were included as covariates, youth activity self-efficacy became potentially important (RI = 0.05) and parent optimism became inconclusive (RI = 0.04). Despite different descriptors, changes in RI were minute (RI = ±0.01). In Model 2, 42% of the variance in youth functional disability was accounted for, with youth activity self-efficacy (RI = 0.23), youth pain acceptance (RI = 0.11), and parent psychosocial self-efficacy (RI = 0.07) as the most robust and significant contributors. In Model 3, 59% of the variance in generic HRQoL was explained by the predictors. Youth activity self-efficacy (RI = 0.19), youth psychological flexibility (RI = 0.10), and youth pain acceptance (RI = 0.09) were the most important and robust predictors. Youth optimism (RI = 0.08) was Potentially Important because the effect size decreased in the regression analyses. In the supplementary analyses with covariates, parent psychological flexibility became a potentially important predictor; however, the RI only increased by 0.02. In Model 4, the predictors accounted for 47% of the variance in rheumatology-specific HRQoL, with youth pain acceptance (RI = 0.15) as the most robust contributor. Youth activity and symptom self-efficacy (RI = 0.14 and RI = 0.08, respectively) and parent psychosocial self-efficacy (RI = 0.06) were Potentially Important because their effect sizes decreased in the regression model. Finally, 28% of the variance in benefit finding was explained by predictors (Model 5). Youth optimism was the only robust contributor (RI = 0.12). Youth emotion self-efficacy and parent symptom self-efficacy were also Potentially Important (RI = 0.05 each), albeit to a less

Table 4. Contributions of resilience resources and mechanisms in order of RI to the outcomes of usual pain intensity, functional disability, generic and rheumatology-specific HRQoL, and benefit finding*

	r	*P* (r)	β	*P* (β)	95% CI (β)	R^2	RI
Usual pain intensity	–	–	–	–	–	0.29	–
Pain acceptance (Y)	−0.34	**<0.001**	−0.34	**<0.001**	−0.52, −0.16	–	0.12[a]
Self-efficacy: psychosocial (P)	−0.20	0.071	−0.35	**0.010**	−0.61, −0.08	–	0.07[a]
Optimism (Y)	−0.31	**0.002**	−0.16	0.229	−0.42, 0.10	–	0.05[a]
Optimism (P)	−0.20	0.072	−0.22	**0.019**	−0.41, −0.04	–	0.05[a]
Self-efficacy: activity (Y)	−0.30	**0.007**	−0.14	0.355	−0.44, 0.16	–	0.04
Self-efficacy: symptom (Y)	−0.22	0.055	−0.14	0.322	−0.41, 0.13	–	0.03
Resilience (Y)	−0.28	**0.002**	−0.10	0.402	−0.32, 0.13	–	0.03
Self-efficacy: symptom (P)	0.00	0.999	0.25	**0.037**	0.02, 0.49	–	0.00
Pain acceptance (P)	−0.02	0.852	0.03	0.876	−0.37, 0.43	–	−0.00
Psychological flexibility (P)	−0.04	0.759	0.22	0.338	−0.23, 0.67	–	−0.01
Psychological flexibility (Y)	0.19	**0.042**	−0.14	0.176	−0.35, 0.06	–	−0.03
Self-efficacy: emotion (Y)	−0.23	**0.017**	0.22	0.053	−0.00, 0.45	–	−0.05
Functional disability	–	–	–	–	–	0.42	–
Self-efficacy: activity (Y)	−0.52	**<0.001**	−0.44	**<0.001**	−0.68, −0.19	–	0.23[a]
Pain acceptance (Y)	−0.45	**<0.001**	−0.25	**0.007**	−0.43, −0.07	–	0.11[a]
Self-efficacy: psychosocial (P)	−0.29	**0.005**	−0.23	**0.029**	−0.44, −0.02	–	0.07[a]
Resilience (Y)	−0.36	**<0.001**	−0.09	0.468	−0.31, 0.14	–	0.03
Psychological flexibility (Y)	0.32	**<0.001**	0.07	0.563	−0.17, 0.31	–	0.02
Psychological flexibility (P)	−0.25	**0.027**	−0.09	0.625	−0.44, 0.27	–	0.02
Optimism (P)	−0.21	0.060	−0.08	0.388	−0.28, 0.11	–	0.02
Optimism (Y)	−0.39	**<0.001**	−0.02	0.845	−0.26, 0.21	–	0.01
Self-efficacy: symptom (Y)	−0.28	**0.006**	0.02	0.875	−0.21, 0.25	–	−0.01
Self-efficacy: symptom (P)	−0.14	0.179	0.07	0.549	−0.16, 0.31	–	−0.01
Pain acceptance (P)	−0.19	0.082	0.14	0.408	−0.19, 0.46	–	−0.03
Self-efficacy: emotion (Y)	−0.34	**<0.001**	0.14	0.259	−0.10, 0.37	–	−0.05
HRQoL: generic	–	–	–	–	–	0.59	–
Self-efficacy: activity (Y)	0.60	**<0.001**	0.32	**0.005**	0.09, 0.55	–	0.19[a]
Psychological flexibility (Y)	−0.53	**<0.001**	−0.20	**0.050**	−0.40, −0.00	–	0.10[a]
Pain acceptance (Y)	0.48	**<0.001**	0.18	**0.015**	0.04, 0.33	–	0.09[a]
Optimism (Y)	0.61	**<0.001**	0.14	0.240	−0.09, 0.37	–	0.08[a]
Psychological flexibility (P)	0.18	0.138	0.27	0.064	−0.02, 0.56	–	0.05
Self-efficacy: emotion (Y)	0.54	**<0.001**	0.08	0.433	−0.12, 0.27	–	0.04
Resilience (Y)	0.52	**<0.001**	0.06	0.479	−0.10, 0.22	–	0.03
Self-efficacy: symptom (Y)	0.44	**<0.001**	0.04	0.749	−0.18, 0.26	–	0.02
Self-efficacy: symptom (P)	0.13	0.221	0.08	0.407	−0.11, 0.27	–	0.01
Self-efficacy: psychosocial (P)	0.19	0.104	0.01	0.906	−0.19, 0.22	–	0.00
Optimism (P)	0.17	0.116	0.01	0.904	−0.15, 0.17	–	0.00
Pain acceptance (P)	0.10	0.381	−0.29	**0.045**	−0.57, −0.01	–	−0.03
HRQoL: rheumatology	–	–	–	–	–	0.47	–
Pain acceptance (Y)	0.45	**<0.001**	0.34	**<0.001**	0.19, 0.49	–	0.15[a]
Self-efficacy: activity (Y)	0.53	**<0.001**	0.26	0.064	−0.01, 0.53	–	0.14[a]
Self-efficacy: symptom (Y)	0.40	**<0.001**	0.21	0.127	−0.06, 0.47	–	0.08[a]
Self-efficacy: psychosocial (P)	0.28	**0.018**	0.21	0.067	−0.02, 0.44	–	0.06[a]
Optimism (P)	0.25	**0.027**	0.18	**0.040**	0.01, 0.35	–	0.04
Resilience (Y)	0.38	**<0.001**	0.09	0.309	−0.08, 0.26	–	0.03
Psychological flexibility (P)	0.16	0.202	0.14	0.437	−0.21, 0.49	–	0.02
Optimism (Y)	0.43	**<0.001**	0.04	0.671	−0.16, 0.25	–	0.02
Psychological flexibility (Y)	−0.33	**<0.001**	−0.01	0.910	−0.21, 0.19	–	0.00
Self-efficacy: symptom (P)	0.15	0.190	−0.02	0.853	−0.24, 0.20	–	−0.00
Pain acceptance (P)	0.09	0.485	−0.30	0.080	−0.63, 0.04	–	−0.03
Self-efficacy: emotion (Y)	0.40	**<0.001**	−0.14	0.257	−0.38, 0.10	–	−0.06
Benefit finding	–	–	–	–	–	0.28	–
Optimism (Y)	0.32	**<0.001**	0.36	**0.006**	0.10, 0.62	–	0.12[a]
Self-efficacy: emotion (Y)	0.23	**0.017**	0.22	0.057	−0.01, 0.45	–	0.05[a]
Self-efficacy: symptom (P)	0.22	**0.038**	0.21	0.063	−0.01, 0.43	–	0.05[a]
Pain acceptance (Y)	−0.12	0.172	−0.29	**<0.001**	−0.44, −0.13	–	0.03
Self-efficacy: symptom (Y)	0.24	**0.012**	0.08	0.572	−0.19, 0.35	–	0.02
Self-efficacy: psychosocial (P)	0.14	0.135	0.07	0.537	−0.16, 0.30	–	0.01
Psychological flexibility (P)	−0.10	0.336	−0.07	0.695	−0.40, 0.27	–	0.01
Optimism (P)	−0.03	0.775	−0.15	0.125	−0.34, 0.04	–	0.00

(Continued)

Table 4. (*Cont'd*)

	r	*P* (r)	β	*P* (β)	95% CI (β)	R^2	RI
Psychological flexibility (Y)	−0.20	**0.024**	−0.02	0.881	−0.23, 0.20	–	0.00
Resilience (Y)	0.23	**0.013**	0.01	0.928	−0.22, 0.25	–	0.00
Pain acceptance (P)	−0.08	0.486	0.03	0.834	−0.28, 0.34	–	−0.00
Self-efficacy: activity (Y)	0.10	0.313	−0.17	0.203	−0.43, 0.09	–	−0.02

* RI was calculated as the product of the reported standardized correlation and regression coefficients. Bold values represent significance values ≤ 0.05. CI, confidence interval; HRQoL, health-related quality of life; P, parent; RI, relative importance; Y, youth.
[a] Variables classified as "important" or "potentially important" based on the size of the RI value and pattern of statistical significance. Minor inconsistencies in RI values are due to rounding; calculations were done with higher precision than two decimal places.

robust degree. In the supplementary analysis with covariates, parent symptom self-efficacy became Important because the β coefficient retained significance in the regression model and the RI increased from 0.04 to 0.05.

DISCUSSION

Resilience is a complex process involving an interaction of risk and resilience resources and mechanisms at the individual, familial, and cultural levels resulting in diverse outcomes pertaining to pain adaptation.[7] This study applied a novel approach to this construct in the context of JIA pain to explore the synergy between evidence-based youth and parent resilience resources and mechanisms and how they collectively interact and contribute to proxy measures of pain adaptation.

Significant correlations were observed between resilience resources and mechanisms, especially within individuals. This supports the notion that these variables conceptually align as resilience (rather than risk) resources and mechanisms. These correlations also emphasize the need to determine which variables are the most important predictors of key outcomes (aim 2). As an example, large correlations were observed between parent psychological flexibility and pain acceptance (r = 0.82), which suggests construct overlap and logically makes sense given that pain acceptance (ie, a willingness to permit pain to be present) is encompassed within psychological flexibility (ie, the capacity to stay present focused and engaged in values-based action).

It was hypothesized the resilience resources and mechanisms would predict positive adaptation in the context of JIA pain across five surrogate markers of recovery, sustainability, and growth. Although this was the case, the variables that were most relevant differed slightly based on the outcome in question. This is consistent with the literature emphasizing that the process of resilience, as well as the outcomes that are considered necessary and sufficient, is dependent on the individual and their context.[10]

The biggest contributor to usual pain intensity was pain acceptance. This was closely followed by parent psychosocial self-efficacy and parent optimism, which had stronger effects with

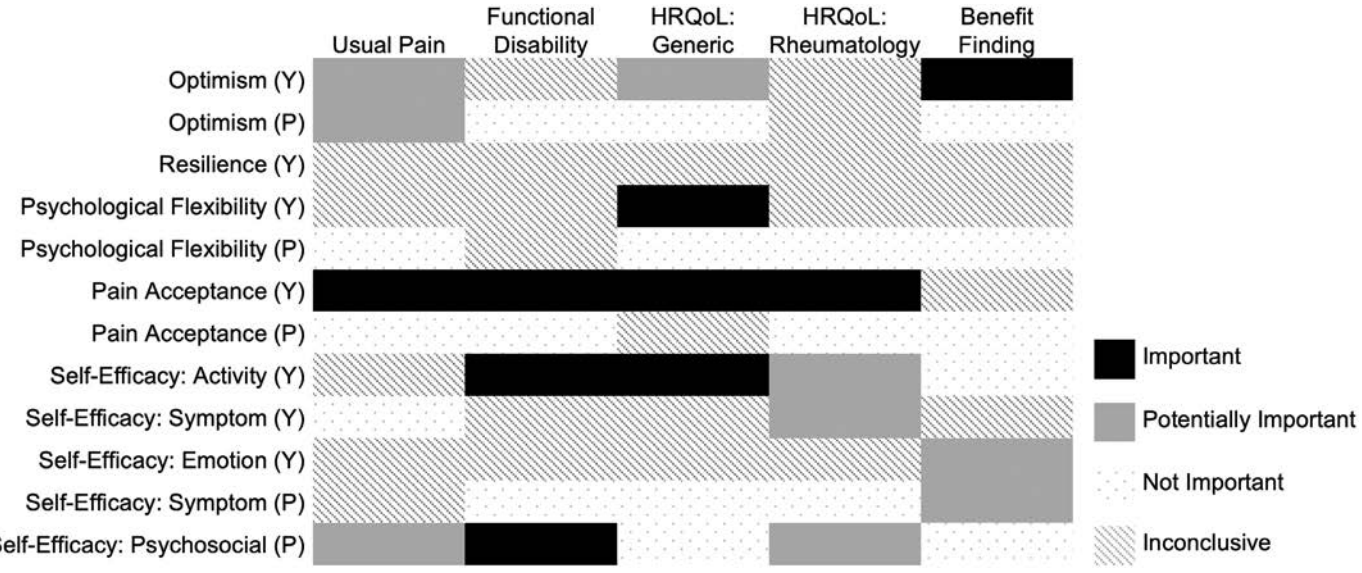


Figure 2. Summary of the significance of each predictor between correlational and regression analyses across models. "Important" means RI ≥ +0.05, and both standardized coefficients are significant. "Potentially Important" means RI ≥ +0.05, and one standardized coefficient is significant. "Not Important" means RI values are between −0.05 and +0.05, and neither standardized coefficient is significant. "Inconclusive" means all other cases where RI < +0.05, but at least one of the two coefficients is statistically significant. HRQoL, health-related quality of life; P, parent; RI, relative importance; Y, youth.

the addition of other predictors, and youth optimism, which decreased in effect size. A recent systematic review conducted by our team explored the psychosocial factors associated with JIA pain.[3] Although pain acceptance and optimism had not been identified, parent psychosocial self-efficacy was a significant protective factor in three of five associations. Other psychosocial factors associated with reduced pain (albeit in fewer associations) included other domains of youth and parent self-efficacy, youth coping via distraction and positive self-statements, and select family factors (eg, family achievement, activities, and cohesion). Interestingly, some of these factors were also significantly correlated with reduced pain intensity in this study; however, when other variables were included, their RI and predictive ability were less stable.

In exploring reduced functional disability, the RI of activity self-efficacy was double that of any other predictor. This is logical because feeling capable of engaging in activities influences one's efforts and actions. Youth pain acceptance and parent psychosocial self-efficacy were also robust contributors to this model. Although trait resilience, youth and parent psychological flexibility, and youth optimism were also relevant, their roles were inconclusive in the regression. These findings are largely consistent with the literature showing support for these predictors independently[13,20,21]; however, there is mixed literature regarding youth pain acceptance[19,28,29] and psychological flexibility[15,17,29], and limited evidence exploring the synergy of these predictors. As such, to improve functional ability in youth with JIA, these findings emphasize the importance of youth prioritizing values-based actions and believing they can engage in activities in the presence of pain, and the importance of parents perceiving they are capable of psychosocially supporting their child in these endeavors.

Generic and rheumatology-specific HRQoL scores are composed of subscales, including physical, emotional, social, and school functioning; and pain and hurt, daily activities, treatments, worry, and communication. Given the array of functional areas addressed, unsurprisingly many of the youth predictors (eg, optimism, pain acceptance, psychological flexibility, self-efficacy) have independently predicted improvements in HRQoL in the literature,[11,12,14,18,24,28,29] which was also seen in Models 3 and 4 to an extent. Interestingly, many youth predictors had weaker effects and lost significance in these regression models. Activity-related self-efficacy, psychological flexibility, and pain acceptance were the most robust predictors, followed by optimism, symptom self-efficacy, and parent psychosocial self-efficacy, which together explained the greatest portion of variance.

Finally, the ability of youth to identify positive consequences of their arthritis was best predicted by youth optimism. Youth emotion self-efficacy and parent symptom self-efficacy were also important but demonstrated collinearity. Although benefit finding has not received much attention in the JIA literature, it has been associated with optimism and self-esteem in pediatric oncology patients.[42] Interestingly, in pediatric patients with chronic pain, it demonstrated an inverse effect, wherein it was associated with reduced quality of life and greater posttraumatic stress disorder, anxiety, and depression symptoms. Authors posited this was due to the complex nature of living with chronic pain.[47] Although an inverse relationship between benefit finding and HRQoL was not seen in this study (all correlations were nonsignificant), it is possible that unlike a diagnosis of chronic pain, which has been associated with a perceived lack of physician understanding,[48] a diagnosis of JIA may facilitate the process of resilience. Given these discrepancies, it will be critical for future research to further explore this outcome.

Despite these differences across models, key patterns emerged. There were generally large r coefficients, which shrunk after controlling for other variables (suggesting construct overlap), with fewer instances of suppressor variables or β coefficients growing with the addition of other predictors. Youth pain acceptance was the most robust predictor across outcomes and was always among the top four contributors in terms of RI. Thus, a willingness to permit pain to be present and persist with valued activities is a key mechanism of change in fostering resilience, regardless of the outcome in question. This maps to the existing literature demonstrating its importance as a main effect or mediator,[14,18,19,28] with the added notion that it maintains its unique contributions even in the presence of other resources and mechanisms. Comparatively, some predictors consistently demonstrated weaker relationships to outcomes, including youth trait resilience (construct overlap) and parent pain acceptance and psychological flexibility (largely classified as "Not Important"). These variables were likely less important because parents overall made weaker contributions to youth outcomes. It may be that a parent's capacity to stay present focused, engaged in values-based actions, and hopeful in their ability to manage their child's symptoms is developmentally less relevant for adolescents as they increase their independence and self-management. Although unimportant in these analyses, they are likely more relevant for the parent's own adaptation. Moreover, at least one parent resilience resource or mechanism (often psychosocial self-efficacy) was significant across most models, indicating that parents have some influence on child outcomes.

These results generally support the Ecological Resilience-Risk Model in Pediatric Chronic Pain[7] in the context of JIA, particularly for within-youth resources and mechanisms. Moreover, this study addressed numerous gaps in the literature.[22] Not only did this study identify the role of novel resilience resources and mechanisms in this population, it illuminated select constructs that hold greater weight in terms of one's pain adaptation and demonstrated the ways in which many of these resources and mechanisms coexist. Clinically, harnessing resilience to promote adaptation in the face of adversity aligns well with Acceptance and Commitment Therapy.[49] This study demonstrated support for numerous protective and/or promotive factors; however, the factors most relevant for clinicians to focus on will largely depend on client goals. Thus, even though a variable may have been

classified as "Not Important" in these models, there may be circumstances in which it is relevant.

Limitations of this study include the use of an internet survey design, which resulted in reliance on self-reported data for disease characteristics, low retention, and missing data. Akin to much of the existing literature, this study was also limited by its sample size. As such, covariates such as diagnosis and treatments and other evidence-based resilience resources and mechanisms (eg, trait mindfulness, positive affect) were not incorporated. Finally, although this sample had adequate representation by way of disease characteristics, this was less true for demographic variables (eg, largely female and adolescent participants). It is possible that results may not generalize to younger children and/or male individuals and that the relevant resources and mechanisms may differ based on disease characteristics and pain experiences.

In addition to promoting data sharing and multisite collaboration to increase sample sizes, there is a need to develop and validate scales to measure other resilience resources and mechanisms (eg, committed action, self-regulation, sense of self)[7]; incorporate perspectives from others within the child's network (eg, siblings)[22]; and statistically account for the biological, developmental, social, and cultural milieu to identify how these factors might interact with those identified in this study. Furthermore, more complex methodological and statistical approaches (eg, longitudinal designs, profile analyses, network analyses) and the use of qualitative and/or mixed-methods approaches[22] would enrich our understanding of resilience in youth with JIA. Finally, there is value in rigorously testing the effects of strengths-based interventions,[7] particularly those incorporating the identified predictors.

This study explored the relationships between, and predictive ability of, various youth and parent resilience resources and mechanisms in predicting pain adaptation in the context of JIA. In addition to demonstrating how predictors depend on the surrogate marker being used, an important pattern emerged wherein pain acceptance was one of the most robust predictors across outcomes. Other predictors, including some parent predictors, also played an important role. These findings further support the Ecological Resilience-Risk Model in Pediatric Chronic Pain and have important implications for the process's interventions should emphasize when helping youth and parents adjust to living with JIA.

ACKNOWLEDGMENTS

We would like to thank Isabel Jordan (Strategic Lead, Patient Partnerships, Chambers' Lab) for helping to create a patient partnership plan for this research. We are grateful for the contributions of Cassie and Friends, a nonprofit organization dedicated to the pediatric rheumatology community, for their partnership in this work. We would also like to acknowledge the contributions of Tiffany Thompson, Abby Mazzone, and Amy Mazzone, patient and family partners who have been instrumental in the study design, execution, and interpretation. We would like to acknowledge IWK Health's rheumatology clinic and staff as well as the numerous arthritis organizations, industry partners, parent groups, and individuals who have helped to share the recruitment callouts for this study. Last, but not least, we would like to thank all of those who took part in this research.

AUTHOR CONTRIBUTIONS

All authors contributed to at least one of the following manuscript preparation roles: conceptualization AND/OR methodology, software, investigation, formal analysis, data curation, visualization, and validation AND drafting or reviewing/editing the final draft. As corresponding author, Ms Brandelli 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/Helsinki Declaration requirements.

REFERENCES

1. Canadian Paediatric Society. Canadian paediatric surveillance program: 2009 results. Public Health Agency of Canada. 2009. https://cpsp.cps.ca/uploads/publications/Results-2009.pdf
2. Correll CK, Dave M, Paul AF, et al. Identifying research priorities among patients and families of children with rheumatic diseases living in the United States. J Rheumatol 2020;47(12):1800–1806.
3. Brandelli YN, Chambers CT, Mackinnon SP, et al. A systematic review of the psychosocial factors associated with pain in children with juvenile idiopathic arthritis. Pediatr Rheumatol Online J 2023;21(1):57.
4. Stinson JN, Stevens BJ, Feldman BM, et al. Using an electronic pain diary to better understand pain in children and adolescents with arthritis. Pain Manag (Lond) 2011;1(2):127–137.
5. Zimmerman MA. Resiliency theory: a strengths-based approach to research and practice for adolescent health. Health Educ Behav 2013;40(4):381–383.
6. Masten AS. Global perspectives on resilience in children and youth. Child Dev 2014;85(1):6–20.
7. Cousins LA, Kalapurakkel S, Cohen LL, et al. Topical review: resilience resources and mechanisms in pediatric chronic pain. J Pediatr Psychol 2015;40(9):840–845.
8. Beukelman T, Patkar NM, Saag KG, et al. 2011 American College of Rheumatology recommendations for the treatment of juvenile idiopathic arthritis: initiation and safety monitoring of therapeutic agents for the treatment of arthritis and systemic features. Arthritis Care Res (Hoboken) 2011;63(4):465–482.
9. Sturgeon JA, Zautra AJ. Resilience: a new paradigm for adaptation to chronic pain. Curr Pain Headache Rep 2010;14(2):105–112.
10. Rosenberg AR, Yi-Frazier JP. Commentary: resilience defined: an alternative perspective. J Pediatr Psychol 2016;41(5):506–509.
11. Tomlinson RM, Bax KC, Ashok D, et al. Health-related quality of life in youth with abdominal pain: an examination of optimism and pain self-efficacy. J Psychosom Res 2021;147:110531.
12. Cousins LA, Cohen LL, Venable C. Risk and resilience in pediatric chronic pain: exploring the protective role of optimism. J Pediatr Psychol 2015;40(9):934–942.
13. Gmuca S, Xiao R, Urquhart A, et al. The role of patient and parental resilience in adolescents with chronic musculoskeletal pain. J Pediatr 2019;210:118–126.e2.
14. Wright LA, Cohen LL, Gise J, et al. Pain and QOL in pediatric sickle cell disease: buffering by resilience processes. J Pediatr Psychol 2021;46(8):1015–1024.

15. Beeckman M, Simons LE, Hughes S, et al. A network analysis of potential antecedents and consequences of pain-related activity avoidance and activity engagement in adolescents. Pain Med 2020; 21(2):e89–e101.
16. Forgeron PA, McGrath P, Stevens B, et al. Social information processing in adolescents with chronic pain: my friends don't really understand me. Pain 2011;152(12):2773–2780.
17. Beeckman M, Simons LE, Hughes S, et al. Investigating how parental instructions and protective responses mediate the relationship between parental psychological flexibility and pain-related behavior in adolescents with chronic pain: a daily diary study. Front Psychol 2019;10:2350.
18. Lee S, McMurtry CM, Summers C, et al. Quality of life in youth with chronic pain: an examination of youth and parent resilience and risk factors. Clin J Pain 2020;36(6):440–448.
19. Weiss KE, Hahn A, Wallace DP, et al. Acceptance of pain: associations with depression, catastrophizing, and functional disability among children and adolescents in an interdisciplinary chronic pain rehabilitation program. J Pediatr Psychol 2013;38(7):756–765.
20. Feinstein AB, Sturgeon JA, Bhandari RP, et al. Risk and resilience in pediatric pain: the roles of parent and adolescent catastrophizing and acceptance. Clin J Pain 2018;34(12):1096–1105.
21. Kalapurakkel S, Carpino EA, Lebel A, et al. "Pain can't stop me": examining pain self-efficacy and acceptance as resilience processes among youth with chronic headache. J Pediatr Psychol 2015;40(9): 926–933.
22. Hynes L, Saetes S, McGuire B, et al. Child and family adaptation to juvenile idiopathic arthritis—a systematic review of the role of resilience resources and mechanisms. Front Psychol 2019;10:2445.
23. Connelly TWJ Jr. Family functioning and hope in children with juvenile rheumatoid arthritis. MCN Am J Matern Child Nurs 2005;30(4):245–250.
24. Seid M, Huang B, Niehaus S, et al. Determinants of health-related quality of life in children newly diagnosed with juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2014;66(2):263–269.
25. Barlow JH, Shaw KL, Wright CC. Development and preliminary validation of a children's arthritis self-efficacy scale. Arthritis Care Res 2001;45(2):159–166.
26. Sawyer MG, Whitham JN, Roberton DM, et al. The relationship between health-related quality of life, pain and coping strategies in juvenile idiopathic arthritis. Rheumatology (Oxford) 2004;43(3):325–330.
27. Barlow JH, Shaw KL, Wright CC. Development and preliminary validation of a self-efficacy measure for use among parents of children with juvenile idiopathic arthritis. Arthritis Care Res 2000;13(4): 227–236.
28. Beeckman M, Hughes S, Van Ryckeghem D, et al. Resilience factors in children with juvenile idiopathic arthritis and their parents: the role of child and parent psychological flexibility. Pain Med 2019;20(6): 1120–1131.
29. Feinstein AB, Forman EM, Masuda A, et al. Pain intensity, psychological inflexibility, and acceptance of pain as predictors of functioning in adolescents with juvenile idiopathic arthritis: a preliminary investigation. J Clin Psychol Med Settings 2011;18(3):291–298.
30. Knafl K, Leeman J, Havill NL, et al. The contribution of parent and family variables to the well-being of youth with arthritis. J Fam Nurs 2015; 21(4):579–616.
31. Canadian Institute for Health Information. Guidance on the use of standards for race-based and indigenous identity data collection and health reporting in Canada. 2022. www.cihi.ca/sites/default/files/document/guidance-and-standards-for-race-based-and-indigenous-identity-data-en.pdf
32. Ey S, Hadley W, Allen DN, et al. A new measure of children's optimism and pessimism: the youth life orientation test. J Child Psychol Psychiatry 2005;46(5):548–558.
33. Scheier MF, Carver CS, Bridges MW. Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): a reevaluation of the Life Orientation Test. J Pers Soc Psychol 1994;67(6): 1063–1078.
34. Smith BW, Dalen J, Wiggins K, et al. The brief resilience scale: assessing the ability to bounce back. Int J Behav Med 2008;15(3):194–200.
35. Greco LA, Lambert W, Baer RA. Psychological inflexibility in childhood and adolescence: development and evaluation of the Avoidance and Fusion Questionnaire for Youth. Psychol Assess 2008;20(2):93–102.
36. Timmers I, Simons LE, Hernandez JM, et al. Parent psychological flexibility in the context of pediatric pain: brief assessment and associations with parent behaviour and child functioning. Eur J Pain 2019; 23(7):1340–1350.
37. Gauntlett-Gilbert J, Alamire B, Duggan GB. Pain acceptance in adolescents: development of a short form of the CPAQ-A. J Pediatr Psychol 2019;44(4):453–462.
38. Smith AM, Sieberg CB, Odell S, et al. Living life with my child's pain: The Parent Pain Acceptance Questionnaire (PPAQ). Clin J Pain 2015;31(7):633–641.
39. Birnie KA, Hundert AS, Lalloo C, et al. Recommendations for selection of self-report pain intensity measures in children and adolescents: a systematic review and quality assessment of measurement properties. Pain 2019;160(1):5–18.
40. Walker LS, Greene JW. The functional disability inventory: measuring a neglected dimension of child health status. J Pediatr Psychol 1991;16(1):39–58.
41. Varni JW, Seid M, Smith Knight T, et al. The PedsQL in pediatric rheumatology: reliability, validity, and responsiveness of the pediatric quality of life inventory generic core scales and rheumatology module. Arthritis Rheum 2002;46(3):714–725.
42. Phipps S, Long AM, Ogden J. Benefit Finding Scale for Children: preliminary findings from a childhood cancer population. J Pediatr Psychol 2007;32(10):1264–1271.
43. Teitcher JE, Bockting WO, Bauermeister JA, et al. Detecting, preventing, and responding to "fraudsters" in internet research: ethics and tradeoffs. J Law Med Ethics 2015;43(1):116–133.
44. Pratt JW. Dividing the indivisible: using simple symmetry to partition variance explained. In: Pukkila T, Puntanen S, eds. Proceedings of the Second International Conference in Statistics. University of Tampere; 1987:245–260.
45. Richard FD, Bond CFJ Jr, Stokes-Zoota JJ. One hundred years of social psychology quantitatively described. Rev Gen Psychol 2003; 7(4):331–363.
46. Thomas DR, Hughes E, Zumbo BD. On variable importance in linear regression. Soc Indic Res 1998;45(1/3):253–275.
47. Soltani S, Neville A, Hurtubise K, et al. Finding silver linings: a preliminary examination of benefit finding in youth with chronic pain. J Pediatr Psychol 2018;43(3):285–293.
48. Meldrum ML, Tsao JCI, Zeltzer LK. "I can't be what I want to be": children's narratives of chronic pain experiences and treatment outcomes. Pain Med 2009;10(6):1018–1034.
49. Pielech M, Vowles KE, Wicksell R. Acceptance and commitment therapy for pediatric chronic pain: theory and application. Children (Basel) 2017;4(2):1–12.
50. Smith BW, de Cruz-Dixon N, Schodt K, et al. Brief Resilience Scale (BRS). In: Medvedev O, Krägeloh CU, Siegert RJ, et al, eds. Handbook of Assessment in Mindfulness Research. Springer; 2023.

Arthritis Care & Research
Vol. 77, No. 3, March 2025, pp 322–329
DOI 10.1002/acr.25433

American College of Rheumatology
Empowering Rheumatology Professionals

Measurable Outcomes of an Ophthalmology and Rheumatology Coordinated Care Clinic

Catherine Lavallee,[1] Monica Ahrens,[2] Stefanie L. Davidson,[3] Haseeb Goheer,[4] Allison Shuster,[4] and Melissa A. Lerman[3]

Objective. We evaluated the impact of an Ophthalmology/Rheumatology multidisciplinary clinic for patients with anterior uveitis by comparing outcomes between those who received traditional care (TC) versus coordinated care (CC).

Methods. We conducted a retrospective cohort study of children with anterior uveitis from a pediatric tertiary care center between 2013 and 2022. Standard descriptive statistics were used; survival analyses explored differences in cohort disease activity and biologic disease-modifying antirheumatic drug (DMARD) treatment. Steroid treatment by cohort was compared using generalized estimating equation model with Poisson distribution and log link. Complications were compared using logistic regression. Number of visits in each cohort were assessed using Poisson generalized estimating equation model adjusted for complications.

Results. We studied 215 patients with anterior uveitis; 66% were female, 53% had juvenile idiopathic arthritis, and 23% were idiopathic, with a median age at diagnosis of 8 years old (interquartile range 5–12). CC was associated with a 60% higher hazard of reaching disease control (hazard ratio 1.6; $P < 0.01$) when controlling for time since diagnosis and anterior chamber cell counts at the beginning of disease activity. CC was associated with starting biologic DMARDs earlier than TC ($P < 0.01$). Compared with the group who received TC, the group who received CC had a 96% lower rate of glucocorticoid reception per appointment within the first year ($P < 0.01$). The visit rate was 64% lower for the group who received CC when controlling for total complications per patient.

Conclusion. Patients who received multidisciplinary care had better outcomes than patients who received TC. Limitations include different cohort start times and absence of defined criteria for CC referral.

INTRODUCTION

Pediatric noninfectious chronic anterior uveitis (CAU) is an intraocular disease characterized by inflammation of the anterior chamber of the eye, or anterior uveitis (AU). Uveitis is also a comorbid feature of many other autoimmune diseases.[1–3] Pediatric CAU frequently has few symptoms and can be difficult to detect until complications have occurred.[4] Both CAU and its first-line treatment, topical glucocorticoids, can cause serious complications, including cataracts and elevations in intraocular pressure (IOP), which may result in glaucoma and vision loss.[5] Standard first-line treatment includes topical glucocorticoids and dilating drops to limit synechiae development.[3,6,7] It is important to diagnose and quickly treat uveitis to prevent permanent damage.[8–10] Consensus-based guidelines have been created for uveitis treatment with similar recommendations: patients resistant to initial treatment should receive systemic immunosuppressive agents, with methotrexate as the first-line agent.[7,11,12] In patients with inflammation resistant to methotrexate, subsequent options include tumor necrosis factor inhibitors, other biologic agents, and other conventional disease-modifying antirheumatic drugs (DMARDs).[7,11,12]

For patients for whom treatment needs to be escalated, many ophthalmologists refer to rheumatologists to prescribe systemic medication.[11] Barriers hindering timely and accurate communication between ophthalmologists and rheumatologists

Supported by Rheumatology Research Foundation Medical Student Preceptorship and the Department of Rheumatology at Children's Hospital of Philadelphia.

[1]Catherine Lavallee, MD: Virginia Tech Carilion School of Medicine, Roanoke; [2]Monica Ahrens, PhD: Virginia Tech, Roanoke; [3]Stefanie L. Davidson, MD, Melissa A. Lerman, MD, PhD, MSCE: Perelman School of Medicine at the University of Pennsylvania and Children's Hospital of Philadelphia; [4]Haseeb Goheer, BS, Allison Shuster, BS: Children's Hospital of Philadelphia, Philadelphia, Pennsylvania.

Author disclosures and graphical abstract are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25433.

Address correspondence via email to Melissa A. Lerman, MD, PhD, MSCE, at lermanm@chop.edu.

Submitted for publication February 8, 2024; accepted in revised form August 19, 2024.

SIGNIFICANCE & INNOVATIONS

- The uveitis registry at the Children's Hospital of Philadelphia provides a unique opportunity to report clinical outcomes of two care modalities: traditional and multidisciplinary care.
- In this study, we demonstrate that coordinated care improved important uveitis outcomes, including time to disease control, minimized glucocorticoid use and accelerated initiation of biologic disease-modifying antirheumatic drugs.

comanaging patients with uveitis include different electronic health records, incorrect documentation, and reliance on patients to facilitate communication.[13] These barriers may delay decisions to change systemic agents, resulting in worse disease activity or outcomes.[13]

Consensus guidelines have recommended multidisciplinary care for improved communication.[12,13] Research has supported the benefits of multidisciplinary clinics in other fields.[14–16] However, there are limited data available on the performance of multidisciplinary clinics for uveitis.[17–20,21] There are multiple models of care that facilitate increased communication between rheumatologists and ophthalmologists. One model consists of regular meetings between ophthalmologists, rheumatologists, and other care team members, to discuss mutual patients. Another model includes an asynchronous team-based approach, wherein ophthalmologists fill out a shared safety monitoring spreadsheet for patients treated with conventional DMARDs, but patients treated with biologics see both a rheumatologist and an ophthalmologist who are in close communication.[22] Both of these models allow for group input in decision-making about management, and the patient then has a visit with only one of the clinicians to make a shared treatment decision. A third model is a combined clinic where the ophthalmologist and rheumatologist see the patient during the same visit.

In the Uveitis Coordinated Care Clinic at the Children's Hospital of Philadelphia (CHOP), a rheumatologist and ophthalmologist evaluate patients with uveitis back-to-back in one visit, and then together meet with the patient and their family. The aim is to increase communication between the physicians, improve family understanding, and optimize medication treatment, resulting in expedited disease management and ultimately improved outcomes.

The uveitis registry at CHOP provides a unique opportunity to report clinical outcomes of these care modalities, to evaluate whether the multidisciplinary model consolidates appointments for patients, and to demonstrate if there are tangible results from the improved communication between the ophthalmologists and rheumatologists of this model. The goal of this study was to evaluate whether receiving care in a multidisciplinary clinic will result in improved disease outcomes. To answer this question, we examined three variables: (1) the time to attain AU control and to initiate biologic DMARD reception, (2) ophthalmic complications, (3) the burden of medical care, including topical glucocorticoid treatment and the number of uveitis-associated visits patients attended.

PATIENTS AND METHODS

Patients. The study was performed using a registry of all patients with uveitis seen in the rheumatology division of a pediatric tertiary care center between 2013 and 2022. This was a retrospective cohort study restricted to children with AU aged ≤ 19 years. The registry recorded four types of visits: coordinated care (CC), CHOP ophthalmology, outside ophthalmology, and CHOP rheumatology. Visits included in this analysis were CC, CHOP ophthalmology, and outside ophthalmology visits.

CC clinic. The Uveitis Coordinated Care Clinic at the CHOP is held three out of four weeks of each month and based in the ophthalmology clinic space at the main campus, with one room modified to include an examination table and other tools for rheumatology examination. This clinic is staffed by one rheumatologist (MAL) and one ophthalmologist (SD).

Charts are reviewed before clinic visits to identify patients who are known in advance to need procedures (eg, dilation, optical coherence tomography, visual field testing, fluorescein angiogram [oral]). During the time frame included in this study, patients were seen by an ophthalmology technician who performed vitals and medicine reconciliation and assessed vision and IOP. Vitals are not part of a standard ophthalmology examination; technicians were trained for this; a scale, stadiometer, and blood pressure cuff were obtained for the coordinated clinic. A small cohort of technicians staff the coordinated clinic and are familiar with its workflow. Pressure is checked first with a tonometer. If the pressure is abnormal (low or high), it is repeated for reproducibility by the tonometer and then applanation. The patient is generally seen by ophthalmology first. As such, anterior chamber activity is available during the rheumatology visit and patients who require dilation have drops instilled. Complete rheumatologic examinations are performed at the initial visit and for any patient with a known rheumatic disease; if necessary, ultrasound and joint injections are performed. If a patient does not require a full joint examination at every visit, such as a patient with a well-established diagnosis of idiopathic uveitis, joint range of motion and strength can be tested in the eye examination chair. After both clinicians have completed their examinations, they meet with the patient and family to discuss disease activity, create a treatment plan, and schedule appointments (providers maintain a scheduling spreadsheet). Effort is made to schedule ophthalmology or CC visits such that they align with infusions (for patients whose insurance enables them to receive infusions at CHOP). Medication teaching and complex care coordination occur during

the visit. About 10 patients are scheduled for coordinated visits each clinic day. The ophthalmologist separately sees an additional 20 patients—some of those are in the CC cohort and some do not ever participate in the CC. The ophthalmologist will often discuss examination findings and care changes with the rheumatologist for patients in CC being seen only by ophthalmology on a given day.

Referrals to the CC clinic must come from a rheumatologist, ophthalmologist, or primary care provider. Common reasons for referral include significant disease burden or side effects of uveitis, and previously requiring a biologic drug. Some families prefer CC, whereas others prefer to remain with their initial providers or find the distance to CC onerous (eg, may be seen at CHOP satellites otherwise). Patients are accepted after records have been reviewed, confirming that the patient has a uveitis diagnosis. Some patients receive CC but visit with local ophthalmology providers for intervening visits.

Data collection. Demographic and clinical characteristics were abstracted from the electronic health record. Patient demographic and baseline disease data was abstracted for all patients and included: sex, age at diagnosis, rheumatologic diagnosis, juvenile idiopathic arthritis (JIA) subtype (for patients with JIA), antinuclear antibody status, anterior chamber (AC) cell count at first visit (baseline) and at subsequent visits. Clinical outcomes included number of glucocorticoid drops used at baseline and subsequent follow-ups, systemic medications, and complications. Complications examined were band keratopathy, cataracts, and elevated intraocular pressure.

Definitions. Uveitis location and activity was defined by the Standardization of Uveitis Nomenclature (SUN) Working Group.[20] Patients could receive traditional care (TC) with separate visits with rheumatologists and ophthalmologists, or care in the Uveitis Coordinated Care Clinic in which they saw a rheumatologist and ophthalmologist together in one visit (CC). Inclusion criteria for the CC cohort: ≥2 visits in the CC clinic within the first 6 months of first CC visit; along with ≥1 additional visit in the subsequent 6 months (minimum of 3 CC visits within year 1). If criteria were not met, patients were assigned to the TC cohort. Start time for the CC cohort was defined as the first CC visit in the registry; it was defined as the first ophthalmology appointment in the registry for the TC cohort. Patients were allowed to transition from the TC cohort to the CC cohort, and time spent in each cohort is included in the disease control analysis. This allows for extra data to be included in the analysis, for a direct baseline comparison and helps reduce confounding by enabling direct comparison across the care types. Patients were included if they had at least one ophthalmology visit with recorded AC cell count and topical glucocorticoid drop amount (any topical glucocorticoid). The topical glucocorticoid prescribing pattern of the CC clinic is to first prescribe prednisolone acetate, then escalate to difluprednate, and finally, for patients who required long term topical glucocorticoids, to switch to fluorometholone.

Disease control was defined as two consecutive visits at which the patient had ≤0.5 AC cell (as per SUN criteria) and ≤2 topical glucocorticoid drops daily.[20,23] JIA subtypes were defined according to the International League of Associations for Rheumatology criteria.[24] Given that uveitis reactivates, patients could have multiple opportunities to achieve disease control. In the primary analyses, all episodes of control were examined. Sensitivity analyses were performed by restricting only to the first episode of control in each cohort. Disease activity was ordinal and was categorized according to the maximal AC cell count in either eye at a visit.

Conventional DMARDs considered for this study were methotrexate, leflunomide, and mycophenolate mofetil. Biologics considered were tumor necrosis factor inhibitors (adalimumab, infliximab, golimumab and certolizumab), tocilizumab, rituximab, and abatacept. Complications were defined as the presence of band keratopathy (+/–) or cataracts (+/–), and a reproducible elevated IOP (>21 mm Hg). The assumption was made that a patient could only develop cataracts in each eye once, and the same assumption was made for band keratopathy.

Statistical analysis. Statistical analyses were performed using R (version 4.3.1).[25] Patient demographics and clinical manifestations were evaluated using descriptive statistics. Recurrent event Cox proportional hazards (Cox PH) models were used to assess time to reach control and time to first biologic DMARD reception in each cohort. Many participants were receiving biologic DMARDs upon entering the CC cohort; to limit confounding effects of previous biologic use, only participants not already treated with a biologic were included in this analysis. The predictor of interest, CC, was tested to determine if there was an association with time to control. Covariates included in the time to control model were AC cell count, time since diagnosis, medication prescribed, and JIA status. Baseline AC count and time since diagnosis were included in the model for time to the first biologic reception. Hazard ratios (HRs) for each of these covariates are estimates in which the hazard refers to the probability the event (reaching control) will happen given it has not yet happened.

The rate of topical glucocorticoid reception, rate of visits per patient year, and total complication rates across the two cohorts were compared using generalized linear mixed models with Poisson distribution and log link, controlling for repeated measures assuming a compound symmetric working correlation structure. Additionally, each complication of interest was modeled separately, using a generalized linear mixed effects model with binomial outcome and logit link to compare the odds of developing cataracts or band keratopathy between the two groups and generalized estimating equations with Poisson distribution and log link to look at the rates of elevated IOP across the two groups. An exemption was granted for this study by The CHOP

Institutional Review Board (22-019708) for the conduct of secondary research for which consent is not required.

RESULTS

The registry had 321 patients, 215 of whom had AU and data available for analysis (Figure 1). The cohort who received TC included 170 patients, and 45 patients were included in the cohort who received CC. The population was predominantly composed of female patients (65%) with a median age at diagnosis of 8 (interquartile range 5–12) years (Table 1). Although about half of the patients who received TC had a JIA diagnosis, patients with JIA comprised a lower percentage of the cohort receiving CC. Those with idiopathic AU comprised 19% of the cohort who received TC versus 36% of the cohort who received CC. Of the patients with JIA, 54% of the cohort receiving TC had the oligoarticular subtype, and 71% of the cohort receiving CC had the oligoarticular subtype. At baseline, fewer patients in the cohort receiving TC had an AC cell count of two or more (41.1%) than in the cohort receiving CC (60%). Within the registry period, 58.5% of patients were prescribed a biologic DMARD, with 52.3% in the cohort who received TC and 86% in the cohort who received CC at least one time. Of all the participants, 64.8% had been prescribed a nonbiologic DMARD, with 52.3% in the cohort who received TC and 86% in the cohort who received CC.

Time to control analysis. The nonadjusted survival curves for each group, as estimated with Kaplan–Meier estimates, demonstrated that patients who received CC achieved disease control more rapidly (log rank test $P < 0.01$) (Figure 2).

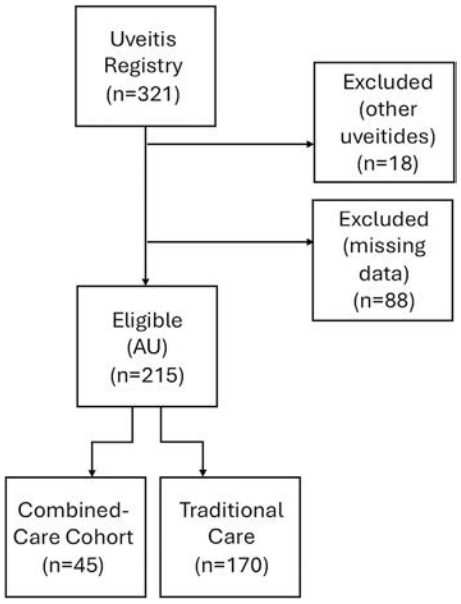


Figure 1. Consort diagram of the study cohorts and sample sizes. AU, anterior uveitis.

Median time to control for the cohort who received CC was 77 days (95% confidence interval [CI] 72–119) and for the cohort who received TC was 136 days (95% CI 118–166). A Cox proportional hazard model examined time until participants reached control across the two groups of interest (Table 2). Being in the group who received CC was associated with 46% higher hazard (HR 1.46; 95% CI 1.01–2.12; $P = 0.02$) of reaching disease control, when controlling for time since diagnosis, AC cell count at beginning of the disease episode, and JIA diagnosis. Time since diagnosis, uveitis activity, and a JIA diagnosis were included in models as covariates to account for their confounding.

In a model that controlled for AC cell count at the beginning of each episode, time since diagnosis at the beginning of the disease episode, JIA status, and medications received as a time-varying covariate, there was no longer a statistically significant association between cohort and time to control (Table 2). Instead, there were statistically significant associations with JIA status and medication seemed to have a relationship with time to control. Patients with JIA had a 29% lower hazard of control than patients without JIA (HR 0.71; 95% CI 0.49–1.03; $P = 0.07$). Treatment with a nonbiologic DMARD was statistically associated with a 53% higher hazard of control compared to no medication (HR 1.53, 95% CI 1.01–2.3; $P = 0.04$). Receiving a biologic DMARD was statistically associated with a 72% higher hazard of control compared to receiving no medication (HR 1.72; 95% CI 1.18–2.52; $P < 0.01$). Finally, receiving both a biologic DMARD and a nonbiologic DMARD was associated with a 2.5 times higher rate of hazard of control than receiving on no medication (HR 2.49; 95% CI 1.68–3.68; $P < 0.01$).

We also analyzed time to control by restricting analysis to the time to the first episode of control after first disease activation (data not shown). The cohort who received CC continued to have statistically significant increase in hazard of success which was higher than in the time to any episode of control analysis (HR 1.87, 95% CI 1.15–3.06, $P < 0.01$) when similarly adjusted for initial AC count and time since diagnosis.

Patterns of biologic treatment. After removing the participants who entered their cohort already receiving a biologic, there were 161 patients in the group who received TC and 27 in the group who received CC. The test comparing time to first reception of biologic DMARDs demonstrated that being in the cohort who received CC, compared to the cohort who received TC, was associated with a difference in survival functions of treatment with biologic DMARDs (HR 2.46; 95% CI 1.43–4.23; $P < 0.01$; Figure 3). The median time to receiving a biologic for the cohort who received TC was 662 days (95% CI 417–1,520) and for the cohort who received CC was 98 days (95% CI 56–273) for. More patients who received CC were treated with biologics at the end of time measured; however, there were also more patients who received CC treated with biologics at the beginning of the CC than at the beginning of TC.

Table 1. Patient demographic and clinical characteristics*

Characteristic	Overall (N = 215)	Cohort who received TC (n = 170)	Cohort who received CC (n = 45)
Sex			
Female	141 (65.6)	112 (65.9)	29 (64.4)
Male	74 (34.4)	58 (34.1)	16 (35.6)
Age at diagnosis, median (IQR), y	8.38 (4.94–12.1)	8.35 (4.92–12)	9.37 (5.2–13.1)
Systemic inflammatory disease			
JIA	113 (52.6)	96 (56.5)	17 (37.8)
Oligoarticular	64 (56.6)	52 (54.2)	12 (70.6)
Polyarticular RF−	13 (11.5)	11 (11.5)	2 (4.4)
Polyarticular RF+	2 (1.8)	2 (2.1)	0 (0)
ERA	20 (17.7)	19 (19.8)	1 (2.2)
PsA	7 (6.2)	6 (6.3)	1 (2.2)
Undifferentiated	5 (4.4)	4 (4.2)	1 (2.2)
Missing/unknown	2 (1.8)	2 (2.1)	0 (0)
TINU	15 (7)	12 (7.1)	3 (6.7)
Sarcoid/Blau	5 (2.3)	3 (1.8)	2 (4.4)
Idiopathic	48 (22.3)	32 (18.8)	16 (35.6)
Vasculitis/MS	4 (1.9)	4 (2.4)	0 (0)
Missing/unknown	30 (14)	23 (13.5)	7 (15.6)
ANA			
Positive	108 (50.2)	87 (51.2)	21 (46.7)
Negative	92 (42.8)	73 (42.9)	19 (42.2)
Not done/unknown	15 (7)	10 (5.9)	5 (11.1)
Baseline AC cell count, cells/hpf[a]			
<1	51 (23.8)	42 (24.7)	9 (20)
≥0.5	22 (10.2)	21 (12.4)	1 (2.2)
≥1	45 (20.9)	37 (21.8)	8 (17.8)
≥2	51 (23.7)	41 (24.1)	10 (22.2)
≥3	32 (14.9)	22 (12.9)	10 (22.2)
≥4	14 (6.5)	7 (4.1)	7 (15.6)
Had cataracts	33 (15.1)	28 (16.2)	5 (10.9)
Had band keratopathy	73 (33.3)	65 (37.6)	8 (17.4)
Had high pressure	8 (5.1)	7 (6.1)	1 (2.4)
Time from diagnosis to entry into cohort, days, median (IQR), d	0 (−287 to 5,090)	0 (−287 to 4,060)	62.0 (−1 to 5,090)

* Values are n (%) unless indicated otherwise. AC, anterior chamber; ANA, antinuclear antibodies; CC, coordinated care; ERA, enthesitis-related arthritis; hpf, high powered field; IQR, interquartile range; JIA, juvenile idiopathic arthritis; MS, multiple sclerosis; PsA, psoriatic arthritis; RF, rheumatoid factor; TC, traditional care; TINU, tubulointerstitial nephritis and uveitis syndrome.
[a] AC cell count is defined per the Standardization of Uveitis Nomenclature criteria.

Steroid exposure. According to the generalized linear mixed effects model, the rate of visits in which a patient was treated with topical glucocorticoid drops was 39% lower in the first 6 months for the cohort who received CC compared to the cohort who received TC (rate ratio [RR] 0.61; 95% CI 0.44–0.84, $P < 0.01$) (data not shown). The rate of glucocorticoid reception per appointment was 96% lower for the cohort who received CC within the first year (RR 0.04; 95% CI 0.03–0.05, $P < 0.01$).

Development of complications. The odds ratio (OR) of developing band keratopathy, while controlling for length of time in the study, was 68% lower for the cohort who received CC (OR 0.32; $P = 0.02$) (Table 3), in which the probability of band keratopathy in the cohort who received CC and the cohort who received TC was 0.06 and 0.18, respectively. When controlling for time in study, there were no statistically significant difference in the odds risk of developing cataracts between the care models. Finally, there was no statistically significant difference for high pressure and overall complication rate (Table 3).

Episodes of uveitis care. Rates of total visits, including to the CC clinic, rheumatologists, and/or ophthalmologists, were 67% lower for patients receiving CC when controlling for the total complications the patient experienced (RR 0.33; 95% CI 0.27–0.42; $P < 0.01$; results not shown). The rate of visits per year (when holding average total complications constant at 1.78) is 1.5 in the cohort who received CC and 4.5 in the cohort who received TC.

DISCUSSION

This study describes outcomes of patients in a large pediatric uveitis cohort, some of whom received TC and some of whom received care in a combined rheumatology and ophthalmology

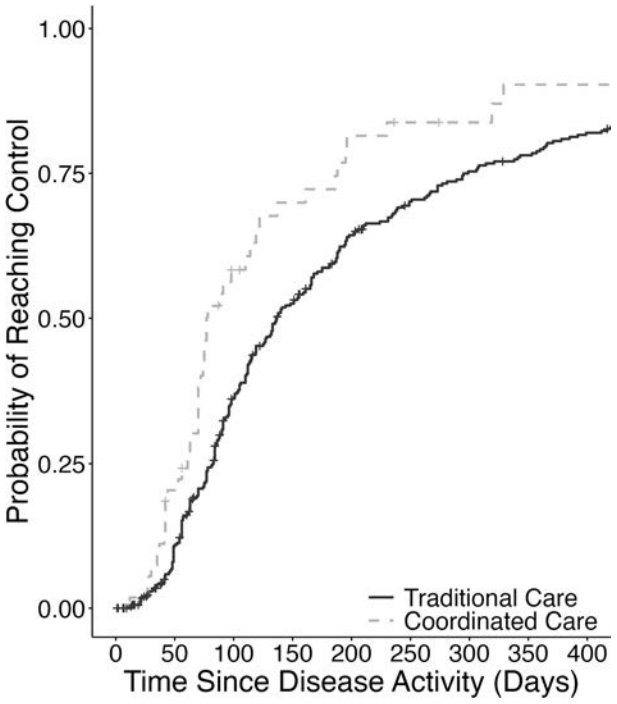


Figure 2. Kaplan–Meier estimates for time to disease control analysis. Log rank test $P < 0.01$.

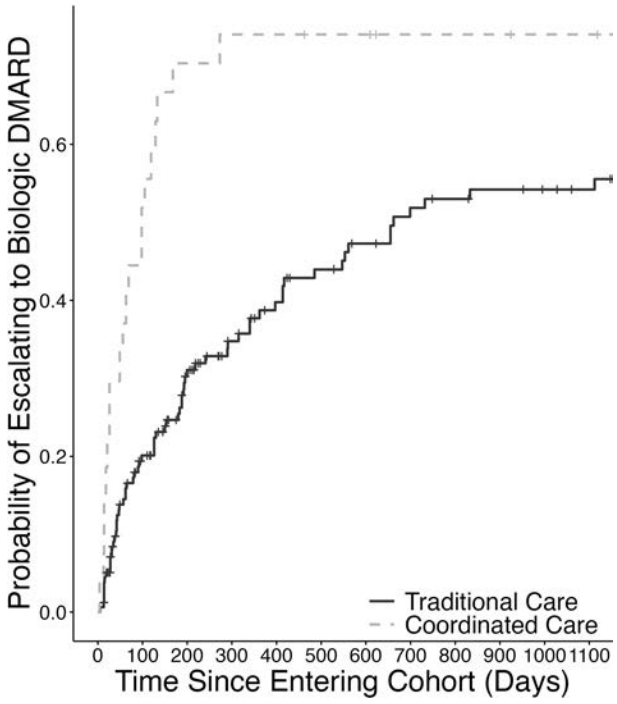


Figure 3. Kaplan–Meier estimates for time to biologic DMARD reception. Log rank test $P < 0.01$. DMARD, disease-modifying antirheumatic drug.

clinic. In our analyses, multidisciplinary care was associated with decreased time to disease control and decreased topical glucocorticoid reception, suggesting that communication between doctors optimized patient medication management and decreased visits. In recent literature, multidisciplinary clinics have been suggested as an intervention to address commonly identified impediments to uveitis care such as limited communication between ophthalmologists and rheumatologists.[22,26] To our knowledge, this is the first analysis that demonstrates improved outcomes from a pediatric multidisciplinary clinic. As mentioned in the introduction, uveitis can cause permanent damage and vision loss, so it is important to treat it rapidly and thoroughly. In this analysis, we demonstrate that under multidisciplinary care, patients had faster disease control both at the first episode of

Table 2. Multivariate models of time-to-disease-control analysis for all events of control*

Characteristic	HR	95% CI	*P* value
CC cohort	1.46	1.01–2.12	0.021
Max AC[a]	1	0.92–1.08	0.911
Time since diagnosis, days	0.98	0.95–1.01	0.136
JIA[b]	0.81	0.61–1.07	0.117
Time-varying models			
CC cohort	1.56	0.93–2.61	0.089
Max AC[a]	1.01	0.91–1.13	0.96
CC cohort	1.38	0.8–2.37	0.244
Max AC	0.97	0.85–1.1	0.923
Time since diagnosis, days	0.90	0.83–0.97	0.009
JIA	0.71	0.5–1.03	0.068
Receiving DMARD[c]	1.53	1.01–2.31	0.043
Receiving biologics[c]	1.93	1.09–3.4	0.024
Receiving dual therapy[c]	2.49	1.68–3.68	<0.001

* Nested models of variables for time-to-control analysis. Time-varying models include medications as a time-varying covariate. AC, anterior chamber; CC, coordinated care; CI, confidence interval; DMARD, disease-modifying antirheumatic drug; HR, hazard ratio; JIA, juvenile idiopathic arthritis.
[a] The time-varying variable of the maximum AC cell count was measured in either eye.
[b] All subtypes of JIA were included.
[c] Comparators were those who did not receive medication. Dual therapy consisted of treatment with both nonbiologic and biologic DMARDs.

Table 3. Total complications*

Characteristic	Ratio	*P* value
Total complications, RR (95% CI)		
Cohort who received CC	0.86 (0.59–1.24)	0.41
Maximum AC[a]	0.91 (0.89–0.9722)	<0.001
Elevated pressure, RR (95% CI)		
Cohort who received CC	1.22 (0.86–1.73)	0.27
Maximum AC	0.93 (0.88–0.99)	0.02
Band keratopathy, OR (95% CI)		
Cohort who received CC	0.32 (0.12–0.81)	0.017
Total observation time, y	1.1 (1.04–1.17)	0.001
Cataracts, OR (95% CI)		
Cohort who received CC	0.31 (0.07–1.45)	0.138
Total observation time, y	1.12 (1.01–1.25)	0.027

* RRs of total count of complications and elevated pressure were calculated. ORs of band keratopathy and cataracts were also calculated. AC, anterior chamber; CC, coordinated care; CI, confidence interval; OR, odds ratio; RR, rate ratio.
[a] The maximum AC cell count in either eye at the baseline visit was used.

disease activity documented in the registry and cumulatively over all episodes of disease reactivation.

Our data suggest that this was due in part to the patients who received CC starting treatment with biologics earlier than those who received TC. The adjusted models for the time-to-disease-control analysis demonstrate that biologic DMARDs were the most important variable in the model. This suggests the greatest impact of multidisciplinary care was the acceleration of initiation of a biologic when needed. Having uveitis associated with JIA was another predictor of time to control, but the multidisciplinary cohort had fewer patients with JIA (27.8% vs 56.6%). It is not currently possible to disentangle whether patients with JIA have worse AU activity because they are typically in TC or whether the CC clinic performed better due to their lower rates of patients with JIA.

Another factor in achieving disease control was time receiving topical glucocorticoids, and the extent of treatment of topical glucocorticoids are part of the definition of control (two or fewer drops per day). Both the rate at which patients were treated with topical glucocorticoids in the first 6 months and the rate of treatment with glucocorticoids per appointment at 12 months were significantly lower in the cohort who received CC. Topical glucocorticoids were also a driver of complications in patients with uveitis, such as elevated ocular pressure and cataracts.[8,27] The odds of developing band keratopathy was 68% lower while receiving multidisciplinary care versus TC. Conversely, there were no significant difference in the odds of developing high pressure or cataracts. A potential explanation for this is that some cataracts were not vision threatening, whereas others were denser. We did not distinguish between these in our single variable of cataracts. It is possible that patients who received CC had more mild cataracts which did not progress, although the absolute odds were not changed.

The burden of disease encompasses a myriad of factors, including medication treatment and disease damage (delineated above) as well as time seeking medical care/away from school-work-home life.[17,28,29] Not only were the former decreased, but the rate of visits was also 64% lower while receiving multidisciplinary care. Our data did not allow us to explore other factors that contribute to burden of disease, including, but not restricted to, financial and psychosocial impact on patients and their families.[29]

There are several limitations of this study, including its cohort size and retrospective nature. Although the cohort was relatively large due to stringent inclusion criteria, only 45 patients were able to be included in the cohort who received CC. It is possible that certain types of patients might be preferentially referred to, or chose to participate in, the multidisciplinary clinic. Although some patients were referred by their ophthalmologist due to refractory disease, others sought out the clinic for several reasons, including the convenience that the model offered. Patients who received CC may have had either (or both) more severe or more recalcitrant disease. We examined this by assessing the impact of AC disease activity at cohort entry (SUN criteria)[20] and time since diagnosis as variables. Neither were statistically associated with time to control. The registry did not collect sociodemographic factors, so we could not analyze to what extent these factors play a role in disease control. An additional limitation of this study is that our results only relate to patients with AU. We could not assess disease control and outcomes in patients with intermediate, posterior, and panuveitis due to only having robust data on AC activity in the registry. Many patients with intermediate, posterior, and panuveitis receive care in the multidisciplinary clinic; consequently, this limited our cohort size.

It is possible that the providers' uveitis expertise impacted outcomes. We did not record individual providers in the abstracted data. Although ophthalmology and rheumatology providers in the CC clinic were fixed, these same providers also see patients who received TC. Some patients who received TC benefited from their expertise as well. In future work, we hope to include individual providers as covariates to evaluate whether we are overestimating the effect of care model versus individual physician.

We had limited follow-up data on patients who received third-line agents (eg, abatacept, tocilizumab, or others) to include them in the study population, so we could not assess the impact of care model on their treatment. We did not evaluate all possible complications, and in future work, we may examine outcomes such as glaucoma and best corrected visual acuity as an outcome.

In conclusion, although CC has been recommended in guidelines and uveitis care reviews, its effects on outcomes for patients with pediatric uveitis have never been examined. The main strength of our study is that it is the first to examine the impact of multidisciplinary care for patients with pediatric uveitis. We demonstrate that multidisciplinary care is associated with improved uveitis outcomes such as time to disease control, at least in part by minimizing glucocorticoid treatment and accelerating the uptake of receiving biologics.

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 Lavallee 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/Helsinki Declaration requirements.

REFERENCES

1. Ravelli A, Martini A. Juvenile idiopathic arthritis. Lancet 2007; 369(9563):767–778.

2. Chang MH, Shantha JG, Fondriest JJ, et al. Uveitis in children and adolescents. Rheum Dis Clin North Am. 2021;47(4):619–641.
3. Agrawal RV, Murthy S, Sangwan V, et al. Current approach in diagnosis and management of anterior uveitis. Indian J Ophthalmol 2010; 58(1):11–19.
4. Maleki A, Anesi SD, Look-Why S, et al. Pediatric uveitis: a comprehensive review. Surv Ophthalmol 2022;67(2):510–529.
5. Thorne JE, Woreta FA, Dunn JP, et al. Risk of cataract development among children with juvenile idiopathic arthritis-related uveitis treated with topical corticosteroids. Ophthalmology 2010;117(7):1436–1441.
6. Angeles-Han ST, Ringold S, Beukelman T, et al. 2019 American College of Rheumatology/Arthritis Foundation Guideline for the screening, monitoring, and treatment of juvenile idiopathic arthritis-associated uveitis. Arthritis Care Res (Hoboken) 2019;71(6):703–716.
7. Constantin T, Foeldvari I, Anton J, et al. Consensus-based recommendations for the management of uveitis associated with juvenile idiopathic arthritis: the SHARE initiative. Ann Rheum Dis 2018;77(8): 1107–1117.
8. Rosenberg KD, Feuer WJ, Davis JL. Ocular complications of pediatric uveitis. Ophthalmology 2004;111(12):2299–2306.
9. Thorne JE, Woreta F, Kedhar SR, et al. Juvenile idiopathic arthritis-associated uveitis: incidence of ocular complications and visual acuity loss. Am J Ophthalmol 2007;143(5):840–846.
10. Gregory AC II, Kempen JH, Daniel E, et al; Systemic Immunosuppressive Therapy for Eye Diseases Cohort Study Research Group. Risk factors for loss of visual acuity among patients with uveitis associated with juvenile idiopathic arthritis: the Systemic Immunosuppressive Therapy for Eye Diseases Study. Ophthalmology 2013;120(1): 186–192.
11. Dick AD, Rosenbaum JT, Al-Dhibi HA, et al; Fundamentals of Care for Uveitis International Consensus Group. Guidance on noncorticosteroid systemic immunomodulatory therapy in noninfectious uveitis: Fundamentals of Care for Uveitis (FOCUS) Initiative. Ophthalmology 2018; 125(5):757–773.
12. Angeles-Han ST, Lo MS, Henderson LA, et al; Juvenile Idiopathic Arthritis Disease-Specific and Uveitis Subcommittee of the Childhood Arthritis Rheumatology and Research Alliance. Childhood Arthritis and Rheumatology Research Alliance consensus treatment plans for juvenile idiopathic arthritis-associated and idiopathic chronic anterior uveitis. Arthritis Care Res (Hoboken) 2019;71(4):482–491.
13. Ballenger LR, Ardoin SP, Driest KD. Assessing barriers to uveitis screening in patients with juvenile idiopathic arthritis through semi-structured interviews. Pediatr Qual Saf 2018;3(3):e084.
14. Soleymani T, Reddy SM, Cohen JM, et al. Early recognition and treatment heralds optimal outcomes: the benefits of combined rheumatology-dermatology clinics and integrative care of psoriasis and psoriatic arthritis patients. Curr Rheumatol Rep 2017;20(1):1.
15. Komenda P, Levin A. Analysis of cardiovascular disease and kidney outcomes in multidisciplinary chronic kidney disease clinics: complex disease requires complex care models. Curr Opin Nephrol Hypertens 2006;15(1):61–66.
16. Parker DM, Angeles-Han ST, Stanton AL, et al. Chronic anterior uveitis in children: psychosocial challenges for patients and their families. Am J Ophthalmol. 2018;191:xvi–xxiv.
17. la Mora A, Aurrecoechea E, del Corral AD, et al. Rheumatology-ophthalmology collaborative uveitis units may improve the diagnostic and therapeutic approach of this pathology: experience from a uveitis unit in a secondary Spanish hospital. J Arthritis 2016;5(5):1000215.
18. Venkatesh P, Macrae M, Fleck BW. Lothian combined paediatric ophthalmology and rheumatology service. Br J Ophthalmol 2006;90(12): 1549–1550.
19. Van Bentum RE, Van den Berg JM, Wolf SE, et al. Multidisciplinary management of auto-immune ocular diseases in adult patients by ophthalmologists and rheumatologists. Acta Ophthalmol 2020;99(2): e164–e170.
20. Jabs DA, Nussenblatt RB, Rosenbaum JT; Standardization of Uveitis Nomenclature (SUN) Working Group. Standardization of Uveitis Nomenclature for reporting clinical data. Results of the First International Workshop. Am J Ophthalmol 2005;140(3):509–516.
21. Leal I, Romão VC, Mano S, et al. A non-infectious uveitis multidisciplinary clinic in a tertiary referral center: clinical impact and added value. J Multidiscip Healthc 2021;14:695–704.
22. Rosenbaum JT, Dick AD. The eyes have it: a rheumatologist's view of uveitis. Arthritis Rheumatol 2018;70(10):1533–1543.
23. Lerman MA, Burnham JM, Chang PY, et al. Response of pediatric uveitis to tumor necrosis factor-α inhibitors. J Rheumatol 2013;40(8): 1394–1403.
24. Petty RE, Southwood TR, Manners P, et al; International League of Associations for Rheumatology. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. J Rheumatol 2004;31(2):390–392.
25. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. 2023. Accessed June 22, 2020. https://www.R-project.org
26. Mueller CM, Ward L, O'Keefe GAD. Health literacy, medication adherence, and quality of life in patients with uveitis. Ocul Immunol Inflamm 2022;30(5):1060–1067.
27. Henderson LA, Zurakowski D, Angeles-Han ST, et al. Medication use in juvenile uveitis patients enrolled in the Childhood Arthritis and Rheumatology Research Alliance Registry. Pediatr Rheumatol Online J 2016;14(1):9.
28. Cimino L, Neri P, Miserocchi E, et al. Non-infectious uveitis burden on quality of life and work impairment assessed through different psychometric questionnaires. Eur J Ophthalmol 2022;32(4):2282–2290.
29. McDonald J, Cassedy A, Altaye M, et al. Comprehensive assessment of quality of life, functioning, and mental health in children with juvenile idiopathic arthritis and noninfectious uveitis. Arthritis Care Res (Hoboken) 2022;74(8):1311–1320.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 77, No. 3, March 2025, pp 330–336
DOI 10.1002/acr.25203

BRIEF REPORT

Bringing Reproductive Health Guidelines Into Fellowship Training: A National Survey of Adult and Pediatric Rheumatology Fellows and Program Directors

Selene Rubino,[1] Michael J. Battistone,[2] Stacy P. Ardoin,[3] Elise D. Berlan,[3] Kristine Carandang,[4] Kate Chiseri,[5] Arthur Kavanaugh,[6] Whitney White,[7] Kelly A. Wise,[3] Andrew L. Wong,[8] Bethany Marston,[9] and Megan E. B. Clowse[10]

Objective. This study seeks to assess rheumatology fellows' (RFs') and program directors' (PDs') interests in different educational tools and methods and to facilitate curriculum development for reproductive health related to rheumatic disease.

Methods. Constructs were conceptualized in four dimensions: 1) RF and PD confidence in their current curriculum relating to the American College of Rheumatology (ACR) Reproductive Health Guidelines (RHGs), 2) personal interest in this topic, 3) opinions of the importance of this topic, and 4) interest in a range of learning materials and educational experiences. The final survey was distributed to 753 RFs and 179 PDs in the United States using the ACR Committee on Training and Workforce email list.

Results. Response rates were 13% (n = 98) for RFs and 25% (n = 44) for PDs. Both groups indicated more interest in the topic than confidence in their curriculum and rated summary sheets, question banks, didactics, and online modules higher than nine other educational tools or methods. Despite interest in the topic, 38% of RF respondents and 24% of PD respondents were unaware of the recently published ACR RHGs.

Conclusion. RFs and PDs consider reproductive health very important and report high personal interest in this topic. In contrast, both groups indicated lower confidence in current curricula, and substantial proportions of both groups were unaware of recently published guidelines. RFs' and PDs' interests in specific educational modalities are aligned. Curriculum development efforts should prioritize summary sheets, question banks, didactics, and online modules. Efforts are needed to address the educational needs of practicing rheumatologists and other professionals caring for patients with rheumatic disease.

INTRODUCTION

Health care concerns relating to reproduction are significant for many patients with rheumatic and musculoskeletal disease (RMD). In response to these concerns, the American College of Rheumatology (ACR) convened a meeting of a broad range of stakeholders that included patient advocates, representatives of the US National Institutes of Health and Food and Federal Drug Administration (FDA), scientific researchers, and an interprofessional and multidisciplinary group of clinicians (1). This Reproductive Health Summit (RHS) emphasized the need for better management of RMD before, during, and after pregnancy with the goal of protecting both birthing parent and fetus from destructive inflammatory processes while minimizing the risk of adverse effects of therapies. In addition to recognizing that "the desire for healthy pregnancies among our patients is a very real and

[1]Selene Rubino, MD: University of Utah, Salt Lake City; [2]Michael J. Battistone, MD: University of Utah, George E. Wahlen Veterans Affairs, and Salt Lake City Health Care System, Salt Lake City, Utah; [3]Stacy P. Ardoin, MD, MSc, Elise D. Berlan, MD, MPH, Kelly A. Wise, PharmD, BCACP: Nationwide Children's Hospital, Columbus, Ohio; [4]Kristine Carandang, PhD, OTR/L: Independent, San Diego, California; [5]Kate Chiseri, MPH, MCHES: American College of Rheumatology, Atlanta, Georgia; [6]Arthur Kavanaugh, MD: University of California San Diego, La Jolla; [7]Whitney White, PharmD, BCPS, FASHP: Birmingham Veterans Affairs Medical Center, Birmingham, Alabama; [8]Andrew L. Wong, MD: University of California Los Angeles Medical Center, Sylmar; [9]Bethany Marston, MD: University of Rochester Medical Center, Rochester, New York; [10]Megan E. B. Clowse, MD, MPH: Duke University, Durham, North Carolina.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25203).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25203.

Address correspondence via email to Michael J. Battistone, MD, at michael.battistone@va.gov.

Submitted for publication March 3, 2023; accepted in revised form July 20, 2023.

SIGNIFICANCE & INNOVATIONS

- The 2020 American College of Rheumatology (ACR) established the first guidelines for reproductive health in the management of rheumatologic disease.
- The responding rheumatology fellows (RFs) and fellowship program directors (PDs) consider reproductive health to be very important in their work and report a high level of personal interest in this topic. Despite their interest, a substantial proportion of RFs (38%) and PDs (24%) were unaware of recently published ACR Reproductive Health Guidelines.
- RFs' and PDs' interests in specific educational modalities are aligned; local and national educational innovations should prioritize the development of summary sheets, question banks, didactics, and online modules.
- Further definition of educational gaps is needed across other aspects of the health professions' education and practice continuum. This includes practicing rheumatologists, obstetricians/gynecologists, women's health specialists, pharmacists, physician assistants, nurse practitioners, and other clinicians whose work is relevant to reproductive health concerns in rheumatic disease.

pressing issue that needs to be addressed," proceedings from the summit called for 1) providing more guidance to patients and prescribers through collection of data to inform updates to FDA product information, 2) creating new research paradigms—such as pregnancy and lactation registries—to facilitate construction of a robust evidence base, and 3) developing effective educational programs focusing on patient care in the intersection of reproductive health and RMD.

An important outcome of the summit was the creation of a sustained, multispecialty, interprofessional effort—the Reproductive Health Initiative—tasked with dissemination of clinical practice guidelines relating to contraception, assisted reproductive technologies, fertility preservation in the context of gonadotoxic therapy, use of peri- and postmenopausal hormone replacement therapy, pregnancy assessment and management, and medication use in patients with RMD. The ACR Reproductive Health Guidelines (RHGs), published in 2020, were the result of a five-year systematic effort that produced 12 ungraded "good practice" statements and 131 graded recommendations (2). The RHGs provide guidance for developing new educational programs and initiatives in the form of curriculum design.

The purpose of this study was two-fold: 1) to design and implement a national survey of rheumatology fellows (RFs) and fellowship program directors (PDs) to inform the development of fellowship curricula and programs relating to ACR RHGs, and 2) to examine evidence of the validity of these surveys drawn from the content and response process (Messick's framework) prior to national sampling (3).

MATERIALS AND METHODS

Content. Using the approach outlined by Artino and Gehlbach, constructs of interest were defined through literature review and reflective critique involving eight reproductive health initiative members, facilitated by a former PD (4). The constructs were defined as "interest in curriculum" and considered separately for RFs and PDs, which led to the development of two separate surveys. Constructs were conceptualized in four dimensions: 1) *confidence* in the current curriculum relating to ACR RHGs; 2) *personal interest* in this topic; 3) *opinions of the importance* of this topic; and 4) *interest in a range of learning materials and educational experiences*.

Response process. Survey items to measure these constructs were developed following the approach outlined by Gehlbach and Artino (4,5). Effort was taken to use construct-specific response options communicated in positive terms, with a single focus, and to avoid reverse-scoring. Survey design emphasized visual consistency and coherence. Expert reviewers (five RFs and four PDs, none of whom had been involved in defining the constructs) evaluated survey items for relevance to the construct and clarity of communication. Descriptive statistical analyses of expert ratings were performed. The content validity index (CVI) for each item—an assessment of each item's clarity and relevance to the construct—was calculated by grading the relevance of each survey question on a five-point scale from 1 (not relevant) to 5 (very relevant). Individual survey responses were divided by 5 and averaged across respondents. A separate CVI was calculated for RFs and PDs. Items with a CVI of <0.7 were discarded. Cognitive interviewing clarified the mental model emerging through survey use. Survey participants' race, ethnicity, and gender were determined by their responses to fixed sets of categories using National Institutes of Health terminology, although an option to self-describe using an open-ended format was also included.

The initial surveys consisted of 27 items for RFs and 28 items for PDs. Expert review led to discarding 10 items for RFs and 12 items for PDs. Cognitive interviewing indicated that RFs considered questions relating to their level of interest in reproductive health as connected to a sense of ownership or responsibility for addressing these issues. Final versions of the surveys, consisting of 17 items for RFs and 16 items for PDs (Supplemental Materials 1 and 2) were presented to the ACR Committee on Training and Workforce (COTW) and distributed nationally via SurveyMonkey through the COTW email list to serve all 753 fellows and 179 PDs representing adult, pediatric, and combined internal medicine and pediatrics (med/peds) rheumatology fellowships. The response window was June 7–30, 2021. Survey response

data were analyzed with descriptive statistics using Excel (Microsoft).

This project was reviewed by the Institutional Review Board of the University of Utah and the Salt Lake City Veterans Affairs Health Care System. It was determined to be a quality improvement project that did not meet the definition of research involving human subjects and was exempt from further review.

RESULTS

Ninety-eight RFs and 44 PDs completed the survey (response rates of 13% and 25%, respectively). Demographic characteristics of respondents and fellowship program characteristics are shown in Table 1 (note that percentages may not sum to 100% because of rounding). A majority of respondents—71% of RFs and 62% of PDs—identified as cisgendered women. A majority of RFs and PDs were Caucasian/White (57% and 76%, respectively), with RFs being more representative of national demographics. RFs were surveyed from the adult (75%), pediatrics (22%), and med/peds specialties (3%). This distribution was proportionate to PD specializations.

Table 1. Demographics of survey respondents*

Demographics	RFs (n = 98) N (%)	PDs (n = 44) N (%)
Sex		
Female	70 (71.4)	28 (63.6)
Male	28 (28.6)	16 (36.4)
Race (multiple responses possible)		
Caucasian/White	55 (56.1)	33 (75)
Asian/Pacific Islander	15 (15.3)	8 (18.2)
Multiracial/biracial	4 (4.1)	0
Asian Indian	2 (2.0)	0
Hispanic	2	0
Middle Eastern	2	0
South Asian	2	0
African American/Black	1 (1.0)	1 (2.3)
American Indian/Alaska Native	1	0
Coptic	1	0
Latino	1	0
South-East Asian	1	0
Prefer not to answer/did not answer	11 (11.2)	2 (4.5)
Hispanic or Latino	9 (9)	2 (4.5)
Year of training		
First	42 (42.9)	-
Second	40 (40.8)	-
Third	15 (15.3)	-
Length of time as PD, years		
<1	-	7 (15.9)
1–5	-	18 (40.9)
6–10	-	7 (15.9)
11–15	-	6 (13.6)
>15	-	6 (13.6)
Number of fellows in program, median (range)	5 (1–12)	4 (0–10)
Specialty		
Adult	71 (75)	30 (68.1)
Pediatrics	21 (22)	10 (22.7)
Med/peds	3 (3)	1 (2.3)

* Med/peds = combined internal medicine and pediatrics; PDs = program directors; RFs = rheumatology fellows.

Overall, responses of RFs and PDs were similar. Of note, 24% of PDs and 38% of RFs responding to the survey were not aware that the ACR's RHGs had been published. Both groups considered knowledge in reproductive health to be "quite important" or "essential" in caring for patients, although they were only "moderately confident" that their fellowship curriculum in reproductive health would provide good preparation for exam certification or clinical practice (Figure 1). A majority (63%) of the 57 RFs who were aware of the RHGs had also used them, whereas 23% were aware of the guidelines but had not used them. When asked about their experiences teaching the guidelines, 56% (32 of 57) of RFs had participated in discussions about them but had not taught them, compared with 10% who had either led a discussion or had taught them in some other way. Of the remainder, 3% (2 of 57) of RFs were aware of the guidelines but had not read them, and 30% had read but not discussed them. PDs were also asked about their experiences teaching: 19% (6 of 32) were aware of the guidelines but had not read them, 25% had read but not discussed them, 28% had participated in a discussion but not as a leader or teacher, and 28% had either led a discussion of the guidelines or taught them in some other way.

Respondents reported their interest in various educational materials and experiences that might be developed as platforms for teaching the guidelines (Table 2). Both RFs and PDs provided their highest ratings for the same four items: summary sheets, question banks, online modules, and didactics.

DISCUSSION

Health care concerns relating to reproduction have particular significance for patients with RMD and the providers who care for them. The 2020 ACR RHGs are a valuable resource for integrating these concerns into individual patient care plans, and incorporating these recommendations into fellowship curricula will provide opportunities for RFs to learn to use them during subspecialty training. Our work highlights four important dimensions of educational need: 1) despite being interested in reproductive health, RFs and PDs do not feel confident in existing curricula, 2) many are unaware of the guidelines, 3) RFs' and PDs' interests in specific educational modalities are generally aligned, and 4) the interest is greatest for summary sheets, question banks, online modules, and didactics, suggesting that developing these tools should be a higher priority.

Although our survey did not investigate the effects of didactics on confidence scores, Canadian medical students offered supplemental rheumatology didactics were found to have an increase in confidence beyond that seen in other internal medicine subspecialties (6). Prior surveys have examined educational experiences in the context of other fellowships (7,8). One study of

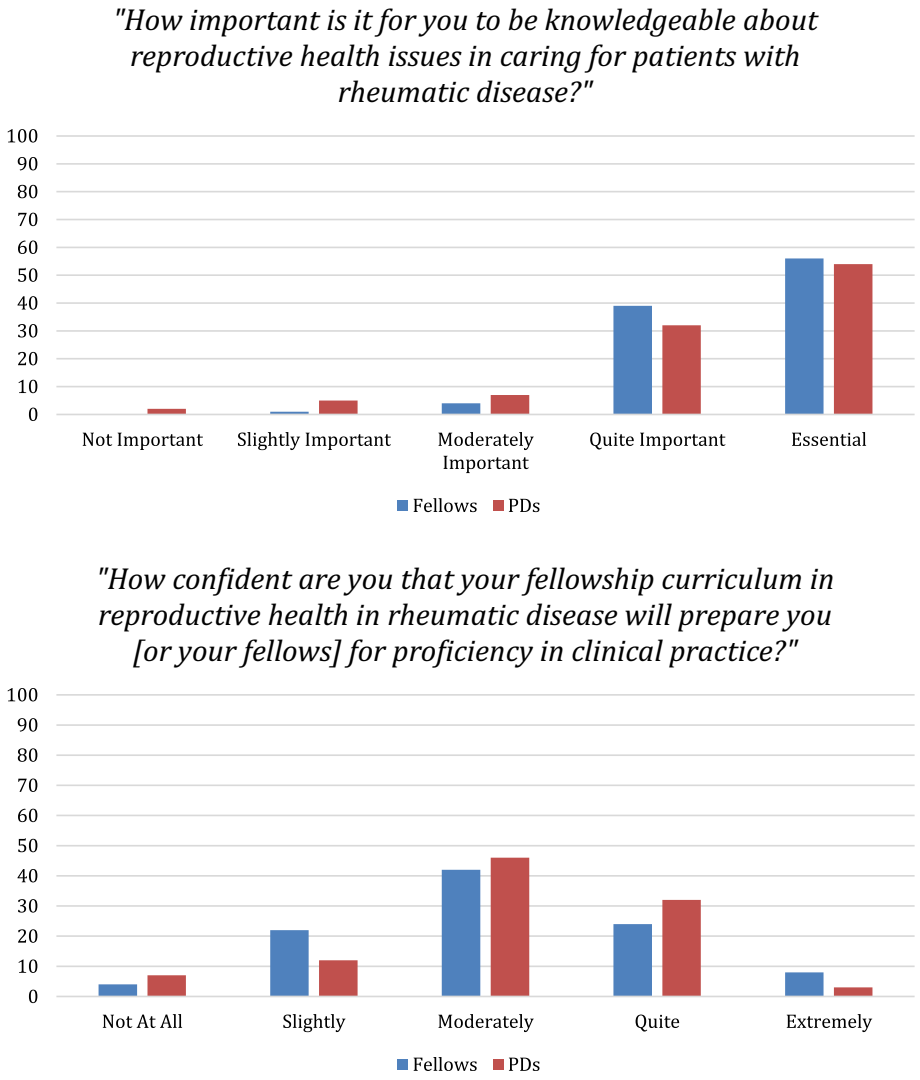


Figure 1. Percentage of responses in each category for two survey items. PDs = program directors.

pediatric hematology and oncology fellowship PDs found that, although 88% of respondents had formal education in fertility, only 30% formally taught sexual health (9). There were several barriers to developing this curriculum, including lack of experts and difficulty fitting it into existing fellowship program curricula.

Although our survey did not address specific challenges to implementation, this is likely an important factor; an earlier ACR survey assessing the need for RF-as-teacher programs found lack of time to be a major barrier to incorporating this curriculum. Only 55% of PDs agreed that fellows have time for teacher training, and only 40% had faculty with time to supervise such programs (10). Our experience with musculoskeletal ultrasound (MSK U/S) may also inform our approaches to developing and implementing new curricula for reproductive health. Although a formal curriculum for MSK U/S is nearly universally desired by rheumatology fellowship PDs, the majority of fellowship programs report lacking such a program because of inadequate time and number of trained faculty (11). It is notable that some educational tools favored by respondents to our survey, summary sheets, for example, are relatively less time-consuming and may be more feasible to implement. Finally, in accordance with the principles of curricular design in medical education literature, we implemented a targeted learner's assessment of RFs and PDs. Both were found to favor the development of concise online question banks or summary sheets. Interestingly, social media was the least favorite option. The creation and distribution of these tools would be a unique opportunity, both to standardize the rollout of new recommendations in rheumatology as well as to spread awareness of RHGs.

Development of program-specific clinical teams to spread knowledge and implementation of RHGs may be important locally, even if educational materials are available. A Cochrane review of five prior randomized controlled trials evaluating dissemination of educational materials specific to antibiotic stewardship demonstrated improvement in prescription patterns but with a high degree of heterogeneity (12). The reviewers recommend the

Table 2. Interest in types of learning materials and educational experiences*

	Not at all interested 1	Slightly interested 2	Moderately interested 3	Quite interested 4	Extremely interested 5	Mean (1–5)
RFs						
Summary sheet[a]	–	–	9 (11)	22 (26)	55 (64)	4.5
Question bank[a]	–	4 (4.7)	14 (16)	17 (20)	51 (59)	4.3
Didactics[a]	1 (1.2)	9 (11)	19 (22)	34 (40)	23 (27)	3.8
Online modules[a]	4 (4.7)	11 (13)	18 (21)	24 (28)	29 (34)	3.7
Visiting professors	3 (3.5)	8 (9.3)	31 (36)	22 (26)	22 (26)	3.6
Podcasts	6 (7)	12 (14)	23 (27)	18 (21)	27 (31)	3.6
Journal articles	2 (2.3)	11 (13)	30 (35)	23 (27)	20 (23)	3.6
Journal club	6 (7)	13 (15)	33 (39)	17 (20)	16 (19)	3.3
Social media	22 (26)	18 (21)	22 (26)	8 (9.3)	16 (19)	2.7
PDs						
Question bank[a]	1 (2.5)	1 (2.5)	3 (7.5)	12 (30)	23 (58)	4.4
Summary sheet[a]	1 (2.6)	1 (2.6)	5 (13)	10 (26)	21 (55)	4.3
Online modules[a]	–	4 (10)	10 (25)	8 (20)	18 (45)	4.0
Didactics[a]	1 (2.6)	3 (7.7)	8 (21)	12 (31)	15 (38)	4.0
Journal articles	1 (2.6)	3 (7.7)	10 (26)	12 (31)	13 (33)	3.9
Visiting professors	3 (7.5)	10 (25)	4 (10)	12 (30)	11 (28)	3.5
Journal club	5 (13)	5 (13)	8 (20)	12 (30)	10 (25)	3.4
Podcasts	4 (10)	5 (13)	15 (38)	7 (18)	9 (23)	3.3
Social media	12 (31)	10 (26)	9 (23)	4 (10)	4 (10)	2.4
Slide deck[b]	1 (2.5)	2 (5)	6 (15)	13 (33)	18 (45)	4.1
EMR alert[b]	4 (10)	7 (18)	7 (18)	9 (23)	12 (31)	3.5
OSCE (formative)[b]	9 (22)	4 (9.8)	9 (22)	8 (20)	11 (27)	3.2
Patient partners[b]	7 (18)	8 (21)	9 (23)	6 (15)	9 (23)	3.1

* Values are the number of RFs or PDs that indicated each level of interest (%) unless indicated otherwise. EMR = electronic medical record; OSCE = objective structured clinical examination; PDs = program directors; RFs = rheumatology fellows.
[a] These four items in the respective category were highest rated by both RFs and PDs.
[b] These four items were discarded from the RF survey because the content validity index from the expert review was <0.7.

inclusion of site-specific providers into the design and subsequent measurement of an intervention. Prior studies assessing implementation of new rheumatology curricula also suggest that the availability of trained faculty is an important factor (10).

The major limitation of our survey project is the relatively low participation—only 25% of PDs and 13% of RFs completed the questionnaire. Multiple factors are known to influence response rates (13). These include 1) format (multiple methods of delivery increase rates—we used only an online survey); 2) acknowledgment (it is possible that some PDs and RFs may not have recognized they had received the survey); 3) incentives (monetary rewards have been shown to increase response rates—we did not offer any form of compensation); 4) length (our process of construction reduced the length of the survey by over 20%, though this may still have been prohibitive for some recipients); 5) timing (the end of the academic year involves transitions for senior fellows—though the majority of responses came from second-year and third-year fellows); and 6) salience (relevance of the topic to the individual respondent). One of the most interesting results was that 24% of PDs and 38% of RFs were unaware of published guidelines. In considering the possible effects of nonresponse bias, we would expect nonresponders to be less aware of these guidelines; this suggests that our survey results may have underestimated this knowledge gap (13). Given that this survey was distributed to academic rheumatologists with formal educational responsibilities, the number of practicing physicians who are not aware of the guidelines is likely higher. Finally, the PDs' and RFs' versions of the surveys differed slightly in some of the options for preferred learning modalities because of differences in the CVI scores for these items. This approach followed best practice recommendations for survey design, though it also raises interesting questions. Do PDs and RFs actually have different perspectives or opinions regarding educational tools and methods? If so, is the variance meaningful at the level of individual programs? How are these different perspectives negotiated successfully? These questions were not addressed in our study, though they might provide interesting work for future scholars.

Our study has several strengths. We used a systematic, rigorous process to define constructs of interest, developed items and response anchors in accordance with current best practices in survey design, and conducted expert validation and cognitive pretesting prior to distribution. In partnership with the ACR COTW, the survey was disseminated nationally to all RFs and PDs, with response data captured and stored electronically.

There was alignment between RFs and PDs regarding their interest in specific educational tools or methods to teach the ACR RHGs because the same four modalities received the highest mean ratings from both groups of survey respondents. Although the specific needs and resources of individual fellowships might encourage use of any of the educational approaches we presented in our survey, there may be broad appeal across multiple programs for those with higher mean ratings. This suggests that the following four areas might be prioritized for development:

First summary sheets encompass a broad range of materials (eg, tables, checklists, algorithms, etc.), and are often created for use in patient care settings. The high ratings from RFs and PDs likely reflect shared perceptions of the value of clinical training experiences and other aspects of workplace learning in developing clinical competence. In considering why summary sheets were found to be most helpful, it may be that both learners and teachers see them as efficient ways of communicating a message—one that can initially be received in a classroom setting and then reinforced through workplace learning in the clinic. Though summary sheets are by definition too limited to contain the detailed concepts and conversations underlying a specific guideline, they provide a carefully constructed "takeaway" that can more easily be implemented in a busy clinical practice setting. If these are to be disseminated, peer review or expert consultation could help ensure that the information presented in a specific summary sheet communicates the guidelines accurately. Discussion guides that meet this description can be found at www.LupusPregnancy.org and www.ReproRheum.Duke.edu, which are both free resources for rheumatologists.

Second, question banks have long been used by students and trainees in formative knowledge assessments, particularly when preparing for board certification examinations and other high-stakes testing. An educational initiative to develop question banks could involve either the expansion of existing assets such as the Continuing Assessment Review Evaluation program, the construction of new resources, or a combination of these approaches. In addition, PDs may explore and share new ways of integrating these questions into their fellowship experiences. Although creating a pool of high-quality questions would likely require resources to support best practices in writing, reviewing, and evaluating questions, the high degree of interest from RFs and PDs alike suggests that this as a powerful tool in developing trainees' knowledge of the guidelines.

Third, although it has become popular to castigate the lecture as an educational practice, enthusiasm for didactics as a tool for teaching and learning the ACR guidelines is clearly shared by RFs and PDs. This may reflect the unique potential of lectures—which are formal, spoken, and social events—when they are prepared and delivered effectively (14). Excellence in didactics could be fostered locally by supporting faculty who are effective in this format and nationally through speakers' boards.

Finally, online modules offer a tremendous range of options for educational content and learning activities. Although these are often self-directed and individual, online modules can be integrated as elements within a larger educational experience with multiple learners and could be conducted either synchronously or asynchronously. Resources would be needed to ensure technical expertise is available not only for the creation but also the maintenance of these modules.

As an assessment of educational needs, this work serves as an important link between the development of the ACR RHGs and the creation of new educational experiences for RFs to bring these "off the shelf and into the clinic." In addition to encouraging educational leaders at individual or collaborative fellowships to build on our work by creating new materials and programs, we suggest that projects in reproductive health in rheumatic disease might provide an excellent focus when considering an application for the Clinician Scholar Educator Award, sponsored by the ACR Rheumatology Research Foundation; the Community Practice Innovation Award might also be a mechanism to support continuing professional development initiatives. Because many curricular innovations are built to fit the needs and purpose of single programs or small consortia, a clear understanding of the educational principles (and the educational philosophy underpinning these principles) supporting the specific techniques of these innovations will be critical when considering how they might be disseminated to the new context of a different program (15). We hope that this brief report will enhance PDs' and other leaders' assessment of educational need, inform the creation of new materials and experiences for our fellows, and ultimately lead to the delivery of better care to our patients, specifically in regard to reproductive health.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Battistone, Ardoin, Berlan, Carandang, Kavanaugh, White, Wise, Wong, Clowse.

Acquisition of data. Battistone, Chiseri.

Analysis and interpretation of data. Rubino, Battistone, Berlan, Kavanaugh, Wong, Marston, Clowse.

REFERENCES

1. Kavanaugh A, Cush JJ, Ahmed MS, et al. Proceedings from the American College of Rheumatology Reproductive Health Summit: the management of fertility, pregnancy, and lactation in women with autoimmune and systemic inflammatory diseases. Arthritis Care Res (Hoboken) 2015;67:313–25.
2. Sammaritano LR, Bermas BL, Chakravarty EE, et al. 2020 American College of Rheumatology guideline for the management of reproductive health in rheumatic and musculoskeletal diseases. Arthritis Rheumatol 2020;72:529–56.
3. Messick S. Validity. In: Linn RL, editor. Educational measurement. 3rd ed. New York (NY): American Council on Education/Macmillan; 1989. p. 13–103.
4. Artino AR, La Rochelle JS, Dezee KJ, et al. Developing questionnaires for educational research: AMEE Guide No. 87. Med Teach 2014;36: 463–74.
5. Gehlbach H, Artino AR. The survey checklist (Manifesto). Acad Med 2018;93:360–6.
6. Katz SJ, Oswald AE. How confident are internal medicine residents in rheumatology versus other common internal medicine clinical skills: an issue of training time or exposure? Clin Rheumatol 2011;30:1081–93.
7. Nageswara Rao AA, Warad DM, Weaver AL, et al. Cross-cultural medical care training and education: a national survey of pediatric hematology/oncology fellows-in-training and fellowship program directors. J Cancer Educ 2019;34:478–87.
8. Mehta AK, May N, Verga S, et al. Palliative care education in U.S. adult neuro-oncology fellowship programs. J Neurooncol 2018;140: 377–83.
9. Frederick NN, Fine E, Michaud A, et al. Pediatric hematology and oncology fellow education in sexual and reproductive health: a survey of fellowship program directors in the United States. Pediatr Blood Cancer 2020;67:e28245.
10. Reid P, Miloslavsky EM, Dua A. Teacher training for rheumatology fellows: a national needs assessment of fellows and program directors. Clin Rheumatol 2020;39:673–80.
11. Torralba KD, Cannella AC, Kissin EY, et al. Musculoskeletal ultrasound instruction in adult rheumatology fellowship programs. Arthritis Care Res (Hoboken) 2020;72:859–70.
12. Davey P, Brown E, Charani E, et al. Interventions to improve antibiotic prescribing practices for hospital inpatients [review]. Cochrane Database Syst Rev 2013;4:CD003543.
13. Phillips AW, Reddy S, Durning SJ. Improving response rates and evaluating nonresponse bias in surveys: AMEE Guide No. 102. Med Teach 2016;38:217–28.
14. Charlton BG. Lectures are such an effective teaching method because they exploit evolved human psychology to improve learning. Med Hypotheses 2006;67:1261–5.
15. Cianciolo AT, Regehr G. Learning theory and educational intervention: producing meaningful evidence of impact through layered analysis. Acad Med 2019;94:789–94.

Arthritis Care & Research
Vol. 77, No. 3, March 2025, pp 337–348
DOI 10.1002/acr.25184

American College of Rheumatology
Empowering Rheumatology Professionals

Multimorbidity Patterns and Rheumatoid Arthritis Disease Outcomes: Findings From a Multicenter, Prospective Cohort

Sarah Dutt,[1] Punyasha Roul,[1] Yangyuna Yang,[1] Tate M. Johnson,[1] Kaleb Michaud,[2] Brian Sauer,[3] Grant W. Cannon,[3] Joshua F. Baker,[4] Jeffrey R. Curtis,[5] Ted R. Mikuls,[1] and Bryant R. England[1]

Objective. To determine whether unique multimorbidity patterns are associated with long-term rheumatoid arthritis (RA) disease severity.

Methods. We conducted a cohort study within the Veterans Affairs Rheumatoid Arthritis registry. We applied previously derived multimorbidity patterns based on the presence of diagnostic codes for relevant conditions prior to enrollment using linked administrative data. Disease activity and functional status were assessed longitudinally up to 5 years after enrollment. The association of multimorbidity patterns with disease activity and functional status were assessed using generalized estimating equations models adjusting for relevant confounders.

Results. We studied 2,956 participants, of which 88.2% were male, 76.9% reported white race, and 79.3% had a smoking history. Mental health and substance abuse (β 0.12 [95% confidence interval {CI} 0.00, 0.23]), cardiovascular (β 0.25 [95% CI 0.12, 0.38]), and chronic pain (β 0.21 [95% CI 0.11, 0.31]) multimorbidity were associated with higher Disease Activity Score in 28 joints (DAS28) scores. Mental health and substance abuse (β 0.09 [0.03, 0.15]), cardiovascular (β 0.11 [95% CI 0.04, 0.17]), and chronic pain multimorbidity (β 0.15 [95% CI 0.10, 0.20]) were also associated with higher Multidimensional Health Assessment Questionnaire (MDHAQ) scores. The metabolic pattern of multimorbidity was not associated with DAS28 or MDHAQ. The number of multimorbidity patterns present was highly associated with DAS28 and MDHAQ (P trend < 0.001), and patients with all four multimorbidity patterns had the highest DAS28 (β 0.59 [95% CI 0.36, 0.83]) and MDHAQ (β 0.27 [95% CI 0.16, 0.39]) scores.

Conclusion. Mental health and substance abuse, chronic pain, and cardiovascular multimorbidity patterns are associated with increased RA disease activity and poorer functional status. Identifying and addressing these multimorbidity patterns may facilitate achieving RA treatment targets.

INTRODUCTION

In addition to joint manifestations, rheumatoid arthritis (RA) predisposes patients to the development of other chronic conditions such as heart disease, lung diseases, and osteoporosis, among others.[1] The development of multiple chronic conditions, often termed multimorbidity, affects the majority of patients with RA.[2,3] Certain shared risk factors, such as smoking and obesity, can predispose to both RA and other chronic conditions that contribute to the burden of multimorbidity. Multimorbidity may also result from the systemic inflammatory responses accompanying RA, which are known to adversely

The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs or the US government.

[1]Sarah Dutt, Punyasha Roul, MS, Yangyuna Yang, MBBS, PhD, Tate M. Johnson, MD, Ted R. Mikuls, MD, MSPH, Bryant R. England, MD, PhD: VA Nebraska-Western Iowa Health Care System and University of Nebraska Medical Center, Omaha; [2]Kaleb Michaud, PhD: University of Nebraska Medical Center, Omaha, and FORWARD–the National Data Bank for Rheumatic Disease, Wichita, Kansas; [3]Brian Sauer, PhD, Grant W. Cannon, MD: VA Salt Lake City and University of Utah, Salt Lake City; [4]Joshua F. Baker, MD, MSCE: Corporal Michael J. Crescenz VA Medical Center and the University of Pennsylvania, Philadelphia; [5]Jeffrey R. Curtis, MD, MPH, MS: University of Alabama at Birmingham, Birmingham.

Dr. Johnson's work was supported by a Rheumatology Research Foundation Scientist Development award. Dr. Baker's work was supported by a VA Merit Award (grant CX-001703). Dr. Curtis's work was supported by the NIH/National Institute of Arthritis and Musculoskeletal and Skin Diseases (grant P30-AR-072583) and the Patient Centered Outcomes Research Institute. Dr. Mikuls's work was supported by a VA Merit Award (grant BX-004600), the Department of Defense (grant PR-200793) and the NIH/National Institute of General Medical Sciences (grant U54-GM-115458). Dr. England's work was supported by a Rheumatology Research Foundation Scientist Development award, the Great Plains IDeA-CTR from the NIH/National Institute of General Medical Sciences (grant U54-GM-115458), and the VA Clinical Science Research & Development (grant IK2-CX-002203).

Additional supplementary information cited in this article can be found online in the Supporting Information section (http://onlinelibrary.wiley.com/doi/10.1002/acr.25184).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25184.

Address correspondence via email to Bryant R. England, MD, PhD, at Bryant.england@unmc.edu.

Submitted for publication April 17, 2023; accepted in revised form June 27, 2023.

SIGNIFICANCE & INNOVATIONS

- Most people with rheumatoid arthritis (RA) are multimorbid, experiencing multiple chronic conditions.
- Multimorbidity patterns are novel measures of multimorbidity occurring in people with RA, but their associations with RA-related outcomes are unknown.
- We characterized cross-sectional and longitudinal associations between different multimorbidity patterns with RA disease activity and functional status in a multicenter, prospective RA cohort.
- Mental health and substance abuse, chronic pain, and cardiovascular multimorbidity patterns are associated with increased RA disease activity and poorer functional status.

impact several organ systems including the heart and lungs.[1,4–7] Although multimorbidity is recognized to portend poor survival and reduced quality of life in RA,[8] there has been limited research investigating how multimorbidity may be associated RA severity and the RA disease course.

In some prior studies, multimorbid patients with RA were less likely to receive biologic disease-modifying antirheumatic drugs (DMARDs) or had biologic DMARD initiation postponed.[9,10] In contrast, in a large, real-world RA cohort study with more stringent eligibility criteria requiring multiple visits with moderate or high disease activity, there were no differences in the initiation of new DMARDs, including biologics, based on multimorbidity burden.[11] Although there are conflicting findings regarding whether multimorbid patients with RA are treated less aggressively, perhaps depending on the health care environment, practice setting, and the predilections of the treating rheumatologist, multimorbid patients do consistently appear to be less likely to achieve the RA treatment targets of remission or low disease activity after beginning new DMARDs.[4,11]

A limitation of the aforementioned studies is that they typically assessed multimorbidity using a simple count of chronic conditions, which does not capture the interconnectedness of chronic conditions core to the concept of multimorbidity.[12,13] Previously, we have used independent, large, real-world data sources and machine learning approaches to derive novel multimorbidity patterns among patients with RA.[12] However, it remains unknown whether these multimorbidity patterns are associated with long-term RA-related outcomes such as RA disease activity and functional status. Therefore, the objective of this study was to determine how unique multimorbidity patterns are associated with the long-term disease course in RA. As an extension of our prior work, we hypothesized that multimorbid patients would have increased RA disease activity and poorer functional status and that these outcomes would differ across unique multimorbidity patterns. Additionally, we hypothesized that patients with a greater number of multimorbidity patterns would demonstrate more severe disease activity and functional status trajectories.

PATIENTS AND METHODS

Study design and data sources. We performed a cohort study within the Veterans Affairs (VA) Rheumatoid Arthritis (VARA) registry. The VARA registry, initiated in 2003, is a multicenter, prospective cohort of US Veterans with RA diagnosed by a rheumatologist and fulfilling the 1987 American College of Rheumatology (ACR) classification criteria.[14] All participants provided written informed consent, and each site received Institutional Review Board (IRB) approval. This study was approved by the VA Nebraska-Western Iowa Health Care System IRB (no. 1576193).

The VARA registry has been previously described in detail.[15] Briefly, at enrollment, patient demographics, smoking status, and RA disease history (eg, RA diagnosis date, prior treatments) are collected. ACR core measures are collected at enrollment and follow-up visits, as dictated by usual care.[16] These ACR core measures include 28-joint tender and swollen joint counts, patient and provider global assessments, pain, physical function, and acute-phase reactants. The registry has also previously been linked to administrative and electronic health record data in the VA Corporate Data Warehouse (CDW).[17,18]

Multimorbidity assessment. The primary exposures were unique multimorbidity patterns at the time of registry enrollment. The presence of 42 conditions used in RA multimorbidity studies[12] was assessed prior to enrollment through diagnostic codes from inpatient and outpatient encounters within the VA CDW. Patients were required to have at least two diagnostic codes for each condition from separate encounters, to reduce misclassification. All available VA data prior to the time of VARA enrollment were used (median 6.2 years). Based on the presence of these individual conditions, we applied previously developed multimorbidity patterns to the current cohort.[12] These multimorbidity patterns were derived through factor analysis of the aforementioned conditions within the VA and MarketScan Commercial Claims and Encounters Database. For this report, our primary analyses focused on multimorbidity patterns derived from the VA: mental health and substance abuse, metabolic, cardiovascular, and chronic pain. The chronic pain pattern of multimorbidity included specific chronic pain conditions and does not incorporate pain scores collected as part of RA management. We applied patterns derived in MarketScan in sensitivity analyses. Multimorbidity patterns and the individual conditions composing these patterns are listed in Supplementary Table. Consistent with prior work,[12] participants were required to have at least two conditions from the respective multimorbidity pattern for it to be considered present. Multimorbidity patterns were not mutually exclusive (ie, patients could have more than one multimorbidity pattern), and the number of multimorbidity patterns present was used as a measure of multimorbidity burden in secondary analyses.

Study outcomes. The primary study outcomes were RA disease activity and physical function over up to five years of follow-up after registry enrollment. Outcome follow-up was limited to five years because of decreasing availability of the outcomes after this time and because additional multimorbidity patterns may have developed. For each year of follow-up after enrollment, we calculated the mean values for each study outcome. Disease activity was measured using the Disease Activity Score in 28 joints (DAS28).[12] DAS28 with erythrocyte sedimentation rate (ESR) was preferentially used, but if only the C-reactive protein (CRP) was available, we calculated the DAS28-CRP. Functional status was measured using the Multidimensional Health Assessment Questionnaire (MDHAQ).[12] Secondary outcomes were the Clinical Disease Activity Index (CDAI)[12] and individual components of RA disease activity measures: patient global assessment (0–100), provider global assessment (0–100), tender joint count (0–28), swollen joint count (0–28), CRP, and ESR.

Study covariate. We selected study covariates a priori that may confound the association between multimorbidity patterns and study outcomes. The potential cofounders accounted for included age, gender, race (self-reported), smoking status (current, former, never), education level, rheumatoid factor (RF) or anti–cyclic citrullinated peptide (CCP) antibody positivity based on standardized assays,[19] RA disease duration, conventional synthetic DMARD use (methotrexate, azathioprine, hydroxychloroquine, sulfasalazine, minocycline, leflunomide), biologic DMARD use (etanercept, adalimumab, infliximab, golimumab, certolizumab, abatacept, interleukin-6 inhibitors, rituximab), and prednisone use. Janus kinase inhibitors were infrequently used at enrollment (0.4%). Covariates were fixed at baseline values and collected from the VARA registry, except for medications which were obtained from linked VA CDW pharmacy dispensing data sources.

Statistical analysis. Baseline characteristics of participants were assessed descriptively overall and by the presence of each multimorbidity pattern. We cross-sectionally assessed the associations of different patterns of multimorbidity with the DAS28 and MDHAQ at enrollment using ordinary least squares regression models adjusted for the aforementioned covariates. Sensitivity analyses were performed using multimorbidity patterns derived in MarketScan. In secondary analyses, we performed similar analyses of CDAI and individual disease activity components.

We longitudinally evaluated the impact of baseline multimorbidity patterns on DAS28 and MDHAQ using generalized estimation equations (GEEs) to account for the correlation of these measures among participants over time. Models were adjusted for covariates and specified an exchangeable covariance matrix. Because the multimorbidity patterns were derived to be unique patterns (ie, not highly correlated), all patterns were assessed in the same model. Interaction terms were tested between multimorbidity patterns and follow-up duration but were not significant and thus not included (all $P > 0.20$; data not shown). Additional analyses adjusted for enrollment disease activity and functional status values as well as restricted the population to patients with RA with a disease duration less than two years. We evaluated the inclusion of age squared to account for nonlinear associations, but this did not affect results and was not included in final models (data not shown). We also evaluated whether the number of multimorbidity patterns (0–4), an indicator of multimorbidity burden, was associated with longitudinal DAS28 and MDHAQ scores in similar GEE models. Pattern count was used in place of specific multimorbidity patterns for these models. The association of multimorbidity patterns with CDAI and individual disease activity components was assessed through similar GEE models in secondary analyses. In these analyses, tender joint count and swollen joint count were modeled using a negative binomial distribution, whereas all other components were modeled using a Gaussian distribution. The missing-indicator method was used to handle missing covariate data. All analyses were completed using Stata version 17 (StataCorp) within the VA Informatics and Computing Infrastructure environment.

RESULTS

Patient characteristics and multimorbidity pattern frequency. We studied 2,956 participants, of which the majority were male (88.2%), White (76.9%), and had a smoking history (79.3%). RF or anti-CCP seropositivity was present in 85.8% of participants. At enrollment, the majority (74.8%) of participants were taking conventional synthetic DMARDs, whereas 26.6% were taking biologic DMARDs and 41.0% were taking prednisone. At the five-year follow-up end of study period, 40.3% of participants had available DAS28 values, whereas 40.7% had MDHAQ values.

The metabolic multimorbidity pattern was the most frequent multimorbidity pattern (64.0%), followed by chronic pain (48.4%), mental health and substance abuse (23.2%), and cardiovascular multimorbidity (12.4%) (Table 1). Participants with cardiovascular multimorbidity were older, were more frequently male, and had a longer duration of RA. Participants with mental health and substance abuse or chronic pain multimorbidity were younger and had a higher level of education. Most (73.4%) participants had at least one multimorbidity pattern, with 24.9% having one, 26.0% having two, 19.0% having three, and 3.5% having all four multimorbidity patterns.

Multimorbidity patterns and RA outcomes at enrollment. In cross-sectional analyses at the time of registry enrollment, the cardiovascular (β 0.33 [95% confidence interval

Table 1. Participant characteristics by multimorbidity patterns at registry enrollment*

Characteristics	Overall	Multimorbidity patterns[a]			
		Metabolic	Chronic pain	Mental health and substance abuse	Cardiovascular
n (%)	2956	1,892 (64.0)	1,431 (48.4)	686 (23.2)	366 (12.4)
Age, y	64.5 (11.0)	67.1 (9.4)	64.9 (10.2)	62.3 (9.7)	70.6 (8.8)
Male, n (%)	2,606 (88.2)	1,722 (91.0)	1,241 (86.7)	572 (83.4)	348 (95.1)
White race, n (%)	2,273 (76.9)	1,450 (76.6)	1,073 (75.0)	483 (70.4)	289 (79.0)
BMI, mean (SD) kg/m^2	28.6 (5.7)	29.3 (5.7)	29.5 (6.0)	29.5 (5.8)	29.1 (5.7)
Smoking status, n (%)					
Never	594 (20.7)	358 (19.5)	266 (19.3)	111 (17.0)	62 (17.3)
Former	1,547 (54.0)	1,083 (58.9)	778 (56.5)	342 (52.3)	239 (66.8)
Current	726 (25.3)	398 (21.6)	332 (24.1)	201 (30.7)	57 (15.9)
Education, n (%)					
<High school	303 (11.6)	211 (12.6)	149 (11.8)	61 (10.3)	55 (17.0)
High school	1,007 (38.6)	645 (38.7)	447 (35.5)	181 (30.5)	128 (39.5)
>High school	1,300 (49.8)	812 (48.7)	664 (52.7)	352 (59.3)	141 (43.5)
RA duration, y	11.4 (11.3)	11.8 (11.5)	11.0 (11.1)	10.2 (10.4)	14.7 (13.7)
RF or anti-CCP positive, n (%)	2,210 (85.8)	1,392 (83.8)	1,008 (82.0)	480 (82.8)	274 (82.8)
DAS28	3.8 (1.6)	3.8 (1.6)	3.9 (1.5)	4.0 (1.6)	4.1 (1.7)
MDHAQ	0.9 (0.6)	0.9 (0.6)	1.0 (0.6)	1.1 (0.6)	1.0 (0.7)
Patient global (0–100)	39.5 (25.8)	40.0 (25.9)	43.5 (25.7)	45.7 (25.6)	42.4 (27.1)
Provider global (0–100)	32.8 (23.0)	32.6 (23.1)	34.3 (23.2)	35.1 (23.4)	36.8 (24.9)
Tender joint count	4.7 (6.5)	4.5 (6.4)	5.1 (6.8)	5.8 (7.2)	4.9 (6.9)
Swollen joint count	3.7 (5.2)	3.4 (4.9)	3.5 (4.9)	3.7 (5.1)	3.9 (5.5)
ESR	26.1 (22.9)	27.7 (23.8)	26.1 (23.0)	25.6 (23.4)	33.7 (27.1)
CRP	1.6 (2.9)	1.7 (3.2)	1.7 (3.3)	1.5 (2.3)	2.2 (3.2)
RA medications, n (%)					
csDMARDs	2,212 (74.8)	1,458 (77.1)	1,084 (75.8)	492 (71.7)	284 (77.6)
Biologic DMARDs	786 (26.6)	487 (25.7)	384 (26.8)	192 (28.0)	82 (22.4)
Prednisone	1,212 (41.0)	772 (40.8)	611 (42.7)	277 (40.4)	163 (44.5)

* BMI, body mass index; CCP, cyclic citrullinated peptide; CRP, C-reactive protein; csDMARD, conventional synthetic disease-modifying antirheumatic drug; DAS28, Disease Activity Score in 28 joints; ESR, erythrocyte sedimentation rate; MDHAQ, Multidimensional Health Assessment Questionnaire; RA, rheumatoid arthritis; RF, rheumatoid factor.
[a] Patients may have more than one multimorbidity pattern. Missing data: n = 98 smoking, n = 347 education, n = 371 seropositivity, n = 249 RA duration, n = 589 DAS28, n = 439 MDHAQ, n = 376 patient global, n = 1,126 provider global, n = 188 tender joint count, n = 188 swollen joint count, n = 330 ESR, n = 533 CRP.

{CI} 0.14, 0.53]) and chronic pain (β 0.17 [95% CI 0.02, 0.32]) patterns of multimorbidity were significantly associated with higher DAS28 scores (Figure 1A). DAS28 scores were numerically higher in those with mental health and substance abuse multimorbidity (β 0.14 [95% CI –0.03, 0.32]), although these were not significant. For functional status, mental health and substance abuse (β 0.11 [95% CI 0.04, 0.17]), cardiovascular (β 0.14 [95% CI 0.06, 0.21]), and chronic pain (β 0.13 [95% CI 0.07, 0.19]) multimorbidity were all significantly associated with higher MDHAQ scores (Figure 1B). Metabolic multimorbidity was not significantly associated with either DAS28 or MDHAQ scores at enrollment.

Multimorbidity patterns and longitudinal RA outcomes. In longitudinal analyses with up to five years of follow-up, cardiovascular (β 0.25 [95% CI 0.12, 0.38]), chronic pain (β 0.21 [95% CI 0.11, 0.31]), and mental health and substance abuse (β 0.12 [95% CI 0.00, 0.23]) multimorbidity were significantly associated with higher DAS28 scores (Figure 2A). However, after adjusting for baseline DAS28, only the chronic pain pattern of multimorbidity remained associated with higher DAS28 scores over follow-up (β 0.16 [95% CI 0.07, 0.25]); Supplementary Figure 1. Mental health and substance abuse (β 0.09 [95% CI [0.03, 0.15]), cardiovascular (β 0.11 [95% CI 0.04, 0.17]), and chronic pain (β 0.15 [95% CI 0.10, 0.20]) multimorbidity patterns were also associated with higher MDHAQ during follow-up (Figure 2B). When adjusted for baseline MDHAQ, only the chronic pain (β 0.09 [95% CI 0.05, 0.13]) pattern remained associated with MDHAQ scores (Supplementary Figure 1). Metabolic multimorbidity was not associated with longitudinal DAS28 or MDHAQ scores. In secondary analyses using the CDAI as the disease activity measure, the mental health and substance abuse (β 1.25 [95% CI 0.29, 2.21]), cardiovascular (β 1.38 [95% CI 0.27, 2.48]), and chronic pain (β 2.02 [95% CI 1.17, 2.86]) patterns of multimorbidity remained associated with RA disease activity over follow-up (Table 2). Similar results were found when restricting the population to participants with an RA duration less than 2 years (Supplementary Table 2).

Relative to those without any multimorbidity pattern, individuals with only one pattern of multimorbidity did not have higher DAS28 or MDHAQ scores over follow-up (Figure 3A, 3B).

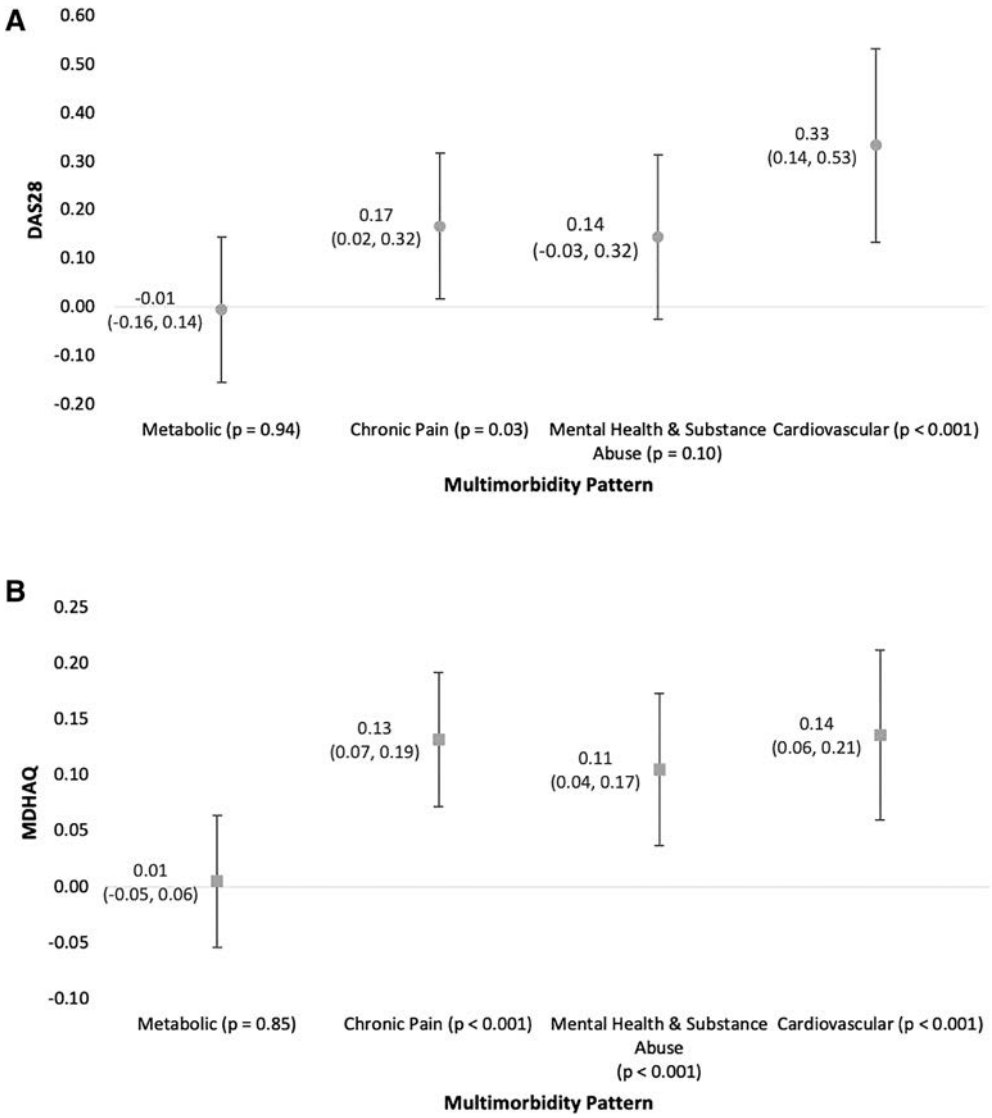


Figure 1. Cross-sectional associations of multimorbidity patterns with (A) disease activity (DAS28) and (B) functional status (MDHAQ) at the time of registry enrollment. Values are beta coefficients and 95% confidence intervals. Models adjusted for age, gender, education, smoking status, race, rheumatoid arthritis duration, rheumatoid factor or anti–cyclic citrullinated peptide seropositivity, conventional synthetic disease-modifying antirheumatic drugs (DMARDs), biologic DMARDs, and prednisone. DAS28, Disease Activity Score in 28 joints; MDHAQ, Multidimensional Health Assessment Questionnaire.

However, participants with more than two or more multimorbidity patterns had higher DAS28 and MDHAQ scores throughout follow-up. Those with all four multimorbidity patterns had the highest DAS28 (β 0.59 [95% CI 0.36, 0.83]) and MDHAQ scores (β 0.27 [95% CI 0.16, 0.39]). A test of linear trend across the number of multimorbidity patterns was highly significant for DAS28 ($P < 0.001$) and MDHAQ ($P < 0.001$). When adjusted for baseline DAS28 or MDHAQ, similar results were obtained although effect sizes were reduced (Supplementary Figure 2).

Multimorbidity patterns and longitudinal disease activity components. Multimorbidity patterns were differentially associated with individual disease activity components throughout follow-up. Mental health and substance abuse (β 3.58 [95% CI 1.56, 5.59]) and chronic pain (β 5.96 [95% CI 4.21, 7.71) patterns of multimorbidity were associated with increased patient global assessment (Table 2). The cardiovascular (β 3.14 [95% CI 1.23, 5.05]) and chronic pain (β 3.50 [95% CI 2.04, 4.95]) patterns of multimorbidity were associated with increased provider global. The metabolic (β −1.53 [95% CI −2.93, −0.13]) pattern of multimorbidity was associated with decreased provider global. The mental health and substance abuse (β 0.19 [95% CI 0.08, 0.29]), cardiovascular (β 0.13 [95% CI 0.01, 0.25]), and chronic pain (β 0.27 [95% CI 0.18, 0.37]) patterns were associated with increased tender joint count over follow-up. The cardiovascular pattern of multimorbidity was the

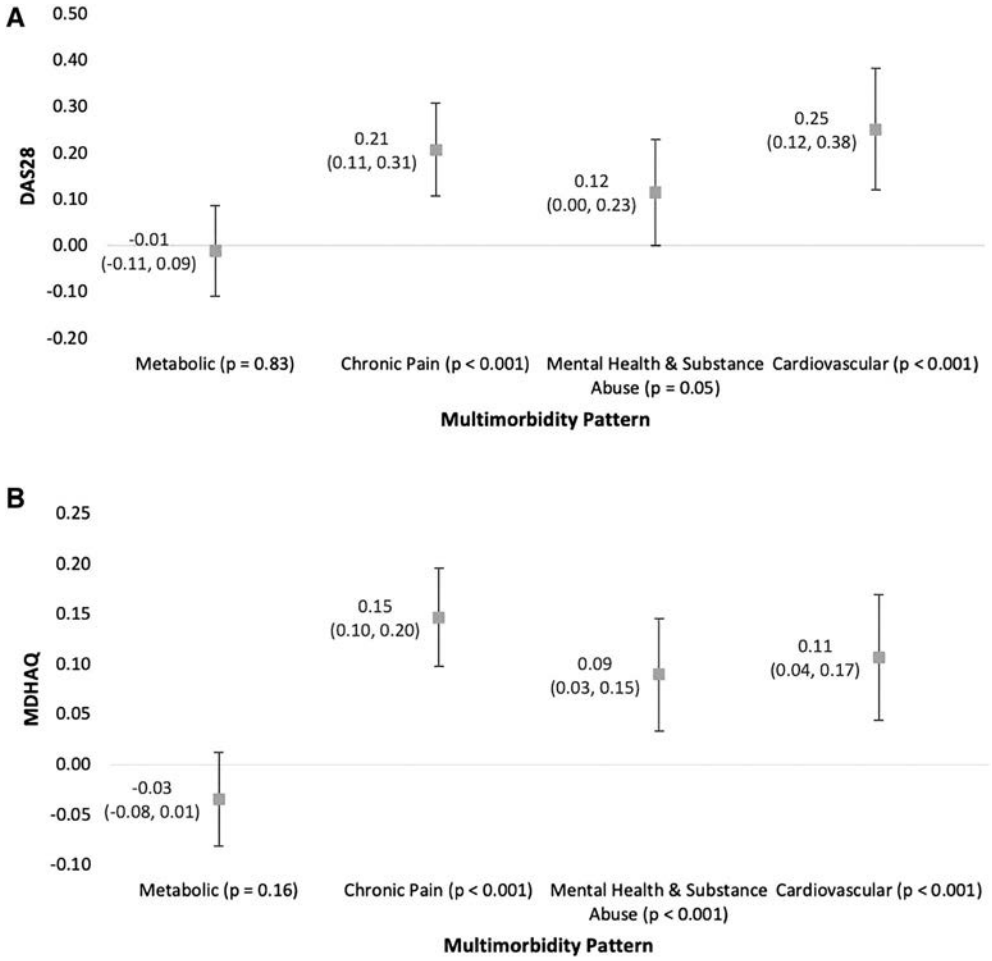


Figure 2. Associations of multimorbidity patterns with rheumatoid arthritis disease activity and functional status over follow-up. Longitudinal association of multimorbidity patterns with (A) disease activity (DAS28) and (B) functional status (MDHAQ) over follow-up (up to 5 years). Values are beta coefficients and 95% confidence intervals. Generalized estimating equations models adjusted for age, gender, education, smoking status, race, rheumatoid arthritis duration, rheumatoid factor or anti–cyclic citrullinated peptide seropositivity, conventional synthetic disease-modifying antirheumatic drugs (DMARDs), biologic DMARDs, and prednisone. DAS28, Disease Activity Score in 28 joints; MDHAQ, Multidimensional Health Assessment Questionnaire.

only pattern of multimorbidity associated with increased swollen joint count (β 0.19 [95% CI 0.07, 0.31]), CRP (β 0.21 [95% CI 0.03, 0.39]), and ESR (β 4.31 [95% CI 2.19, 6.44]) over follow-up. The metabolic pattern (β –0.10 [95% CI –0.19, –0.01]) of multimorbidity was associated with a lower swollen joint count.

Sensitivity analyses with alternative multimorbidity patterns. To assess the robustness of findings to the definition of multimorbidity patterns, we performed sensitivity analyses applying multimorbidity patterns originally derived in MarketScan.[12] The metabolic pattern of multimorbidity was most frequent (52.3%), followed by mental health and chronic pain (48.0%), vascular and neurologic (35.4%), and cardiopulmonary (22.6%) (Supplementary Table 3). Similar to the primary analyses, mental health and chronic pain multimorbidity and cardiopulmonary multimorbidity were associated with higher disease activity (DAS28 and CDAI) and MDHAQ scores at baseline and follow-up (Table 3, Supplementary Table 4), including among those with an RA duration less than 2 years (Supplementary Table 2). Moreover, a greater number of multimorbidity patterns were associated with higher DAS28 and MDHAQ scores. As in primary analyses, the mental health and chronic pain multimorbidity was associated with higher patient and provider global scores, whereas cardiopulmonary multimorbidity was associated more broadly with global scores, joint counts, and acute-phase reactants (Supplementary Table 4).

DISCUSSION

Multimorbidity is a growing public health problem that preferentially affects people with RA and contributes to several poor long-term health outcomes.[8] Although multimorbidity has primarily been studied using chronic disease counts or comorbidity indices,[2,8,11] novel tools have been developed in the general

Table 2. Associations of multimorbidity patterns with Clinical Disease Activity Index and individual disease activity components over follow-up*

Multimorbidity pattern	CDAI	Patient global (0–100)	Provider global (0–100)	Tender joint count	Swollen joint count	CRP	ESR
Metabolic	−0.70 (−1.51, 0.12); $P = 0.10$	−0.72 (−2.41, 0.97); $P = 0.40$	−1.53 (−2.93, −0.13); $P = 0.03$[a]	−0.08 (−0.17, 0.01); $P = 0.09$	−0.10 (−0.19, −0.01); $P = 0.04$[a]	−0.09 (−0.22, 0.04); $P = 0.18$	0.71 (−0.87, 2.28); $P = 0.38$
Chronic pain	2.02 (1.17, 2.86); $P < 0.001$[a]	5.96 (4.21, 7.71); $P < 0.001$[a]	3.50 (2.04, 4.95); $P < 0.001$[a]	0.27 (0.18, 0.37); $P < 0.001$[a]	0.06 (−0.03, 0.15); $P = 0.21$	0.02 (−0.12, 0.15); $P = 0.81$	−0.36 (−2.00, 1.27); $P = 0.67$
Mental health and substance abuse	1.25 (0.29, 2.21); $P = 0.01$[a]	3.58 (1.56, 5.60); $P < 0.001$[a]	0.43 (−1.23, 2.09); $P = 0.62$	0.19 (0.08, 0.29); $P < 0.001$[a]	0.03 (−0.08, 0.14); $P = 0.59$	0.02 (−0.13, 0.18); $P = 0.77$	−1.02 (−2.90, 0.85); $P = 0.29$
Cardiovascular	1.38 (0.27, 2.48); $P = 0.01$[a]	1.17 (−1.11, 3.46); $P = 0.32$	3.14 (1.23, 5.05); $P < 0.001$[a]	0.13 (0.01, 0.25); $P = 0.03$[a]	0.19 (0.07, 0.31); $P < 0.001$[a]	0.21 (0.03, 0.39); $P = 0.02$[a]	4.31 (2.19, 6.44); $P < 0.001$[a]

* Values are shown as β (95% confidence interval) from generalized estimation equation models adjusted for age, gender, education, smoking status, race, rheumatoid arthritis duration, seropositivity, conventional synthetic disease-modifying antirheumatic drugs (DMARDs), biologic DMARDs, and prednisone. CDAI = Clinical Disease Activity Index; CRP = C-reactive protein; ESR = erythrocyte sedimentation rate.
[a] Significant at $P < 0.05$.

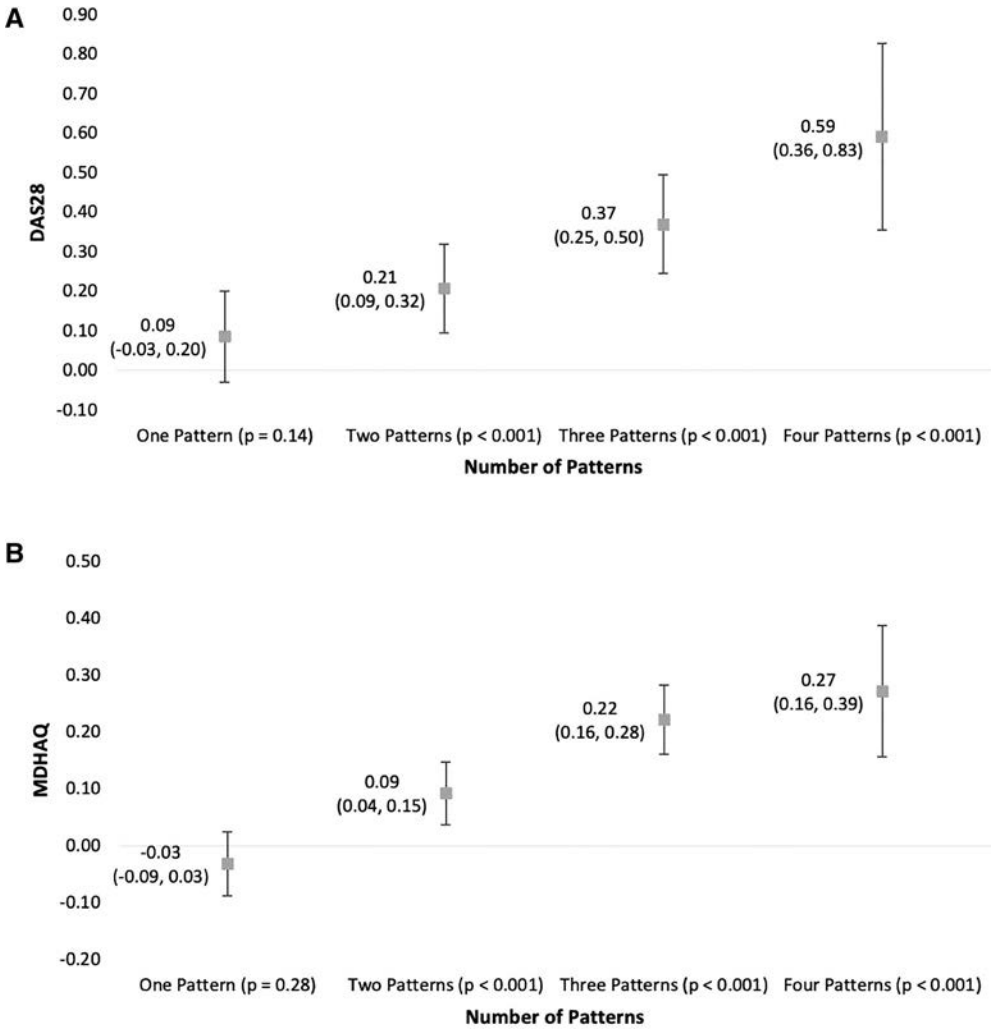


Figure 3. Associations of the number of multimorbidity patterns with rheumatoid arthritis disease activity and functional status over follow-up. Longitudinal association of the number of multimorbidity patterns present with (A) disease activity (DAS28) and (B) functional status (MDHAQ) over follow-up (up to 5 years). Values are beta coefficients and 95% confidence intervals. Generalized estimating equations models adjusted for age, gender, education, smoking status, race, rheumatoid arthritis duration, rheumatoid factor or anti–cyclic citrullinated peptide seropositivity, conventional synthetic disease-modifying antirheumatic drugs (DMARDs), biologic DMARDs, and prednisone. The reference group were those with no multimorbidity patterns. DAS28, Disease Activity Score in 28 joints; MDHAQ, Multidimensional Health Assessment Questionnaire.

population and RA to better assess multimorbidity.[12,13,20] In this study, we applied machine learning–derived multimorbidity patterns to a large prospective RA cohort to evaluate the association of these distinct multimorbidity patterns with RA disease activity and functional status over time. Patients with a greater burden of multimorbidity based on possessing several distinct multimorbidity patterns (ie, multidimensional multimorbidity) had a more severe disease course over follow-up. The multimorbidity patterns of mental health and substance abuse, cardiovascular, and chronic pain were all associated with higher longitudinal disease activity and worse functional status scores, but their effect on the individual components of disease activity scores varied by multimorbidity pattern. These findings illustrate the important differential associations of multimorbidity with RA-related outcomes and the potential for gains that could be realized by identifying and treating (or better, preventing) multimorbidity as part of holistic RA management.

We found that patients with the cardiovascular, mental health and substance abuse, and chronic pain patterns of multimorbidity were associated with worse disease activity and functional status over follow-up. Demonstrating the robustness of these findings, similar results were obtained when using multimorbidity patterns that were derived in an alternative data set. Together, these findings clearly illustrate that patients with RA with multimorbidity can have a more difficult disease course. Although we are unaware of prior studies that have assessed multimorbidity patterns and RA outcomes, these findings build on prior reports of individual comorbidities. Chronic mental health conditions such

Table 3. Association of multimorbidity patterns with disease activity and functional status at baseline and over follow-up using patterns derived in MarketScan*

Multimorbidity assessment	DAS28 at enrollment	MDHAQ at enrollment	DAS28 over follow-up	DAS28 over follow-up adj. for enrollment DAS28	MDHAQ over follow-up	MDHAQ over follow-up adj. for enrollment MDHAQ
Multimorbidity pattern						
Metabolic	−0.06 (−0.20, 0.08); $P = 0.42$	0.00 (−0.05, 0.06); $P = 0.93$	−0.01 (−0.10, 0.09); $P = 0.86$	0.01 (−0.08, 0.09); $P = 0.87$	−0.03 (−0.07, 0.02); $P = 0.26$	−0.05 (−0.09, −0.01); $P = 0.01$
Mental health and chronic pain	0.28 (0.15, 0.41); $P < 0.001$	0.17 (0.12, 0.22); $P < 0.001$	0.23 (0.15, 0.32); $P < 0.001$	0.12 (0.04, 0.20); $P = 0.00$	0.18 (0.14, 0.22); $P < 0.001$	0.10 (0.06, 0.13); $P < 0.001$
Vascular and neurologic	0.04 (−0.12, 0.20); $P = 0.64$	0.06 (−0.00, 0.12); $P = 0.05$	0.001 (−0.11, 0.11); $P = 0.99$	0.03 (−0.07, 0.13); $P = 0.58$	−0.02 (−0.07, 0.03); $P = 0.40$	−0.03 (−0.07, 0.01); $P = 0.19$
Cardiopulmonary	0.24 (0.07, 0.42); $P = 0.01$	0.13 (0.06, 0.19); $P < 0.001$	0.29 (0.17, 0.40); $P < 0.001$	0.12 (0.01, 0.23); $P = 0.03$	0.12 (0.06, 0.18); $P < 0.001$	0.06 (0.02, 0.11); $P = 0.01$
Number of multimorbidity patterns						
No patterns			Referent	Referent	Referent	Referent
1 pattern	–	–	0.15 (0.04, 0.27); $P = 0.01$	0.13 (0.03, 0.23); $P = 0.01$	0.07 (0.01, 0.12); $P = 0.02$	0.02 (−0.02, 0.07); $P = 0.31$
2 patterns	–	–	0.22 (0.10, 0.33); $P < 0.001$	0.14 (0.03, 0.25); $P = 0.01$	0.12 (0.06, 0.18); $P < 0.001$	0.05 (−0.00, 0.09); $P = 0.06$
3 patterns	–	–	0.14 (0.01, 0.27); $P = 0.04$	0.07 (−0.06, 0.19); $P = 0.28$	0.12 (0.05, 0.18); $P < 0.001$	0.02 (−0.04, 0.07); $P = 0.52$
4 patterns	–	–	0.71 (0.56, 0.87); $P < 0.001$	0.43 (0.28, 0.57); $P < 0.001$	0.30 (0.22, 0.38); $P < 0.001$	0.09 (0.03, 0.16); $P = 0.01$

* Values shown as β (95% confidence interval); *P* value. Models adjusted for age, gender, education, smoking status, race, rheumatoid arthritis duration, seropositivity, conventional synthetic disease-modifying antirheumatic drugs (DMARDs), biologic DMARDs, and prednisone. DAS28, 28-Joint Disease Activity Score; MDHAQ, Multidimensional Health Assessment Questionnaire.

as depression, anxiety, and posttraumatic stress disorder are associated with increased RA disease activity and proinflammatory cytokine expression.[21–24] Links between RA disease severity and cardiovascular disease are also well established,[7,25] as are those with chronic pain disorders such as fibromyalgia.[26] In contrast, the metabolic pattern of multimorbidity was not associated with worse disease activity or functional status in our study. This contrasts with prior studies that have found metabolic syndrome[27,28] and obesity[29] to be associated with RA disease severity. We speculate that prior studies' focus on the specific metabolic conditions of obesity and metabolic syndrome rather than our broader evaluation of metabolic multimorbidity, which includes several other related conditions (eg, vision changes, sleep disorders), as well as less ability to adjust for relevant covariates in prior studies could explain these differences. Additionally, we may not have been able to detect a meaningful impact of metabolic multimorbidity on RA outcomes because the majority of participants in our study (64%) had metabolic multimorbidity.

With several multimorbidity patterns being associated with longitudinal disease activity, we further evaluated whether they differentially affected the individual components of RA disease activity measures. The chronic pain and mental health and substance abuse patterns were associated with higher global scores and tender joint counts. In contrast, the cardiovascular pattern of multimorbidity was associated with higher provider global scores, tender and swollen joint counts, and acute-phase reactants. These findings extend insights gained from prior studies of individual comorbidities. The patient-derived components of composite disease activity measures are higher in patients with RA with fibromyalgia, but swollen joint counts and inflammatory measures do not seem to be altered.[30] Similar results are seen in patients with RA with posttraumatic stress disorder.[23] CVD in RA appears to be more closely related to inflammatory burden, both in the joints (eg, joint counts) and systemically (eg, acute-phase reactants).[7,25] Thus, the unique relationships between these individual conditions and disease activity components is captured by multimorbidity patterns and may aid in disease activity assessment.

Most assessments of multimorbidity or comorbidity in RA have used comorbidity counts or indices. Although these intend to measure multimorbidity or comorbidity burden, they may be heavily influenced by a group of highly related conditions (eg, several cardiovascular diseases or cardiovascular disease risk factors). We have recently proposed an alternative method of multimorbidity assessment, the number of distinct multimorbidity patterns present.[12] In this study, we applied this multimorbidity burden measure and assessed its relationship with RA disease outcomes. Participants with a greater number of multimorbidity patterns were more likely to have higher disease activity and poorer functional status throughout follow-up. Furthermore, the associations with RA disease outcomes persisted even when adjusting for disease activity or functional status values at enrollment. These findings provide construct validity for the novel multimorbidity measure and provide the first evidence that independent multimorbidity patterns are prognostic of the RA disease course. This is important for clinicians assessing real-world patients with RA who are typically multimorbid[2,3] and also for risk-adjustment purposes, as achievement of disease activity and functional status thresholds is considered a measure of quality of care.

Although we found persistent associations between multimorbidity and RA disease outcomes, there are several reasons these findings do not imply causation. As with all observational studies, unmeasured and residual confounding are potentially problematic. Additionally, interaction terms between multimorbidity patterns and time were not significant. This indicates that the trajectory of disease outcomes over time did not significantly differ by multimorbidity pattern. In fact, most differences in disease activity and functional status throughout follow-up were present at enrollment, which accounted for our decision to not adjust for baseline disease activity and functional status in our primary analysis. Once we adjusted for enrollment values, our findings were significantly attenuated, with only the chronic pain and highest number of multimorbidity pattern groups continuing to have poorer RA disease outcome measures throughout follow-up. However, this should not diminish the findings because we assessed prevalent multimorbidity at enrollment. Thus, an influence of multimorbidity on these RA outcome measures may already have occurred. Additional studies with complex time-varying assessments of multimorbidity and RA outcomes are needed to address this issue, to understand whether preventing multimorbidity may be possible, and to determine whether such interventions can influence their impact on subsequent RA outcomes.

There are limitations to this study. This study focused on the primary multimorbidity patterns found in patients with RA but was unable to assess all potential patterns of chronic diseases a patient with RA may have. Assessment of conditions in these patterns relied on linked administrative claims data, which may result in misclassification and prevents evaluation of the severity of conditions. This study did not evaluate whether conditions comprising the multimorbidity patterns were being appropriately managed. The study population was predominantly male, consistent with the characteristics of enrollees in the VA, but this may affect the generalizability. Additionally, this study did not investigate how different combinations of multimorbidity patterns may have a greater or lesser impact on RA functional status and disease activity than other combinations. Estimates for associations between multimorbidity patterns and RA disease activity did not exceed minimum important differences, suggesting that at a population level, the magnitude of association may be modest. RA disease severity and multimorbidity are likely intertwined in a bidirectional relationship. Therefore, in this prevalent RA cohort, it is possible that preceding RA disease activity and functional status were the drivers of multimorbidity at enrollment. Because the

availability of DAS28 and MDHAQ measures decreased over follow-up, our results may more closely reflect the associations between multimorbidity patterns and these measures over the earlier follow-up period. Finally, changes in RA treatments throughout follow-up were not assessed as mediators of observed associations, which will require future study.

In conclusion, patients with RA suffering from a greater burden of multimorbidity and the specific patterns of mental health and substance abuse, cardiovascular, and chronic pain multimorbidity were found to have an RA disease course with higher disease activity and poorer functional status. Targeting the identification and management of multimorbidity in patients with RA could facilitate achieving RA treatment targets and optimize long-term patient outcomes.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Dutt, England.
Acquisition of data. Roul, Yang, Sauer, Cannon, Baker, Mikuls, England.
Analysis and interpretation of data. Dutt, Roul, Yang, Johnson, Michaud, Sauer, Cannon, Baker, Curtis, Mikuls, England.

REFERENCES

1. Barnett K, Mercer SW, Norbury M, et al. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. Lancet 2012;380:37–43.
2. England BR, Roul P, Yang Y, et al. Burden and trajectory of multimorbidity in rheumatoid arthritis: a matched cohort study from 2006 to 2015. Ann Rheum Dis 2020;80:286–92.
3. Crowson CS, Gunderson TM, Dykhoff HJ, et al. Comprehensive assessment of multimorbidity burden in a population-based cohort of patients with rheumatoid arthritis. RMD Open 2022;8:e002022.
4. Radner H, Yoshida K, Frits M, et al. The impact of multimorbidity status on treatment response in rheumatoid arthritis patients initiating disease-modifying anti-rheumatic drugs. Rheumatology (Oxford) 2015;54:2076–84.
5. Spagnolo P, Lee JS, Sverzellati N, et al. The lung in rheumatoid arthritis. Arthritis Rheum 2018;70:1544–54.
6. England BR, Hershberger D. Management issues in rheumatoid arthritis-associated interstitial lung disease [review]. Curr Opin Rheumatol 2020;32:255–63.
7. England BR, Thiele GM, Anderson DR, et al. Increased cardiovascular risk in rheumatoid arthritis: mechanisms and implications [review]. BMJ 2018;361:k1036.
8. Canning J, Siebert S, Jani BD, et al. Examining the relationship between rheumatoid arthritis, multimorbidity and adverse health-related outcomes: a systematic review. Arthritis Care Res (Hoboken) 2022;74:1500–12.
9. Radner H, Yoshida K, Hmamouchi I, et al. Treatment patterns of multimorbid patients with rheumatoid arthritis: results from an international cross-sectional study. J Rheumatol 2015;42:1099–104.
10. Armagan B, Sari A, Erden A, et al. Starting of biological disease modifying antirheumatic drugs may be postponed in rheumatoid arthritis patients with multimorbidity: single center real life results. Medicine (Baltimore) 2018;97:e9930.
11. England BR, Yun H, Chen L, et al. Influence of multimorbidity on new treatment initiation and achieving target disease activity thresholds in active rheumatoid arthritis: a cohort study using the Rheumatology Informatics System for Effectiveness (RISE) Registry. Arthritis Care Res (Hoboken) 2021;75:231–9.
12. England BR, Yang Y, Roul P, et al. Identification of multimorbidity patterns in rheumatoid arthritis through machine learning. Arthritis Care Res (Hoboken) 2023;75:220–30.
13. Crowson CS, Gunderson TM, Davis JM III, et al. Using unsupervised machine learning methods to cluster comorbidities in a population-based cohort of patients with rheumatoid arthritis. Arthritis Care Res (Hoboken) 2023;75:210–9.
14. Arnett FC, Edworthy SM, Bloch DA, et al. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. Arthritis Rheum 1988;31:315–24.
15. Mikuls TR, Reimold A, Kerr GS, et al. Insights and implications of the VA rheumatoid arthritis registry [review]. Fed Pract 2015;32:24–9.
16. Felson DT, Anderson JJ, Boers M, et al. The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. Arthritis Rheum 1993;36:729–40.
17. George MD, Sauer BC, Teng CC, et al. Biologic and glucocorticoid use after methotrexate initiation in patients with rheumatoid arthritis. J Rheumatol 2018;46:343–50.
18. England BR, Sayles H, Michaud K, et al. Chronic lung disease in U.S. veterans with rheumatoid arthritis and the impact on survival. Clin Rheumatol 2018;37:2907–15.
19. Miriovsky BJ, Michaud K, Thiele GM, et al. Anti-CCP antibody and rheumatoid factor concentrations predict greater disease activity in men with rheumatoid arthritis. Ann Rheum Dis 2010;69:1292–7.
20. Prados-Torres A, Calderon-Larranaga A, Hancco-Saavedra J, et al. Multimorbidity patterns: a systematic review [review]. J Clin Epidemiol 2014;67:254–66.
21. Matcham F, Ali S, Irving K, et al. Are depression and anxiety associated with disease activity in rheumatoid arthritis? A prospective study. BMC Musculoskelet Disord 2016;17:155.
22. Michelsen B, Kristianslund EK, Sexton J, et al. Do depression and anxiety reduce the likelihood of remission in rheumatoid arthritis and psoriatic arthritis? Data from the prospective multicentre NOR-DMARD study. Ann Rheum Dis 2017;76:1906–10.
23. Mikuls TR, Padala PR, Sayles HR, et al. Prospective study of posttraumatic stress disorder and disease activity outcomes in US veterans with rheumatoid arthritis. Arthritis Care Res (Hoboken) 2013;65:227–34.
24. Maloley PM, England BR, Sayles H, et al. Post-traumatic stress disorder and serum cytokine and chemokine concentrations in patients with rheumatoid arthritis. Semin Arthritis Rheum 2019;49:229–35.
25. Crepaldi G, Scire CA, Carrara G, et al. Cardiovascular comorbidities relate more than others with disease activity in rheumatoid arthritis. PLoS One 2016;11:e0146991.
26. Duffield SJ, Miller N, Zhao S, et al. Concomitant fibromyalgia complicating chronic inflammatory arthritis: a systematic review and meta-analysis [review]. Rheumatology (Oxford) 2018;57:1453–60.
27. Da Cunha VR, Brenol CV, Brenol JC, et al. Metabolic syndrome prevalence is increased in rheumatoid arthritis patients and is associated with disease activity. Scand J Rheumatol 2012;41:186–91.
28. Karvounaris SA, Sidiropoulos PI, Papadakis JA, et al. Metabolic syndrome is common among middle-to-older aged Mediterranean patients with rheumatoid arthritis and correlates with disease activity:

a retrospective, cross-sectional, controlled, study. Ann Rheum Dis 2007;66:28–33.

29. Liu Y, Hazlewood GS, Kaplan GG, et al. Impact of obesity on remission and disease activity in rheumatoid arthritis: a systematic review and meta-analysis [review]. Arthritis Care Res (Hoboken) 2016; 69:157–65.

30. Lee YC, Hackett J, Frits M, et al. Multibiomarker disease activity score and C-reactive protein in a cross-sectional observational study of patients with rheumatoid arthritis with and without concomitant fibromyalgia. Rheumatology (Oxford) 2016;55: 640–8.

Arthritis Care & Research
Vol. 77, No. 3, March 2025, pp 349–358
DOI 10.1002/acr.25437

American College of Rheumatology
Empowering Rheumatology Professionals

Preferences for Tapering Biologic Disease-Modifying Antirheumatic Drugs Among People With Rheumatoid Arthritis: A Discrete Choice Experiment

Suz Jack Chan,[1] Lisa K. Stamp,[2] Gareth J. Treharne,[1] Janet M. Y. Cheung,[3] Nicola Dalbeth,[4] Rebecca Grainger,[5] Simon Stebbings,[1] and Carlo A. Marra[6]

Objective. Little is known about the preferences of people with rheumatoid arthritis (RA) regarding tapering of biologic disease-modifying antirheumatic drugs (bDMARDs). The aim of this study was to assess the preferences of people with RA in relation to potential treatment-related benefits and risks of bDMARD tapering and the health care service–related attributes that affect tapering.

Methods. Participants with RA who had experience taking a bDMARD completed an online discrete choice experiment. Participants were asked their preferences when given three hypothetical treatment scenarios in which varying the frequency of treatment might alter their chance of adverse effects, of regaining disease control, and of other health care service–related effects. Preference weights were estimated using a multinomial logit model.

Results. There were 142 complete responses. Reduced dosing frequency of bDMARD treatment had the largest impact on preference (mean 1.0, 95% confidence interval [CI] 0.8–1.2), followed by chance of disease flare (mean 0.7, 95% CI 0.6–0.9). Participants were willing to accept an increased risk of flare between 10.6% (95% CI 3.2–17.9) and 60.6% (95% CI 48.1–72.9) in exchange for benefits associated with tapering bDMARDs. Participants with better quality of life were more likely to choose to remain on current treatment. The predicted uptake of bDMARD tapering was high among people with RA, suggesting bDMARD tapering was a favored option.

Conclusion. For individuals with RA, making decisions about tapering bDMARDs involves considering several factors, with the most important determinants identified as dosing frequency and the risk of disease flare. Understanding patient perspectives of bDMARD tapering may enable physicians to make patient-focused shared health care decisions.

INTRODUCTION

Current treatment paradigms for managing rheumatoid arthritis (RA) aim to prevent joint damage and disability by frequently assessing disease activity and altering disease-modifying antirheumatic drug (DMARD) therapy to achieve either remission or low disease activity (LDA).[1] Since the introduction of biologic DMARDs (bDMARDs), more people with RA are able to achieve a state of LDA or remission.[2] Although bDMARDs are effective, their high cost and adverse event profile have led to debate about whether they can be tapered or ceased in people who achieve sustained RA disease remission. The tapering etanercept in RA trial randomized people with RA receiving etanercept 50 mg weekly for at least a year and who had been in Disease Activity Score in 28 joints remission for at least 6 months to continue weekly etanercept or to change to etanercept to 50 mg every other week. After 6 months, 26 out of 34 patients (76%) in the weekly and 19 out of 32 patients (59%) in the every other week group maintained disease control ($P = 0.136$).[3] However, cessation of DMARDs compared with tapering bDMARDs may lead to

Supported by the Health Research Council of New Zealand (18/776) and the University of Otago Research Fund.

[1]Suz Jack Chan, PhD, Gareth J. Treharne, PhD, Simon Stebbings, FRACP, MD: University of Otago, Dunedin, New Zealand; [2]Lisa K. Stamp, FRACP, PhD: University of Otago, Christchurch, Christchurch, New Zealand; [3]Janet M. Y. Cheung, PhD: University of Sydney, Sydney, New South Wales, Australia; [4]Nicola Dalbeth, FRACP, MD: The University of Auckland, Auckland, New Zealand; [5]Rebecca Grainger, FRACP, PhD: University of Otago, Wellington, Wellington, New Zealand; [6]Carlo A. Marra, PharmD, PhD: University of Otago, Dunedin, New Zealand, and Curtin University, Perth, Western Australia, Australia.

Drs Chan and Stamp contributed equally to this work.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25437).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25437.

Address correspondence via email to Carlo A. Marra, PharmD, PhD, at carlo.marra@curtin.edu.au.

Submitted for publication March 26, 2024; accepted in revised form September 16, 2024.

SIGNIFICANCE & INNOVATIONS

- Understanding the perspectives of people with rheumatoid arthritis (RA) on biologic disease-modifying antirheumatic drug (bDMARD) tapering may enable policymakers to make patient-focused policy and physicians to facilitate health care decision-making that supports patients' values and preferences.
- Reduced dosing frequency and lower risk of disease flare were the most important determinants for people with RA when facing choices to taper bDMARDs.

disease flare.[4] The 2022 EULAR Recommendations for the Management of RA state, "After glucocorticoids have been discontinued and a patient is in sustained remission, dose reduction of DMARDs (bDMARDs/[targeted synthetic] DMARDs and/or [conventional synthetic] DMARDs) may be considered."[5] Furthermore, the task force noted that tapering, with either dose reduction or an increase in dosing interval, was preferred over cessation.

To make an informed decision about tapering bDMARDs, people with RA need to consider and understand the trade-offs among the benefits of tapering (including potential reduced adverse effects, medication burden, and cost savings) and the potential risks of a disease flare. They also need to consider the potential burden on health care service that may affect them and others in relation to such issues as time to access health care should a flare occur. To date, there is limited research assessing people's preferences for tapering or withdrawing bDMARDs. Verhoef et al identified the major concern regarding tapering for people with RA was the possible increase in disease activity and its influence on pain and function.[6] Participants also needed to know that a return to the higher dose was possible if their RA flared and whether the bDMARD would be effective in controlling their disease after restarting.[6] No study has examined the trade-offs people with RA are willing to make among the risks and benefits of tapering bDMARD therapy after achieving remission. Key factors influencing people's attitudes to RA remission and bDMARD tapering include fear of uncertain outcomes of tapering, especially flare and joint damage, prioritized quality of life (QoL) from continuing bDMARDs over the risk of adverse effects, relief from the inconvenience of taking bDMARDs regularly, and assurance of prompt access to health care if their RA were to flare when tapering.[7] The aim of this study was to measure the preferences of people with RA when balancing treatment-related benefits, the risks of bDMARD tapering, and the health care burden that may result including delayed access to health care during a flare, which all affect patient attitudes to tapering.

PATIENTS AND METHODS

A discrete choice experiment (DCE) was employed to assess the preferences for tapering bDMARDs among people with RA. This method is underpinned by the random use framework,[8] which assumes decision-makers are rational and choose a product or service that provides the highest value (use) among other competing alternatives. The design and development of the DCE was conducted according to good research practices for stated-preference studies.[9–11] Ethical approval was obtained from the Health and Disability Ethics Committee of New Zealand (18/CEN/40). Patients consented to participate in the online survey by indicating that they had read, understood, and wanted to participate at the start of the survey.

Participant selection and recruitment. Potential participants were identified from rheumatology clinics in New Zealand. To be eligible, participants had to fulfill the following criteria: age 18 years or older, RA as defined by the American College of Rheumatology 2010 RA classification criteria,[12] and previously taken or were currently taking a bDMARD. Eligible participants were given or sent an invitation letter and participant information sheet describing the study's purpose, along with a weblink to participate in the survey. Data collection took place between January and September 2020.

Identification of attributes and levels selection. A multistep, funnel-shaped approach was used to identify relevant attributes in the design phase of the study.[13] A broad literature search using keywords of "rheumatoid arthritis," "biologic tapering," and "patients' perspectives" was performed to retrieve relevant articles on tapering bDMARDs in RA.[14] Next, as previously described,[7] 45 people with RA participated in focus group discussions and individual interviews to obtain insights into their perspectives on bDMARD tapering. The list of attributes identified from the literature search[15] was refined based on the qualitative analysis findings, which were then synthesized into relevant attributes and incorporated into the DCE. The levels and attributes are summarized in Table 1.

Experimental design. Seven attributes of four levels each generated 16,384 (4^7) possible profiles. Because it would not be feasible to include all profile questions, a fractional factorial experimental design was proposed. The experimental design was constructed using Ngene software (version 1.2.1)[16] to estimate a main-effect multinomial logit (MNL) model.

A Bayesian D-efficient model was computed using the Modified Federov algorithm, which maximizes the precision of the estimates for the unknown parameters.[17] The signs and magnitude of each of the priors used were informed by a pilot study with 14 people with RA. The final design has a mean Bayesian MNL d-error of 0.01022.

Table 1. List of the discrete choice experiment attributes and levels

Themes derived from qualitative work	Attributes	Levels	A priori expectation
Fear of the uncertainty of outcomes	Chance of a flare within 1 year of reducing biologic dose	25 of 100 (reference level), and 30, 50, and 70 of 100	Negative with increasing risk of flare
	Chance of regaining disease control after a flare within the next 6 months	50 of 100 (reference level), and 60, 80, and 100 of 100	Positive with an increasing chance of regaining disease control
Prioritizing the quality of life	Chance of serious infection within the next year	2, 4, and 6 of 100, and 8 of 100 (reference level)	Negative with increasing risk of serious infection
	Chance of skin cancer within the next 10 years	8, 10, and 12 of 100, and 14 of 100 (reference level)	Negative with increasing risk of skin cancer
Relief from inconvenience	Frequency of biologic treatment	Once every 4 weeks (reference level), and once every 6, 8, and 12 weeks	Positive with an increasing interval time between dosing
Prompt access to health care	Time to seeing my rheumatology team after a flare	Same day (reference level) Within 48 hours Within 5 days Within 10 days	Negative with increasing time to see a rheumatology team
	Chance a blood test would predict a flare in time for me to restart full treatment	50 of 100 (reference level), and 70, 80, and 90 of 100	Positive with an increasing chance of predicting a flare

The resulting design contained 36 choice tasks blocked into three versions, each with 12 choice tasks. Overlapping of some attribute levels was allowed to reduce task complexity and improve choice consistency.[18] Internal validity was checked by including a choice task to assess whether participants chose the alternative with all the attributes unambiguously better than all the other options.[14] In total, participants completed 14 choice tasks, including a warm-up choice task.

An unlabeled choice task was chosen in which two hypothetical treatment alternatives made up of the relevant attributes are assigned a generic description: "Option A" and "Option B." A status quo option indicating staying on current treatment was included to better reflect the actual choice process faced by respondents.[19,20] To minimize the loss of statistical power with the additional status quo option,[21,22] a dual-response design was used.[23,24]

Survey instrument development. The survey was programmed in Qualtrics (https://www.qualtrics.com). An introduction page explained the purpose of the study, followed by a comprehensive tutorial on how to answer the survey. In the preamble to each choice task, a hypothetical background scenario was provided, requesting participants to imagine they were doing very well on bDMARD treatment for RA, and their rheumatologist has suggested reducing the bDMARD dose. Participants were then asked to choose their preferred option within each choice set.

Participants were assigned to one of the three versions of the choice task. Randomization was done centrally in randomization blocks of six. The sequence of attributes was randomized in each choice task version to minimize attribute ordering effects.[25] Pictographs were used to improve participants' understanding of the benefits and risks presented within each choice. A background color coding design was used to indicate the differences among attribute levels for each alternative and facilitate a more straightforward comparison among them.[18] Darker shades of purple denote a "less desirable" level. Supplementary Figure 1 provides an example of the choice task.

Demographic and clinical information were collected, including age, sex, disease duration, and current and previous DMARDs, as well as visual analog scales of pain, global well-being, fatigue, stiffness, and the health assessment questionnaire. The EuroQol 5-domain (EQ-5D) questionnaire was included to investigate participants' health-related QoL.[26] A self-reported subjective numeracy scale (SNS)[27] and risk propensity scale (RPS)[28] was employed to measure participants' perceived numeracy skills and their general risk-taking tendencies.

Sample size. The minimum sample size necessary for the DCE was established according to a rule-of-thumb proposed by Orme[29] using the formula $nta/c \geq 500$, where n is the number of respondents, t is the number of tasks (t = 12), a is the number of alternatives per task (a = 2 excluding the none alternative), and c is the largest number of levels for any one attribute (c = 4 maximum number of levels). A sample size of 105 people with RA was deemed sufficient to estimate the main effects in the statistical model.

Statistical analyses. The attribute levels included in the DCE were effects coded. Effects coding captures nonlinearities in the marginal use for levels of attributes and was chosen over dummy coding to avoid an identification problem inherent in dummy coding in which the use (preference weight) associated with individual attribute level is confounded with the constant term (or grand mean). Descriptive statistics were performed with Microsoft Excel. The DCE data were estimated using an MNL

model with NLOGIT 6.0. For details of the MNL Model, see Supplementary Material 1. A relatively larger, positive preference weight indicated a greater preference for the attribute level, whereas a relatively smaller, negative preference weight was interpreted as less preferred.

The Wald test was used to test for statistically significant differences among the β-coefficients. The mean relative importance of an attribute was determined by the absolute difference among the β-coefficient for the best and worst levels of that attribute. As the differences in β-coefficients were measured on an arbitrary scale, the difference within an attribute with the largest magnitude was assigned a value of one. The overall mean importance of other attributes was measured relative to this change.[30] The extent to which participants were willing to trade off levels of risks associated with bDMARD tapering in exchange for a specific increase in treatment benefits was determined using marginal rates of substitution (MRS). The MRS was presented as maximum acceptable risk (MAR) and is calculated as the absolute value of the ratio between the coefficient of one benefit measure and coefficient of an adverse event. The probability of an individual choosing the tapering option with the specified attributes and corresponding levels was estimated using the formula $P = 1 / (1 + e^{-V})$,[31] where V is defined as the deterministic element that is specified as a linear index of the attributes. The current treatment was chosen as the base case described by the reference level of each attribute. Best-case and worst-case scenarios for tapering bDMARDs assumed the most and the least desirable levels for each attribute.

Participants' sociodemographic characteristics and their responses to the numeracy and risk propensity scores were included in the MNL to determine the effects of respondents' characteristics on the likelihood of choosing the current treatment. The variables were entered up to J−1 (where J = number of alternatives) directly into the use functions and subsequently estimated using the backward regression approach, whereby all variables were entered into the model initially.[32] Models were iteratively assessed and evaluated for the goodness of fit by comparing log-likelihood, McFadden's pseudo-R^2, and Akaike information criterion (AIC) estimates among models.[10]

RESULTS

A total of 736 potential participants were invited to complete the survey. Of the 160 people who responded (21.7%), 18 were excluded (16 incomplete responses, and 2 gave the same response for all questions). Data on the 576 who did not respond were not available. Of the 142 respondents, 26 (18.3%) failed the dominance test. Two MNL analyses were conducted to compare the models including and excluding the data of those 26 respondents. These analyses indicated that excluding respondents who failed the dominance test had no impact on the magnitude of the attributes and did not substantially change the results. Therefore, the data of all 142 respondents were included in subsequent analyses.

Demographics and clinical features. Table 2 summarizes the characteristics of 142 participants. The majority were female (81%) and New Zealand European ethnicity (93%). Mean age was 60.3 years (23–89 years) with a mean RA disease duration of 20.8 years (1.5–58 years). There were 133 out of 142 participants (93.7%) who were currently receiving bDMARD monotherapy or in combination with other medications (including nonsteroidal anti-inflammatory drugs, prednisone, and targeted synthetic DMARDs and conventional synthetic DMARDs). Eight participants were previously prescribed bDMARDs. Of those currently receiving bDMARDs, 42 (31.3%) were receiving adalimumab, 29 (21.6%) were receiving rituximab, 27 (20.1%) were receiving tocilizumab, 25 (18.7%) were receiving etanercept, and 10 (7.5%) were receiving infliximab.

Participants' preferences. Data analysis started with an attribute-only MNL model (Supplementary Table 1) as the base model. Next, an MNL model fitted with participants' sociodemographic characteristics (Table 3) was compared with the base model. The fully fitted MNL showed improved model fit and model parsimony (log-likelihood, pseudo-R^2, and AIC) over the attribute-only MNL. In total, 3,408 observations were recorded. Options A or B were chosen 95.1% of the time compared with the current option (4.9%). In Supplementary Figure 2, the direction of β-coefficient for most of the attribute levels was consistent with the natural ordering of the categories, in which better outcomes are preferred to worse outcomes. The majority of attribute levels significantly influenced respondents' treatment decisions ($P < 0.05$). Respondents preferred to have a longer dosing interval, lower chances of RA flare, lower chances of serious infection, a shorter time to rheumatology team consultation after a flare, higher chances of regaining RA disease control after a flare, and a higher blood test accuracy in predicting a flare. Table 4 shows the results of the Wald tests for comparison among the β-coefficients within the attributes.

Overall relative importance. Comparing the differences among preference weights for best and worst levels of each attribute yielded an estimate of the relative importance of that attribute over the range of levels included relative to any other attribute. Frequency of bDMARDs treatment dosing was the most important attribute (1.0; 95% confidence interval [CI] 0.81–1.19), followed by the chance of flare (0.74; 95% CI 0.59–0.80). Next, the chance of serious infection (0.37; 95% CI 0.25–0.49) had equal relative importance as the chance of regaining disease control after a flare (0.37; 95% CI 0.23–0.52). Other results included the chance of infection (0.37; 95% CI 0.23–0.49) and chance of cancer (0.28; 95% CI 0.18–0.39). The chance a blood test would

Table 2. Demographic characteristics and psychometric data of participants (n = 142)*

Variable	Value
Demographics	
Age, mean (SD)	60.34 (12.16)
RA disease duration, mean (SD)[a]	20.75 (12.09)
Female, n (%)	115 (80.9)
Ethnicity, n (%)	
NZ European	124 (87.3)
Māori	8 (5.6)
Pacific peoples	5 (3.5)
Other	4 (2.8)
Asian	1 (0.08)
Education, n (%)	
Less than high school	2 (1.4)
High school	48 (33.6)
Postsecondary	12 (8.4)
Diploma	31 (22.4)
Bachelor	30 (21.0)
Postgraduate	13 (9.8)
Prefer not to answer	6 (3.5)
Employment, n (%)	
Employed	62 (44.8)
Self-employed	15 (10.5)
Retired	47 (32.9)
Homemaker	7 (4.9)
Student	2 (1.4)
Unemployed	4 (2.8)
Other	5 (2.8)
Income (NZD), n (%)	
<20,000	8 (5.6)
20,000–49,000	42 (29.4)
50,000–100,000	38 (26.6)
>100,000	31 (22.4)
Prefer not to answer	23 (16.1)
Current medication, n (%)[a]	
NSAIDs	51 (36.2)
Prednisone	42 (29.8)
Methotrexate	63 (44.7)
Leflunomide	13 (9.3)
Sulfasalazine	3 (2.1)
Azathioprine	2 (1.4)
Hydroxychloroquine	17 (12.1)
Rituximab	29 (20.6)
Adalimumab	42 (29.8)
Etanercept	25 (17.8)
Infliximab	10 (7.1)
Tocilizumab	27 (19.1)
Not taking bDMARDs	8 (5.7)
Disease activity, mean (SD)[a]	
HAQ-II	0.93 (0.67)
Pain VAS	27.4 (25.81)
Fatigue VAS	45.05 (30.39)
Patient global assessment VAS	32.0 (25.29)
Stiffness VAS	32.12 (25.82)
EQ-5D 3L, mean (SD)[a]	1.4 (0.52)
Self-care	1.21 (0.44)
Usual activities	1.58 (0.59)
Pain/discomfort	1.73 (0.49)
Anxiety/depression	1.31 (0.49)
Health today VAS	65.92 (20.1)
Index value[b]	0.63 (0.169)
Risk propensity score[a,c]	3.42 (1.41)
Subjective numeracy score[a,d]	3.99 (1.13)

* bDMARD, biologic disease-modifying antirheumatic drug; EQ-5D 3L, EuroQol 5-domain 3-level; HAQ, health assessment questionnaire; NSAID, nonsteroidal anti-inflammatory drug; NZ, New Zealand; NZD, New Zealand dollar; RA, rheumatoid arthritis; VAS, visual analog scale.
[a] Data are only available for 141 participants.
[b] Calculated based on EQ-5D-3L value set for NZ.[45]
[c] Risk propensity scale measures risk-taking tendencies across seven items on a scoring range of one to nine. A higher score is interpreted as being more likely to take risks.
[d] Subjective numeracy scale measures perceived ability to perform mathematical tasks and preference to use numerical or prose information across eight items on a scoring range of one to six. A higher score is interpreted as higher numeracy skill and preference for numerical information.

predict a flare (0.26; 95% CI 0.14–0.38) was ranked the least important among the seven attributes.

MAR. The MAR of disease flare in exchange for treatment benefits are presented in Table 5. Therefore, each 1% chance of a disease flare corresponds to an absolute 0.027-unit change in use (95% CI –0.099 to 0.055). The difference in the use among the levels of the other attributes of interest was then measured on the same scale to determine the MAR among attributes. For bDMARD dosing frequency, the difference between the use of "once every 4 weeks" and "once every 6 weeks" is 1.446 (0.237 to –1.209). Thus, the MAR equates to 1.446/0.027, which corresponds to 53.1% (95% CI 41.2%–65.0%) increase in the chance of disease flare to trade for reducing the frequency of bDMARD treatment from once every four weeks to once every six weeks. For the "risk of adverse events," MAR for a flare was 22.15% (95% CI 16.2%–28.1%) in return for a reduction in the chance of serious infection from 8 to 2 out of 100 within a year and 17.2% (95% CI 12.2%–22.3%) in exchange for a reduction in the chance of skin cancer from 14 to 8 out of 100 in the next 10 years.

Predicted uptake of bDMARD tapering. The uptake probabilities for bDMARD tapering ranged from 16.5% to 85.4%, depending on the combination of attribute levels (Table 6). Current treatment, presented as the base case, had an uptake probability of 20.3% (95% CI 12.2%–32.0%). The worst-case scenario had the lowest uptake probability of 16.5% (95% CI 11.4%–23.3%). In contrast, the uptake probability of the best-case scenario with all the most desirable levels was 85.4% (95% CI 79.3%–90.0%). Realistic scenarios (ie, middle ground between best-case and worst-case scenarios) were found to have a relatively high uptake probability between 68% and 83%.

Table 3. Fitted model: multinomial logit analysis fitted with sociodemographic characteristics based on preference for biologic disease-modifying antirheumatic drugs tapering (n = 142)*

	Estimate (95% CI)	t ratio	*P* value
Alternative specific constant			
Status quo	0.21 (−0.40 to 0.81)	0.67	0.50
Option A	0.18 (0.04–0.33)	2.44	0.02
Frequency of bDMARD treatment			
Once every 4 weeks[a]	−1.21 (−1.44 to −0.97)	−10.08	<0.001
Once every 6 weeks	0.24 (0.12–0.36)	3.84	<0.001
Once every 8 weeks	0.53 (0.37–0.68)	6.60	<0.001
Once every 12 weeks	0.44 (0.32–0.56)	7.12	<0.001
Chance of a flare within one year of reducing biologic dose			
25 out of 100[a]	0.43 (0.31–0.55)	5.84	<0.001
30 out of 100	0.27 (0.11–0.42)	3.38	<0.001
50 out of 100	0.10 (−0.11 to 0.32)	0.93	0.35
70 out of 100	−0.80 (−0.93 to −0.67)	−11.78	<0.001
Chance of serious infection within the year			
2 out of 100	0.31 (0.19–0.42)	5.16	<0.001
4 out of 100	0.09 (−0.07 to 0.24)	1.09	0.27
6 out of 100	−0.09 (−0.25 to 0.06)	−1.18	0.24
8 out of 100[a]	−0.30 (−0.41 to −0.19)	−5.22	<0.001
Chance of skin cancer within the next 10 years			
8 out of 100	0.162 (0.03–0.29)	2.40	0.02
10 out of 100	0.24 (0.04–0.45)	2.32	0.02
12 out of 100	−0.10 (−0.28 to 0.09)	−1.00	0.32
14 out of 100[a]	−0.31 (−0.44 to −0.17)	−4.46	<0.001
Chance of regaining disease control after flare within six months			
50 out of 100[a]	−0.30 (−0.45 to −0.16)	−4.00	<0.001
60 out of 100	−0.36 (−0.56 to −0.16)	−3.56	<0.001
80 out of 100	0.36 (0.17–0.54)	3.77	<0.001
100 out of 100	0.31 (0.18–0.43)	4.86	<0.001
Time to see my rheumatology team after a flare			
Within a day[a]	0.31 (0.14–0.48)	3.50	<0.001
Within 2 days	0.10 (−0.07 to 0.27)	1.2	0.23
Within 5 days	−0.26 (−0.51 to −0.01)	−2.03	0.04
Within 10 days	−0.15 (−0.36 to 0.06)	−1.40	0.16
Chance a blood test will predict a flare in time for me to restart full treatment			
50 out of 100[a]	−0.18 (−0.32 to −0.05)	−2.70	0.01
70 out of 100	−0.06 (−0.33 to 0.22)	−0.41	0.68
80 out of 100	−0.01 (−0.20 to 0.18)	−0.10	0.92
90 out of 100	0.25 (0.3–0.37)	4.00	<0.001
Covariates (reference to option A/B)			
Disease duration	0.01 (0.00–0.02)	2.56	0.01
Education level	0.34 (0.10–0.58)	2.82	0.01
Income and employment	0.16 (−0.08 to 0.39)	1.31	0.19
RPS score[b]	−0.04 (−0.11 to 0.04)	−0.94	0.35
SNS score[c]	−0.02 (−0.12 to 0.09)	−0.29	0.77
EQ-5D 3L index score	−0.70 (−1.33 to −0.07)	−2.17	0.03
Model fit			
Log-likelihood	−1,909.96		
McFadden pseudo-R^2	0.36		
AIC	1.43		
BIC	1.49		

* AIC, Akaike information criterion; bDMARD, biologic disease-modifying antirheumatic drug; BIC, Bayesian information criterion; CI, confidence interval; EQ-5D 3L, EuroQol 5-domain 3-level; RPS, risk propensity scale; SNS, subjective numeracy scale.

[a] Reference level. The coefficient for each reference level was calculated as the negative sum of the other levels' coefficient. Negative coefficient represents disutility. Grand mean had an expected use of zero.

[b] Risk propensity scale measured risk-taking tendencies across seven items on a scoring range of one to nine. A higher score was interpreted as being more likely to take risks.

[c] Subjective numeracy scale measured perceived ability to perform mathematical tasks and preference to use numerical or prose information across eight items on a scoring range of one to six. A higher score was interpreted as higher numeracy skills and a preference for numerical information.

Table 4. Statistical comparison of differences among the β-coefficients (Wald test)

Attribute	Chi-square	*P* value
Frequency		
Changing from 6 weeks to 8 weeks	6.45	0.01
Changing from 8 weeks to 12 weeks	0.68	0.41
Changing from 6 to 12 weeks	4.63	0.03
Chance of flare		
Increase from 30% to 50%	2.30	0.13
Increase from 50% to 70%	34.77	0.00
Increase from 30% to 70%	180.29	0.00
Chance of infection		
Increase from 2% to 4%	8.53	0.004
Increase from 4% to 6%	0.58	0.45
Increase from 2% to 6%	13.18	0.0003
Chance of cancer		
Increase from 8% to 10%	0.01	0.93
Increase from 10% to 12%	3.87	0.04
Increase from 8% to 12%	5.95	0.01
Chance of regaining disease control		
Increase from 60% to 80%	24.94	0.00
Increase from 80% to 100%	0.50	0.48
Increase from 60% to 100%	39.95	0.00
Time to see a doctor		
Increase from 2 days to 5 days	5.26	0.02
Increase from 5 days to 10 days	0.08	0.78
Increase from 2 days to 10 days	4.50	0.03
Chance of test		
Increase from 70% to 80%	0.07	0.79
Increase from 80% to 90%	9.33	0.002
Increase from 70% to 90%	2.83	0.09

Influence of respondents' sociodemographic and other characteristics on treatment choice. Results from univariate analyses are shown in Supplementary Table 2. In the multivariate analysis, three covariates were found to be statistically significant ($P < 0.05$) in determining the likelihood of respondents choosing to stay on current treatment. Specifically, respondents with RA for a longer period and those with higher education levels (bachelor's degrees or higher) were more likely to choose to taper their bDMARD. In contrast, respondents with a higher EQ-5D score (perceived as better QoL) were more likely to prefer status quo treatment. Four other covariates included in the analysis, income status, employment status, the RPS score, and the SNS score, did not significantly affect respondents' choices among the three alternatives.

Table 5. Maximum acceptable risk of flare that participants are willing to accept for benefits of tapering biologic disease-modifying antirheumatic drugs*

Improvement in perceived treatment benefit	Mean maximum acceptable risk of flare	95% CI
Decrease in bDMARDs dosing frequency from 4 to 6 weeks	53.13	41.24–65.02
Decrease in bDMARDs dosing frequency from 6 to 8 weeks	10.55	3.16–17.93
Decrease in bDMARDs dosing frequency from 4 to 12 weeks	60.55	48.14–72.95
Decrease in chance of serious infection from 8 to 2 out of 100 within a year	22.15	16.20–28.10
Decrease in chance of skin cancer from 14 to 8 out of 100 in the next 10 years	17.22	12.20–22.25

* bDMARD, biologic disease-modifying antirheumatic drug; CI, confidence interval.

DISCUSSION

This study quantitatively assesses the preferences for bDMARDs among people with RA who have experience in using bDMARDs for tapering their therapy. Seven factors influencing choices for bDMARD tapering were identified, specifically, in the order of importance, the frequency of bDMARD treatment, the chance of a RA flare, the chance of regaining disease control after a flare, the chance of serious infection, the chance of blood test predicting a flare, the chance of skin cancer, and time to see a rheumatologist after a flare. Respondents were willing to trade an increased risk of flare in exchange for possible benefits associated with bDMARD tapering. In addition, those people who had RA for longer and those with higher education levels expressed a stronger preference for bDMARD tapering. In contrast, those who indicated higher health-related QoL were more likely to prefer to remain on bDMARD treatment.

Frequency of bDMARD dosing was the most important attribute influencing treatment choices. Poulos et al reported people with RA were willing to accept a greater risk of adverse effects and lower treatment efficacy in exchange for treatments with a longer dosing interval.[33] In the context of bDMARD tapering, reducing the dosing frequency may alleviate the burden of frequent, long-term treatment[7,34] and provide an opportunity to lead a more flexible and normal lifestyle[35] that could positively impact health-related QoL. Of note, the highest impact on utility was for reducing the current treatment frequency of once every four weeks to once every six weeks compared with changes from current treatment to either once every 8 or 12 weeks, all else being equal. This suggests that people with RA would prefer to have the frequency of treatment gradually reduced rather than an immediate large reduction or cessation, supporting the current EULAR guidelines.

As anticipated, the chance of an RA flare was an important factor in participants' decisions to accept a recommendation to taper bDMARDs. Fear of returning to a state of uncontrolled disease remains one of the major concerns for patients considering tapering.[6,36] Although current evidence suggests that bDMARD tapering can be safe and effective,[37] addressing patients' concerns about flaring is critical. For example, patients should be provided information on the early signs of an RA flare and provided

Table 6. Predicted biologic disease-modifying antirheumatic drug tapering uptake based on varying treatment scenario*

Attribute	Base case (current)	Worst-case scenario	Realistic scenario					Best-case scenario	
			1	2	3	4	5	6	
Frequency of treatment, weeks	4	6	6	6	8	8	12	12	12
Chance of a flare (out of 100)	25	70	50	30	50	30	70	30	25
Chance of serious infection (out of 100)	8	8	6	4	4	8	2	4	2
Chance of skin cancer (out of 100)	14	14	12	10	10	12	8	10	8
Chance of regaining disease control (out of 100)	50	50	100	100	80	100	80	80	100
Time to see a rheumatologist	Within 1 day	Within 10 days	Within 1 day	Within 5 days	Within 2 days	Within 5 days	Within 1 day	Within 2 days	Within a day
Chance of test predicting a flare (out of 100)	50	50	50	70	80	90	50	70	90
Predicted uptake, %	20.3	16.5	68.3	73.3	83.1	70.8	68.4	83.5	85.4
95% CI	12.2–32.0	11.4–23.3	59.4–75.9	64.6–80.5	72.9–89.9	60.6–79.2	58.2–76.9	75.3–89.4	79.3–90.0

* CI, confidence interval.

with plans to monitor and manage flares. Regaining disease control after a flare was an important attribute for tapering decisions; the effectiveness of reinitiating previous bDMARDs dose or other treatment strategies, such as switching bDMARDs, to regain the previous well-controlled disease state should be discussed in the context of recent evidence on bDMARD tapering.[34] Accordingly, a collaborative decision-making approach can facilitate a greater understanding that bDMARD tapering is feasible as a long-term treatment plan and aligns patient and physician expectations and timeline of bDMARD tapering.

Among the three treatment benefits of bDMARD tapering, reducing dosing frequency had the most impact on participants' decision, whereas chance of serious infection and skin cancer had the lowest impact. Reducing the frequency of treatment has a more immediate and tangible impact on the participant than on the uncertainties surrounding the future risk of a serious infection or skin cancer risk, consistent with the previous qualitative work.[7] Similar results were observed in other studies, in which on average, people with RA tended to prioritize dosing frequency of treatment over the potential risk of adverse effects.[38–40] Another possible explanation is the relatively narrow range of levels selected for both chance of serious infection and skin cancer. Levels for both attributes were carefully chosen based on a thorough literature search and were considered realistic and meaningful. However, participants may find the relatively small changes across the levels comparable and, therefore, less sensitive to changes in the attribute levels.

This study investigated the preference of people with RA regarding health care services by focusing on the attitudes related to the likelihood of an RA flare posed as "chance of a blood test predicting a flare" and "time to a see my rheumatology team after a flare" within the DCE. At present, an accurate flare prediction test is currently unavailable, but advances in imaging and biomarkers show promise for predicting flare risk when tapering bDMARDs, and these could increase uptake of tapering among those with stable disease.

The current study predicted that a significant portion of the RA population (68%–83%) would consider tapering bDMARDs, aligning with previous findings indicating favorable attitudes toward tapering.[7] Clinicians and health care funders can use these insights to improve their communication plans regarding bDMARD tapering and tailor the tapering information for eligible individuals who have excellent RA disease control or RA disease remission.

Strengths of the current study include the extensive use of qualitative and quantitative methods to identify, refine, and develop key attributes relevant to people with RA when faced with decisions to accept the recommendation of tapering bDMARDs. Although a previous study specifically focused on ranking factors that influence patients' decisions,[6] this study further explored the extent of risk of flare, which people with RA were willing to accept in exchange for a reduction in dosing frequency of treatment and risk of serious infection and skin cancer.

Limitations to this study include a low response rate to the online DCE survey of 21.7%, resulting in sampling bias. The study occurred during the first nationwide COVID-19 lockdown in New Zealand, which had considerable impact on people with RA.[41–43] We were not able to determine the reasons for nonparticipation nor the patient characteristics of the people with RA invited that did not participate. There may be additional sampling

bias given the topic of the study and because the survey was conducted online, limiting participation to people with personal electronic devices and internet access. The participants' sociodemographic profile reported high mean age, a predominance of females, and longer disease duration, all of which reflect the population of people with RA in Aotearoa/New Zealand. Thus, these results may not be generalizable to younger people, males, and those with shorter disease duration. Additionally, identifying participants from only one institution may limit the diversity of the sample and the overall generalizability of the findings. Lastly, as with all DCE studies, there may be a gap between stated preferences and actual decisions because of the evaluation of hypothetical choices with limited attributes.[44]

In summary, this study provides evidence that choices for tapering bDMARDs among people with RA are most influenced by the dosing frequency and perceived chance of flare. People are willing to accept a greater chance of flare in exchange for a reduction in dosing frequency but were less willing to risk flare-ups in exchange for a reduction in the risk of serious infections or skin cancer. Those with self-reported better QoL prefer to choose current treatment than to taper. Understanding the perspectives of people with RA on bDMARD tapering can inform the development of policy around medication access and enable physicians to support shared decision-making that aligns with patients' values and preferences.

ACKNOWLEDGMENTS

We thank Professor Michiel Bliemer (University of Sydney) for his expertise and assistance with the experimental design of DCE, Gin Nie Chua (University of Aberdeen) for providing advice on data analysis, Judy Chiu (University of British Columbia) for designing the survey webpage, and Janine Haslett and Niamh Fanning (University of Otago, Christchurch, New Zealand) for their assistance with recruiting participants. We also thank all the people with RA who participated in this study. Open access publishing facilitated by University of Otago, as part of the Wiley - University of Otago 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 Marra 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/Helsinki Declaration requirements.

REFERENCES

1. Smolen JS, Aletaha D, Bijlsma JW, et al; T2T Expert Committee. Treating rheumatoid arthritis to target: recommendations of an international task force. Ann Rheum Dis 2010;69(4):631–637.
2. Aletaha D, Smolen JS. Achieving clinical remission for patients with rheumatoid arthritis. JAMA 2019;321(5):457–458.
3. Bertrand D, Stouten V, De Cock D, et al. Tapering of etanercept is feasible in patients with rheumatoid arthritis in sustained remission: a pragmatic randomized controlled trial. Scand J Rheumatol 2022; 51(6):470–480.
4. Smolen JS, Nash P, Durez P, et al. Maintenance, reduction, or withdrawal of etanercept after treatment with etanercept and methotrexate in patients with moderate rheumatoid arthritis (PRESERVE): a randomised controlled trial. Lancet 2013;381(9870):918–929.
5. Smolen JS, Landewé RBM, Bergstra SA, et al. EULAR recommendations for the management of rheumatoid arthritis with synthetic and biological disease-modifying antirheumatic drugs: 2022 update. Ann Rheum Dis 2023;82(1):3–18.
6. Verhoef LM, Selten EMH, Vriezekolk JE, et al. The patient perspective on biologic DMARD dose reduction in rheumatoid arthritis: a mixed methods study. Rheumatology (Oxford) 2018;57(11):1947–1955.
7. Chan SJ, Stamp LK, Liebergreen N, et al. Tapering biologic therapy for rheumatoid arthritis: a qualitative study of patient perspectives. Patient 2020;13(2):225–234.
8. McFadden D. Conditional logit analysis of qualitative choice behavior. Academic Press; 1973.
9. Bridges JF, Hauber AB, Marshall D, et al. Conjoint analysis applications in health–a checklist: a report of the ISPOR Good Research Practices for Conjoint Analysis Task Force. Value Health 2011;14(4): 403–413.
10. Hauber AB, González JM, Groothuis-Oudshoorn CG, et al. Statistical methods for the analysis of discrete choice experiments: a report of the ISPOR Conjoint Analysis Good Research Practices Task Force. Value Health 2016;19(4):300–315.
11. Reed Johnson F, Lancsar E, Marshall D, et al. Constructing experimental designs for discrete-choice experiments: report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force. Value Health 2013;16(1):3–13.
12. Aletaha D, Neogi T, Silman AJ, et al. 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 2010;62(9):2569–2581.
13. Günther OH, Kürstein B, Riedel-Heller SG, et al. The role of monetary and nonmonetary incentives on the choice of practice establishment: a stated preference study of young physicians in Germany. Health Serv Res 2010;45(1):212–229.
14. Johnson FR, Yang JC, Reed SD. The internal validity of discrete choice experiment data: a testing tool for quantitative assessments. Value Health 2019;22(2):157–160.
15. Chan SJ, Yeo HY, Stamp LK, et al. What are the preferences of patients with rheumatoid arthritis for treatment modification? a scoping review. Patient 2021;14(5):505–532.
16. Ngene C. *Ngene 1.1. 1 user manual & reference guide.* Choice-Metrics; 2018.
17. Rose J, Bliemer M. Constructing efficient stated choice experimental designs. Transp Rev 2009;29(5):587–617.
18. Jonker MF, Donkers B, de Bekker-Grob E, et al. Attribute level overlap (and color coding) can reduce task complexity, improve choice consistency, and decrease the dropout rate in discrete choice experiments. Health Econ 2019;28(3):350–363.
19. Louviere J, Hensher D, Swait J. Stated choice methods: analysis and applications. Cambridge University Press; 2000.
20. Bennett J, Blamey R. The choice modelling approach to environmental valuation. Edward Elgar Publishing; 2001.
21. Dhar R, Simonson I. The effect of forced choice on choice. J Mark Res 2018;40(2):146–160.

22. Veldwijk J, Lambooij MS, de Bekker-Grob EW, et al. The effect of including an opt-out option in discrete choice experiments. PLoS One 2014;9(11):e111805.
23. Brazell J, Diener C, Karniouchina E, et al. The no-choice option and dual response choice designs. Mark Lett 2006;17(4):255–268.
24. Rose J, Hess S. Dual-response choices in pivoted stated choice experiments. Trans Res Record 2135(1):25–33.
25. Heidenreich S, Phillips-Beyer A, Flamion B, et al. Benefit-risk or risk-benefit trade-offs? another look at attribute ordering effects in a pilot choice experiment. Patient 2021;14(1):65–74.
26. Hurst NP, Jobanputra P, Hunter M, et al; Economic and Health Outcomes Research Group. Validity of Euroqol–a generic health status instrument–in patients with rheumatoid arthritis. Br J Rheumatol 1994;33(7):655–662.
27. Fagerlin A, Zikmund-Fisher BJ, Ubel PA, et al. Measuring numeracy without a math test: development of the Subjective Numeracy Scale. Med Decis Making 2007;27(5):672–680.
28. Meertens R, Lion R. Measuring an individual's tendency to take risks: the risk propensity scale. J Appl Soc Psychol 2008;38(6):1506–1520.
29. Orme B. Getting Started with Conjoint Analysis: Strategies for Product Design and Pricing Research. Research Publishers; 2006.
30. Bridges JF, Mohamed AF, Finnern HW, et al. Patients' preferences for treatment outcomes for advanced non-small cell lung cancer: a conjoint analysis. Lung Cancer 2012;77(1):224–231.
31. Hall J, Kenny P, King M, et al. Using stated preference discrete choice modelling to evaluate the introduction of varicella vaccination. Health Econ 2002;11(5):457–465.
32. Hensher D, Rose J, Greene W. Applied choice analysis: a primer. Cambridge University Press; 2005.
33. Poulos C, Hauber AB, González JM, et al. Patients' willingness to trade off between the duration and frequency of rheumatoid arthritis treatments. Arthritis Care Res (Hoboken) 2014;66(7):1008–1015.
34. Fautrel B, Den Broeder AA. De-intensifying treatment in established rheumatoid arthritis (RA): Why, how, when and in whom can DMARDs be tapered? Best Pract Res Clin Rheumatol 2015;29(4–5):550–565.
35. Hewlett S, Haig-Ferguson A, Rose-Parfitt E, et al. Dose reduction of biologic therapy in inflammatory arthritis: A qualitative study of patients' perceptions and needs. Musculoskeletal Care 2019;17(1):63–71.
36. Wallis D, Holmes C, Holroyd C, et al. Dose reduction of biological therapies for inflammatory rheumatic diseases: what do patients think? Scand J Rheumatol 2019;48(3):251–252.
37. Verhoef LM, van den Bemt BJ, van der Maas A, et al. Down-titration and discontinuation strategies of tumour necrosis factor-blocking agents for rheumatoid arthritis in patients with low disease activity. Cochrane Database Syst Rev 2019;5(5):CD010455.
38. Alten R, Krüger K, Rellecke J, et al. Examining patient preferences in the treatment of rheumatoid arthritis using a discrete-choice approach. Patient Prefer Adherence 2016;10:2217–2228.
39. Louder AM, Singh A, Saverno K, et al. Patient preferences regarding rheumatoid arthritis therapies: a conjoint analysis. Am Health Drug Benefits 2016;9(2):84–93.
40. Husni ME, Betts KA, Griffith J, et al. Benefit-risk trade-offs for treatment decisions in moderate-to-severe rheumatoid arthritis: focus on the patient perspective. Rheumatol Int 2017;37(9):1423–1434.
41. Brooks J, Montgomery A, Dalbeth N, et al. Omicron variant infection in inflammatory rheumatological conditions - outcomes from a COVID-19 naive population in Aotearoa New Zealand. Lancet Reg Health West Pac 2023;38:100843.
42. Johnstone G, Treharne GJ, Fletcher BD, et al. Mental health and quality of life for people with rheumatoid arthritis or ankylosing spondylitis in Aotearoa New Zealand following the COVID-19 national lockdown. Rheumatol Int 2021;41(10):1763–1772.
43. Treharne GJ, Johnstone G, Fletcher BD, et al. Fears about COVID-19 and perceived risk among people with rheumatoid arthritis or ankylosing spondylitis following the initial lockdown in Aotearoa New Zealand. Musculoskeletal Care 2022;20(2):290–298.
44. Fifer S, Rose J, Greaves S. Hypothetical bias in Stated Choice Experiments: is it a problem? And if so, how do we deal with it? Transp Res Part A Policy Pract 2014;61:164–177.
45. Devlin NJ, Hansen P, Kind P, et al. Logical inconsistencies in survey respondents' health state valuations – a methodological challenge for estimating social tariffs. Health Econ 2003;12(7):529–544.

Arthritis Care & Research
Vol. 77, No. 3, March 2025, pp 359–365
DOI 10.1002/acr.25197

American College of Rheumatology
Empowering Rheumatology Professionals

Radiographic Changes Five Years After Treatment of Meniscal Tear and Osteoarthritic Changes

Jeffrey N. Katz,[1] Jamie E. Collins,[1] Robert H. Brophy,[2] Brian J. Cole,[3] Charles L. Cox,[4] Ali Guermazi,[5] Morgan H. Jones,[1] Bruce A. Levy,[6] Lindsey A. MacFarlane,[1] Lisa A. Mandl,[7] Robert G. Marx,[7] Faith Selzer,[1] Kurt P. Spindler,[8] Rick W. Wright,[4] Elena Losina,[1] and Yuchiao Chang[9]

Objective. Meniscal tear in persons aged ≥45 years is typically managed with physical therapy (PT), and arthroscopic partial meniscectomy (APM) is offered to those who do not respond. Prior studies suggest APM may be associated with greater progression of radiographic changes.

Methods. We assessed changes between baseline and 60 months in the Kellgren-Lawrence (KL) grade and OARSI radiographic score (including subscores for joint space narrowing and osteophytes) in subjects aged 45–85 years enrolled into a seven-center randomized trial comparing outcomes of APM with PT for meniscal tear, osteoarthritis changes, and knee pain. The primary analysis classified subjects according to treatment received. To balance APM and PT groups, we developed a propensity score and used inverse probability weighting (IPW). We imputed a 60-month change in the OARSI score for subjects who underwent total knee replacement (TKR). In a sensitivity analysis, we classified subjects by randomization group.

Results. We analyzed data from 142 subjects (100 APM, 42 PT). The mean ± SD weighted baseline OARSI radiographic score was 3.8 ± 3.5 in the APM group and 4.0 ± 4.9 in the PT group. OARSI scores increased by a mean of 4.1 (95% confidence interval [95% CI] 3.5–4.7) in the APM group and 2.4 (95% CI 1.7–3.2) in the PT group ($P < 0.001$) due to changes in the osteophyte component. We did not observe statistically significant differences in the KL grade. Sensitivity analyses yielded similar findings to the primary analysis.

Conclusion. Subjects treated with APM had greater progression in the OARSI score because of osteophyte progression but not in the KL grade. The clinical implications of these findings require investigation.

INTRODUCTION

Meniscal tears are prevalent, especially in persons with knee osteoarthritis (OA). They are seen in the magnetic resonance imaging (MRIs) in 30–40% of persons 60–69 years old in community-based samples (1) and are present in over 80% of persons with established knee OA (1–3). The clinical syndrome of knee pain in middle-aged persons with imaging evidence of a meniscal tear is typically managed with exercises, physical therapy (PT), activity modifications, and pain control (4–7). Patients who do not respond to these measures are sometimes offered arthroscopic partial meniscectomy (APM) (4–7), which is conducted over 300,000 times annually in the United States and frequently in other countries (8–12).

The impact of resecting a damaged or torn meniscus on the progression of underlying cartilage damage and osteophyte formation is not clear. One study showed greater worsening on MRI measures of cartilage damage both 18 and 60 months following APM than following PT (13,14). Of the four studies that assessed changes in the Kellgren-Lawrence (KL) grade on plain

Clinical Trials.gov identifier: NCT00597012.

Supported by the NIH/National Institute of Arthritis and Musculoskeletal (grants R01-AR-055557, P30-AR-072577, K01-AR-075879, and K23-AR-080206).

[1]Jeffrey N. Katz, MD, MSc, Jamie E. Collins, PhD, Morgan H. Jones, MD, MPH, Lindsey A. MacFarlane, MD, MPH, Faith Selzer, PhD, Elena Losina, PhD: Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts; [2]Robert H. Brophy, MD: Washington University School of Medicine, St. Louis, Missouri; [3]Brian J. Cole, MD, MBA: Rush Medical Center, Chicago, Illinois; [4]Charles L. Cox, MD, Rick W. Wright, MD: Vanderbilt Medical Center, Nashville, Tennessee; [5]Ali Guermazi, MD, PhD, MSc: Boston Veteran's Medical Center and Boston University Medical Center, Boston, Massachusetts; [6]Bruce A. Levy, MD: Mayo School of Medicine, Rochester, Minnesota; [7]Lisa A. Mandl, MD, Robert G. Marx, MD: Hospital for Special Surgery, New York, New York; [8]Kurt P. Spindler, MD: Cleveland Clinic, Cleveland, Ohio; [9]Yuchiao Chang, PhD: Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts.

Additional supplementary information cited in this article can be found online in the Supporting Information section (https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25197).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25197.

Address correspondence via email to Jeffrey N. Katz, MD, MSc, at jnkatz@bwh.harvard.edu.

Submitted for publication April 7, 2023; accepted in revised form July 6, 2023.

SIGNIFICANCE & INNOVATIONS

- Patients with knee pain, meniscal tear, and osteoarthritic changes on x-ray or MRI are generally treated with physical therapy (PT). Those with persistent pain after PT are often treated with arthroscopic partial meniscectomy (APM). Some studies have suggested that APM is associated with greater progression of radiographic findings.
- Subjects with meniscal tear and osteoarthritic changes treated with APM had greater worsening in osteophyte burden (though not in joint space narrowing) on knee radiographs than subjects treated with PT.
- The clinical significance of these radiographic findings merits longer follow-up studies.

radiographs 5 years after randomization to APM or PT or placebo surgery, two of the studies documented greater changes in surgically treated subjects (15,16) and two did not find evidence of differences between subjects randomized to APM or to the comparator (17,18). Notably, it is unclear whether radiographic progression is associated with worse clinical outcomes.

Radiographic changes in this setting have generally been assessed with advancement in KL grade (19), a metric for assessing OA severity based on the presence of osteophytes and the severity of joint space narrowing (JSN) (15–17). KL grades range from 0 (no JSN or osteophytes) to 4 (severe OA). One study also assessed radiographic change with the OARSI score (20), which sums six individual scores (JSN; rated 0–3) in the medial and lateral compartment and osteophyte formation (rated 0–3) in the medial and lateral tibia and medial and lateral femur (15). Because the OARSI score has a much wider range (0–18) than the KL grade (0–4), it has the potential to be more responsive to change.

In this article, we examine changes in radiographs between subjects treated with APM and those treated with PT in the Meniscal Tear in Osteoarthritis Research (MeTeOR) Trial, a multicenter randomized controlled trial (RCT) that compared APM with PT in persons with knee pain, meniscal tear, and imaging evidence of degenerative cartilage damage. We hypothesized that treatment with APM would be associated with greater worsening in OARSI radiographic score than treatment with PT (21,21).

METHODS

Sample. We examined baseline and 60-month radiographs of subjects in MeTeOR, a seven-center RCT that compared outcomes of APM with those of a standardized PT regimen in subjects at least 45 years old with meniscal tear and osteoarthritic changes documented on MRI (21,22). Subjects with severe OA (KLgrade 4) were excluded. For this analysis, we also excluded subjects who were randomized at one of the MeTeOR centers that did not participate in the radiographic follow-up activities, subjects who died before the 60-month follow-up visit, and subjects who did not have a baseline radiograph available for analysis. We also excluded those who did not have a 60-month radiograph unless they underwent TKR prior to 60 months. Finally, we excluded two subjects who crossed over from PT to APM greater than 184 days after randomization because their follow-up time included substantial periods before and after APM, blurring the distinction between “nonoperative” and “operative" status. We included subjects who went on to have TKR irrespective of whether they met other exclusions because they were likely to have experienced progressive knee pain due to worsening of OA (Figure 1). The Mass General Brigham Institutional Review Board approved this study (protocol no. 2020P002004).

Acquisition of radiographs. Each subject underwent a standing bilateral posterior–anterior (PA) knee radiograph at study entry as a component of clinical care. We invited subjects to return for study-ordered follow-up standing bilateral knee radiographs at 18 and 60 months after randomization. These follow-up radiographs were performed with a Synaflexer frame (Bioclinica), ensuring a standard PA fixed-flexed view.

Radiographic grading. Radiographs were sent to the coordinating center at Brigham and Women’s Hospital where they were assessed by one of three raters (two rheumatologists and a fourth-year medical student). Raters were blinded to treatment and temporal order. After a rigorous training period, but before the raters assessed the MeTeOR radiographs, we assessed reliability across the three raters in an independent sample of 25 radiographs. Weighted kappa statistics for pairwise agreement among the three raters for KL grading ranged from 0.75 to 0.79, whereas the Spearman correlation among the three pairs of raters on the OARSI score ranged from 0.88 to 0.89.

Radiograph scores. We assessed radiographs with two commonly used radiographic scores: The OARSI score (20) permits the summation of osteophyte scores (rated 0–3) across four regions (medial tibia, medial femur, lateral tibia, lateral femur) and joint space narrowing scores (also 0–3) across the medial and lateral tibiofemoral compartments. The possible range of the OARSI score is 0–18. We created an osteophyte subscale of the OARSI score, consisting of the sum of the osteophyte ratings across four regions (medial and lateral tibiae, medial and lateral femoral; possible range 0–12), and a JSN subscale consisting of the sum of the two JSN ratings (medial, lateral; possible range 0–6). We assigned a KL grade (19) to each radiograph. To ensure the categories were collectively exhaustive and mutually exclusive, we regarded grade 0 as no JSN or osteophytes; grade 1 denotes a possible osteophyte with none or possible JSN; grade 2 denotes a definite osteophyte with none or possible

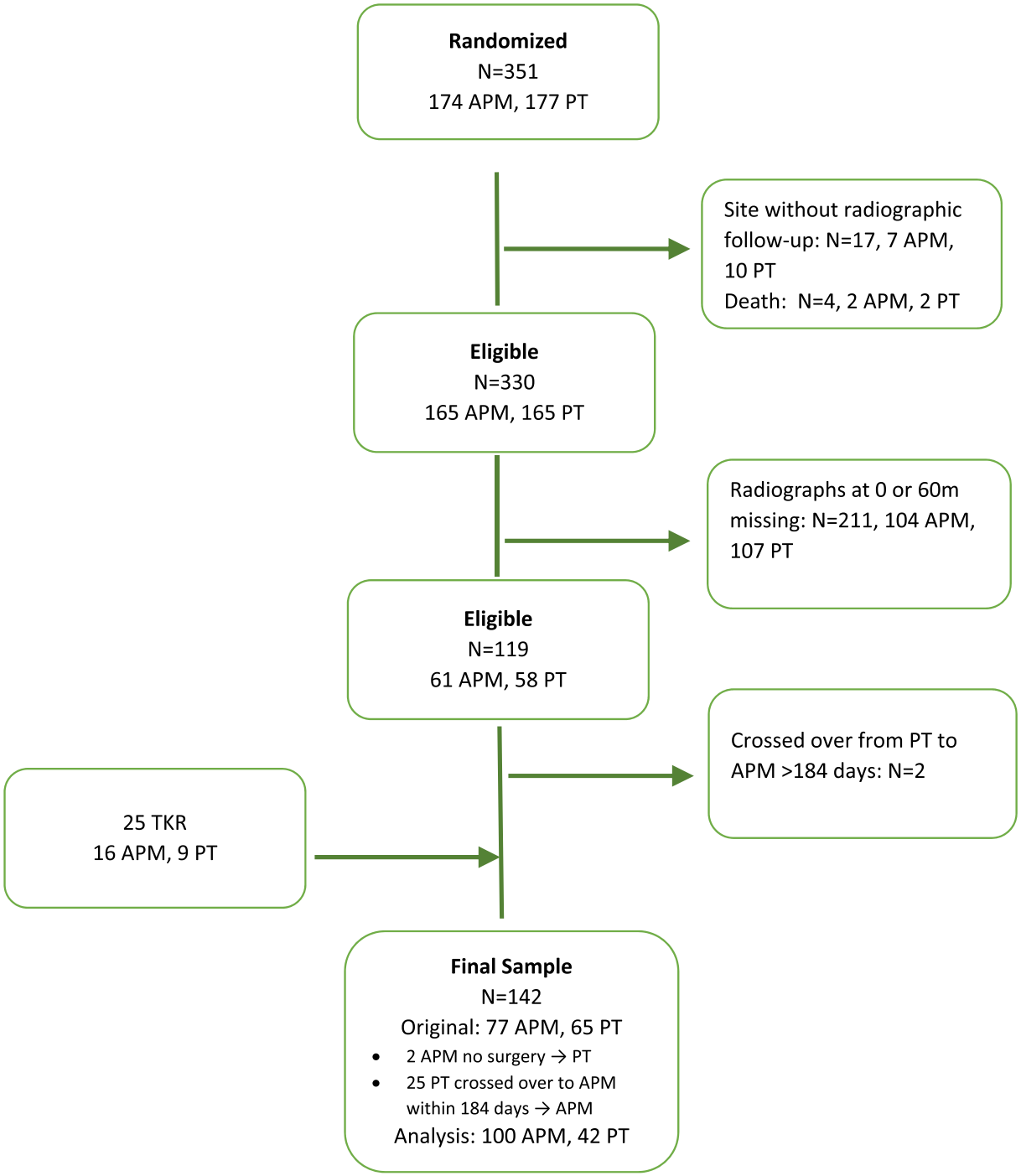


Figure 1. Flow chart of study sample selection. APM = arthroscopic partial meniscectomy; PT = physical therapy; TKR = total knee replacement.

JSN; grade 3 denotes JSN but not bone on bone; and grade 4 denotes bone on bone JSN (23).

Analysis. We assessed changes between baseline and 60 months (60-month score – baseline score) in overall OARSI score and separately for the OARSI osteophyte and JSN components and in the KL grade. To address potential informative loss to follow-up among subjects who had TKR before the 60-month visit, we imputed their 60-month KL grade as KL 4 (advanced OA) and imputed their change of OARSI score from baseline to 60-month as the 90th percentile of the distribution from subjects with OARSI scores available.

In the primary analysis, we sought to examine the association between the treatment received (APM vs. PT) and radiographic changes. Thus, the APM group included subjects who underwent surgery after being randomized to APM *and* those who were randomized to PT but crossed over to APM within 184 days of randomization. The PT group was composed of subjects randomized to PT

and subjects who were randomized to but did not receive APM. In this as-treated analysis, the study sample no longer benefited from the balance of confounders achieved by the original randomization. To reduce potential bias, we used the propensity score approach to balance the APM and PT groups with respect to prognostically important variables (24). The variables included in the propensity score model (modeling propensity to be in the APM group) included baseline demographic variables (age, sex, race, body mass index), pain severity, and KL grade. We then compared changes from baseline to 60 months in KL grade (changes had a normal appearing distribution) and in the OARSI score between the two treatment groups using two-sample *t*-tests, both unadjusted and with the inverse probability weights. We performed an additional analysis that compared the proportion of subjects in each group that increased 1 SD more than the mean increase in OARSI score, and the proportion that advanced by at least one KL grade using chi-square tests.

We performed a sensitivity analysis based on the intention-to-treat principal and analyzed subjects according to the group they were randomized to. In a second set of sensitivity analyses, we reran the primary as-treated analysis and the analysis by randomization group after excluding subjects who received TKR before the 60-month visit. This exclusion obviated the need to impute radiographic scores in these subjects. For each of the sensitivity analyses, we derived a propensity score and recalculated weights.

RESULTS

MeTeOR randomized 351 subjects. One center with 17 subjects did not participate in the imaging follow-up portion of the study, and 4 subjects died, resulting in 330 subjects potentially eligible for this analysis. Of these, 211 did not have paired baseline and 60-month radiographs, and 2 subjects crossed over from PT to APM less than 184 days following randomization. Thus, the final sample included 117 subjects with paired baseline and 60-month radiographs, and another 25 subjects who underwent TKR, yielding 142 subjects (100 APM, 42 PT) who were included for this analysis. There were no clinically important baseline differences in sex, race, self-report pain and function, body mass index, and randomization assignment between MeTeOR subjects included versus those not included in this analysis (Supplementary Table 1). The subjects included in the analysis had somewhat greater radiographic severity (79% vs. 65% of those not included had KL grade ≥2).

Baseline characteristics by treatment are shown in Table 1. The analytic sample had mean ± SD age 59.1 ± 7.7 years, 39% were male, 87% White, and 21% had KL0 or KL1 radiographs. There were no meaningful differences between the APM and PT groups in baseline demographics, pain scores, or KL grade after inverse probability adjustment for the propensity score (Table 1). The mean ± SD baseline OARSI radiographic score was 3.8 ± 3.5 in the APM group and 4.0 ± 4.9 in the PT group after inverse probability weighting.

Results of the primary and sensitivity analyses are shown in Table 2. We highlight findings of the weighted analysis here. The OARSI score increased by a mean of 4.1 (95% CI 3.5–4.7) in the APM group and by 2.4 (95% CI 1.7–3.2) in the PT group ($P < 0.001$). The OARSI JSN component increased by 1.4 (95% CI 1.1–1.6) in the surgically treated group and 1.1 (95% CI 0.8–1.5) in the PT group ($P = 0.36$), whereas the OARSI osteophyte component increased 2.9 (95% CI 2.4–3.3) in the APM group and 1.4 (95% CI 1.0–1.9) in the PT group ($P < 0.001$). The OARSI score increased by at least 7 points (approximately 1 SD more than the cohort's mean progression) in 23% of the surgically treated group and 5% of the PT group ($P < 0.001$) (Table 2). The KL grade increased by a mean of 1.2 (95% CI 1.0–1.4) grades in the surgically treated group and 1.1 (95% CI 0.8–1.5) in the nonoperative group ($P = 0.87$). The

Table 1. Baseline features of subjects treated with APM and those treated with PT; unweighted and propensity-weighted samples*

	Unweighted			Weighted		
Subject features	APM (n = 100)	PT (n = 42)	*P*	APM (n = 100)	PT (n = 42)	*P*
Age (at enrollment), mean ± SD years	59.3 ± 8.3	58.6 ± 6	0.57	59.2 ± 9.9	58.6 ± 12.3	0.63
Male, n (%)	40 (40.0)	16 (38.1)	0.83	56 (39.3)	52 (37.6)	0.85
White, n (%)	88 (88.0)	35 (83.3)	0.49	123 (86.2)	118 (85.3)	0.89
BMI, mean ± SD	29.6 ± 6.3	30.4 ± 6.4	0.47	29.3 ± 7.4	30.7 ± 12.9	0.30
KL grade, mean ± SD	1.9 ± 1.1	1.6 ± 1.0	0.20	1.8 ± 1.3	1.7 ± 2.0	0.68
0, n (%)	14 (14.0)	7 (16.7)	0.70	21 (15.1)	23 (16.7)	0.81
1, n (%)	23 (23.0)	9 (21.4)	0.84	32 (22.4)	32 (23.2)	0.92
2, n (%)	32 (32.0)	21 (50.0)	0.05	53 (37.4)	53 (38.2)	0.93
3, n (%)	24 (24.0)	3 (7.1)	0.005	27 (18.9)	23 (16.8)	0.76
4, n (%)	7 (7.0)	2 (4.8)	0.60	9 (6.2)	7 (5.1)	0.79
OARSI score, mean ± SD†	3.9 ± 2.8	4.2 ± 2.8	0.54	3.8 ± 3.5	4.0 ± 4.9	0.77
KOOS pain, mean ± SD‡	46.4 ± 14.1	46.2 ± 17.8	0.94	45.5 ± 16.7	46.8 ± 31.5	0.68
WOMAC pain, mean ± SD§	40.7 ± 16.7	41.3 ± 18.8	0.86	39.7 ± 19.9	42.3 ± 33.9	0.44

* APM = arthroscopic partial meniscectomy; BMI = body mass index; KOOS = Knee Injury and Osteoarthritis Outcome Score; OARSI = Osteoarthritis Research Society International; PT = physical therapy; WOMAC = Western Ontario and McMaster Universities Arthritis Index.
† OARSI score possible range 0–18, with 18 being worse.
‡ KOOS Pain score possible range 0–100, with 100 being worse.
§ WOMAC Pain score possible range 0–100, with 100 being worse.

Table 2. Changes over 5 years in KL and OARSI scores in subjects treated with APM versus PT*

	Unadjusted, mean change (95% CI)			Inverse probability weighted, mean change (95% CI)		
Groups classified by treatment received	APM (n = 100)	PT (n = 42)	*P*	APM (n = 100)	PT (n = 42)	*P*
ΔKL	1.2 (1.0–1.4)	1.2 (0.9–1.6)	0.83	1.2 (1.0–1.4)	1.1 (0.8–1.5)	0.87
% with ΔKL ≥ 1	68.0 (58.9–77.1)	61.9 (47.2–76.6)	0.48	67.9 (60.2–75.6)	56.5 (48.2–64.7)	0.048
ΔOARSI	4.2 (3.6–4.8)	2.6 (1.9–3.4)	0.002	4.1 (3.5–4.7)	2.4 (1.7–3.2)	<0.001
% with ΔOARSI ≥ 7	23.0 (14.8–31.2)	4.8 (0.0–11.2)	0.009	22.7 (15.8–29.6)	4.9 (1.3–8.5)	<0.001
ΔOARSI JSN	1.4 (1.1–1.6)	1.2 (0.8–1.6)	0.49	1.4 (1.1–1.6)	1.1 (0.8–1.5)	0.36
ΔOARSI OST	2.9 (2.5–3.4)	1.6 (1.1–2.1)	<0.001	2.9 (2.4–3.3)	1.4 (1.0–1.9)	<0.001
Groups classified by randomization arm	APM (n = 77)	PT (n = 65)		APM (n = 77)	PT (n = 65)	
ΔKL	1.1 (0.9–1.4)	1.2 (1.0–1.5)	0.64	1.2 (0.9–1.4)	1.2 (0.9–1.5)	0.85
% with ΔKL ≥ 1	66.2 (55.7–76.8)	66.2 (54.7–77.7)	0.99	66.2 (58.4–73.9)	65.2 (57.4–73.1)	0.87
ΔOARSI	4.3 (3.6–5.0)	3.0 (2.3–3.7)	0.008	4.3 (3.7–5.0)	3.0 (2.3–3.7)	0.005
% with ΔOARSI ≥ 7	24.7 (15.0–34.3)	9.2 (2.2–16.3)	0.016	25.1 (18.0–32.2)	9.3 (4.6–14.1)	<0.001
ΔOARSI JSN	1.4 (1.1–1.7)	1.2 (0.9–1.5)	0.32	1.4 (1.1–1.7)	1.2 (0.9–1.5)	0.33
ΔOARSI OST	3.0 (2.5–3.5)	2.0 (1.5–2.5)	0.004	3.0 (2.5–3.5)	1.9 (1.4–2.4)	0.003
Groups classified by treatment received with TKRs excluded	APM (n = 77)	PT (n = 40)		APM (n = 77)	PT (n = 40)	
ΔKL	1.1 (0.9–1.4)	1.2 (0.9–1.6)	0.63	1.1 (0.9–1.4)	1.2 (0.8–1.6)	0.69
% with ΔKL ≥ 1	63.6 (52.9–74.4)	62.5 (47.5–77.5)	0.90	64.0 (55.3–72.7)	60.2 (51.3–69.1)	0.55
ΔOARSI	3.6 (2.9–4.3)	2.5 (1.8–3.3)	0.035	3.6 (2.9–4.2)	2.4 (1.7–3.2)	0.028
% with ΔOARSI ≥ 7	16.9 (8.5–25.3)	5.0 (0.0–11.8)	0.07	16.8 (10.0–23.6)	5.4 (1.3–9.4)	0.005
ΔOARSI JSN	1.1 (0.8–1.3)	1.2 (0.8–1.6)	0.73	1.1 (0.9–1.4)	1.1 (0.7–1.5)	0.92
ΔOARSI OST	2.6 (2.2–3.1)	1.5 (1.0–2.0)	0.002	2.6 (2.1–3.1)	1.5 (1.0–2.0)	0.002
Groups classified by randomization arm with TKRs excluded	APM (n = 61)	PT (n = 56)		APM (n = 61)	PT (n = 56)	
ΔKL	1.1 (0.8–1.4)	1.2 (0.9–1.5)	0.70	1.1 (0.8–1.4)	1.2 (0.9–1.5)	0.72
% with ΔKL ≥ 1	62.3 (50.1–74.5)	64.3 (51.7–76.8)	0.82	62.2 (53.4–71.0)	63.8 (55.1–72.5)	0.79
Δ OARSI	3.9 (3.1–4.6)	2.6 (1.9–3.2)	0.010	3.8 (3.1–4.6)	2.6 (1.9–3.2)	0.011
% with ΔOARSI ≥ 7	18.0 (8.4–27.7)	7.1 (0.4–13.9)	0.08	18.3 (11.3–25.3)	7.6 (2.8–12.4)	0.015
ΔOARSI JSN	1.1 (0.9–1.4)	1.1 (0.8–1.4)	0.79	1.2 (0.9–1.4)	1.1 (0.8–1.4)	0.78
ΔOARSI OST	2.8 (2.3–3.4)	1.6 (1.2–2.1)	0.002	2.8 (2.2–3.4)	1.6 (1.2–2.1)	0.002

* APM = arthroscopic partial meniscectomy; 95% CI = 95% confidence interval; JSN = joint space narrowing; KL = Kellgren-Lawrence; OARSI = Osteoarthritis Research Society International; OST = osteophyte; PT = physical therapy; TKR = total knee replacement.

proportion of subjects who advanced by at least one KL grade over follow-up was 68% in the APM group and 56% in the PT group (*P* = 0.048) (Table 2).

Sensitivity analyses that classified subjects according to the treatment they were randomized to rather than treatment received showed statistically significantly greater progression in those randomized to APM than in those randomized to PT in mean overall OARSI score (4.3 vs. 3.0) and in mean OARSI osteophyte score (3.0 vs. 1.9) (Table 2). The magnitude of the effect was smaller than that observed in the as-treated analysis. Similarly, the sensitivity analyses that excluded subjects who had TKR showed statistically significantly greater increase in OARSI score in the APM group due to greater worsening of the OARSI osteophyte score (Table 2).

DISCUSSION

We compared changes in radiographs over 5 years between MeTeOR subjects who received APM and those who were treated with PT. We found that, overall, OARSI radiographic scores increased more in the APM group than in the nonoperative group due to changes in the OARSI osteophyte score. We did not find statistically significant differences between the surgically treated and nonoperative groups in changes in the mean KL grade (the commonly used metric for OA structural damage) or in the OARSI JSN score.

Our findings align with those of Collins et al, who documented greater progression in MRI-defined cartilage defects, osteophytes, bone marrow lesions, and synovitis in MeTeOR trial participants treated with APM than in those treated nonoperatively (13,14). Sonesson et al documented that 3 years following randomization to APM versus exercise, 60% of APM-treated subjects had a KL grade of ≤2 compared with 37% of nonoperatively treated subjects (*P* = 0.06) (16). Sihvonen and colleagues demonstrated greater progression in OARSI score in subjects treated with bona fide APM than in those treated with sham (between group difference of 0.7 points; 95% CI 0.1–1.3) (15). These

authors did not examine the JSN and osteophyte components separately. In Sihvonen et al, 72% of APM-treated subjects advanced by at least one KL grade over 5 years, compared with 60% of those treated with sham surgery ($P = 0.16$) (15). Our findings (68% advancing in the APM group and 56% in the PT group) were similar to those of Sihvonen and colleagues. In randomized trials of APM versus PT, Herrlin et al and Yim et al documented lower rates of progression in radiographic score than observed in the studies of Sonesson et al and Sihvonen et al, with no statistically significant difference between APM and PT groups (17,18).

Our study demonstrated that surgically treated subjects had significantly greater advancement in osteophyte score compared with nonoperatively treated subjects, but no significant difference in JSN score. This observation may reflect inherent biologic differences in the processes of JSN and osteophyte development, or alternatively greater measurement responsiveness due to the greater potential range of osteophyte scores (0–12) than joint space narrowing scores (0–6) in the OARSI grading rubric.

Importantly, the clinical significance of the differences we observed is unknown. In a prior study, we showed that changes over 18 months in MRI scores for cartilage damage, osteophytes, effusion synovitis, and bone marrow lesions had no clinically significant associations with changes over the subsequent 42 months in KOOS Pain or KOOS Function scores (25). Further follow-up of the MeTeOR cohort will reveal whether the changes we observed in this study in osteophyte scores are associated with subsequent pain, functional limitation, risk of TKR, or other clinically relevant endpoints.

This study has important limitations. First, fewer than 50% of subjects had both baseline and 5-year radiographs. Although subjects with and without follow-up radiographs had similar demographic and clinical features at baseline, we acknowledge the risk of bias that may arise from drop-out. The baseline radiographs were clinically ordered weight-bearing films but were not performed with frames (such as Synaflexer) to ensure a standard semiflexed view. This factor may have introduced misclassification of JSN, biasing the differences in radiographic progression toward the null. Although the study was performed in multiple centers, the largely White racial composition of the trial cohort may limit generalizability of these findings. We recognize that osteophytes may occur below the joint line, and we did not assess the location of the osteophytes relative to the joint line. Finally, although propensity weighting eliminated differences between treatment groups in measured variables, we cannot exclude residual confounding by unmeasured variables.

We conclude that APM is associated with greater advancement in osteophyte scores—but not in JSN—compared to nonoperative therapy among persons with knee pain, meniscal tear, and osteoarthritic imaging findings. The clinical significance of these results is unknown; thus, we encourage further research on the long-term clinical correlates of radiographic changes following APM.

ACKNOWLEDGMENTS

The authors thank Lauren Margaret Mitchell and Mahima Kumara for editorial assistance.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Katz, Collins, Brophy, Guermazi, Jones, Levy, MacFarlane, Mandl, Marx, Selzer, Spindler, Losina, Chang.

Acquisition of data. Katz, Brophy, Cole, Cox, Guermazi, Jones, Levy, MacFarlane, Mandl, Marx, Selzer, Spindler, Wright, Losina.

Analysis and interpretation of data. Katz, Collins, Brophy, Cole, Cox, Guermazi, Jones, Levy, MacFarlane, Mandl, Marx, Selzer, Spindler, Wright, Losina, Chang.

REFERENCES

1. Englund M, Guermazi A, Gale D, et al. Incidental meniscal findings on knee MRI in middle-aged and elderly persons. N Engl J Med 2008; 359:1108–15.
2. Bhattacharyya T, Gale D, Dewire P, et al. The clinical importance of meniscal tears demonstrated by magnetic resonance imaging in osteoarthritis of the knee. J Bone Joint Surg Am 2003;85:4–9.
3. Berthiaume MJ, Raynauld JP, Martel-Pelletier J, et al. Meniscal tear and extrusion are strongly associated with progression of symptomatic knee osteoarthritis as assessed by quantitative magnetic resonance imaging. Ann Rheum Dis 2005;64:556–63.
4. Australian Knee Society. Position statement from the Australian Knee Society on arthroscopic surgery of the knee, including reference to the presence of osteoarthritis or degenerative joint disease: updated October 2016. Orthop J Sports Med 2017;5:2325967117728677.
5. Arthroscopy Association of Canada, Wong I, Hiemstra L, et al. Position statement of the Arthroscopy Association of Canada (AAC) concerning arthroscopy of the knee joint—September 2017. Orthop J Sports Med 2018;6:2325967118756597.
6. Beaufils P, Becker R, Kopf S, et al. Surgical management of degenerative meniscus lesions: the 2016 ESSKA meniscus consensus. Joints 2017;5:59–69.
7. American Academy of Orthopaedic Surgeons. Management of osteoarthritis of the knee (non-arthroplasty): evidence-based clinical practice guideline. URL: https://www.aaos.org/oak3cpg.
8. Kim S, Bosque J, Meehan JP, et al. Increase in outpatient knee arthroscopy in the United States: a comparison of National Surveys of Ambulatory Surgery, 1996 and 2006. J Bone Joint Surg Am 2011;93:994–1000.
9. Howard DH. Trends in the use of knee arthroscopy in adults. JAMA Intern Med 2018;178:1557–8.
10. Jacquet C, Pujol N, Pauly V, et al. Analysis of the trends in arthroscopic meniscectomy and meniscus repair procedures in France from 2005 to 2017. Orthop Traumatol Surg Res 2019;105:677–82.
11. Chung KS, Ha JK, Kim YS, et al. National trends of meniscectomy and meniscus repair in Korea. J Korean Med Sci 2019;34:e206.
12. Rongen JJ, van Tienen TG, Buma P, et al. Meniscus surgery is still widely performed in the treatment of degenerative meniscus tears in The Netherlands. Knee Surg Sports Traumatol Arthrosc 2018;26: 1123–9.
13. Collins JE, Losina E, Marx RG, et al. Early magnetic resonance imaging-based changes in patients with meniscal tear and

osteoarthritis: eighteen-month data from a randomized controlled trial of arthroscopic partial meniscectomy versus physical therapy. Arthritis Care Res (Hoboken) 2020;72:630–40.

14. Collins JE, Shrestha S, Losina E, et al. Five-year structural changes in the knee among patients with meniscal tear and osteoarthritis: data from a randomized controlled trial of arthroscopic partial meniscectomy versus physical therapy. Arthritis Rheumatol 2022;74:1333–42.
15. Sihvonen R, Paavola M, Malmivaara A, et al. Arthroscopic partial meniscectomy for a degenerative meniscus tear: a 5 year follow-up of the placebo-surgery controlled FIDELITY (Finnish Degenerative Meniscus Lesion Study) trial. Br J Sports Med 2020;54:1332–9.
16. Sonesson S, Kvist J, Yakob J, et al. Knee arthroscopic surgery in middle-aged patients with meniscal symptoms: a 5-year follow-up of a prospective, randomized study. Orthop J Sports Med 2020;8: 2325967119893920.
17. Yim JH, Seon JK, Song EK, et al. A comparative study of meniscectomy and nonoperative treatment for degenerative horizontal tears of the medial meniscus. Am J Sports Med 2013;41:1565–70.
18. Herrlin SV, Wange PO, Lapidus G, et al. Is arthroscopic surgery beneficial in treating non-traumatic, degenerative medial meniscal tears? A five year follow-up. Knee Surg Sports Traumatol Arthrosc 2013;21: 358–64.
19. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis 1957;16:494–502.
20. Altman RD, Gold GE. Atlas of individual radiographic features in osteoarthritis, revised. Osteoarthritis Cartilage 2007;15 Suppl A:A1–56.
21. Katz JN, Chaisson CE, Cole B, et al. The MeTeOR trial (Meniscal Tear in Osteoarthritis Research): rationale and design features. Contemp Clin Trials 2012;33:1189–96.
22. Katz JN, Brophy RH, Chaisson CE, et al. Surgery versus physical therapy for a meniscal tear and osteoarthritis. N Engl J Med 2013;368: 1675–84.
23. Guermazi A, Hunter DJ, Roemer FW. Plain radiography and magnetic resonance imaging diagnostics in osteoarthritis: validated staging and scoring. J Bone Joint Surg Am 2009;91 Suppl 1:54–62.
24. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivariate Behav Res 2011;46:399–424.
25. Katz JN, Collins JE, Jones M, et al. Association between structural change over eighteen months and subsequent symptom change in middle-aged patients treated for meniscal tear. Arthritis Care Res (Hoboken) 2023;75:340–7.

Arthritis Care & Research
Vol. 77, No. 3, March 2025, pp 366–375
DOI 10.1002/acr.25186

American College of Rheumatology
Empowering Rheumatology Professionals

Association of Foot Symptoms With Decreased Time to All-Cause Mortality: The Johnston County Osteoarthritis Project

Skylar Harmon,[1] Carolina Alvarez,[2] Marian T. Hannan,[3] Leigh F. Callahan,[2] Lucy S. Gates,[4] Catherine J. Bowen,[5] Hylton B. Menz,[6] Amanda E. Nelson,[2] and Yvonne M. Golightly[7]

Objective. Adults with foot symptoms (ie, pain, aching, or stiffness) may be at increased risk of reduced time to all-cause mortality. The purpose of this study was to evaluate whether foot symptoms are independently associated with all-cause mortality in older adults.

Methods. We analyzed longitudinal data from 2613 participants from the Johnston County Osteoarthritis Project, a longitudinal population-based cohort of adults 45 years of age and older. Participants completed questionnaires at baseline to determine presence of foot symptoms and covariable status. Baseline walking speed was measured via an 8-foot walk test. To examine the association of foot symptoms with time to mortality, hazard ratios (HRs) and 95% confidence intervals (CIs) were calculated using Cox regression models, adjusted for potential confounders.

Results. We observed 813 deaths over 4 to 14.5 years of follow-up. At baseline, 37% of participants had foot symptoms, mean age was 63 years, mean body mass index was approximately 31 kg/m^2, 65% were women, and 33% were Black. Moderate to severe foot symptoms were associated with reduced time to mortality after adjustment for demographics, comorbidities, physical activity, and knee and hip symptoms (HR = 1.30, 95% CI 1.09–1.54). Importantly, this association was not modified by walking speed or diabetes.

Conclusion. Individuals with foot symptoms had an increased hazard of all-cause mortality compared with those with no foot symptoms. These effects were independent of key confounders and were not moderated by walking speed. Effective interventions to identify and manage at least moderate foot symptoms may reduce the risk of decreased time to mortality.

INTRODUCTION

As many as one in three middle-aged to older adults have foot symptoms (ie, pain, aching, or stiffness on most days) (1,2), with greater prevalence among women, persons who identify as Black, those who are obese, those of older age, and those with routine/manual occupations (3). Foot symptoms are associated with decreased physical function and disability, even when controlling for important covariates (1,4–6). Foot symptoms pose a burden that likely affects locomotor function and participation in daily physical, occupational, and social activities, making it a significant public health concern (7,8).

Previous studies have highlighted the association of knee or hip symptoms with all-cause mortality. In 2018, Cleveland et al examined the impact of knee symptoms on excess mortality using data from the Johnston County Osteoarthritis Project (JoCoOA) (9). The investigators found that knee symptoms in the presence or absence of radiographic osteoarthritis (OA) were associated with an increase in all-cause mortality of greater than 15%. Among patients with knee or hip symptoms with OA,

Supported by the CDC/Association of Schools of Public Health (grants S-043, S-1734, S-3486, and U01-DP-003206), the NIH/National Institute of Arthritis and Musculoskeletal and Skin Diseases (grants P60-AR-30701, P60-AR-049465, P60-AR-064166, P30-AR-072580, and R01-AR-065937), the 2022 Medical Students in Aging Research program at University of North Carolina at Chapel Hill through the NIH/National Institute of Aging (grant 5T35-AG-038047).

[1]Skylar Harmon, BA: University of North Carolina at Chapel Hill, Chapel Hill, and Nova Southeastern University and Dr. Kiran C. Patel College of Allopathic Medicine, Fort Lauderdale, Florida; [2]Carolina Alvarez, MS, Leigh F. Callahan, PhD, Amanda E. Nelson, MD, MPH: University of North Carolina at Chapel Hill, Chapel Hill; [3]Marian T. Hannan, DSc, MPH: Harvard T. H. Chan School of Public Health, Beth Israel Deaconess Medical Center, and Harvard Medical School, Boston, Massachusetts; [4]Lucy S. Gates, PhD: University of Southampton and Southampton General Hospital, Southampton, UK; [5]Catherine J. Bowen, PhD: University of Southampton, Southampton, UK; [6]Hylton B. Menz, PhD, DSc: La Trobe University, Bundoora, Victoria, Australia; [7]Yvonne M. Golightly, PT, PhD: University of North Carolina at Chapel Hill, Chapel Hill, and University of Nebraska Medical Center, Omaha.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25186.

Address correspondence via email to Yvonne M. Golightly, PT, PhD, at ygolightly@unmc.edu.

Submitted for publication January 9, 2023; accepted in revised form June 27, 2023.

SIGNIFICANCE & INNOVATIONS

- Foot symptoms (ie, pain, aching, stiffness) are common among middle-aged to older adults, limiting physical function, physical activity, and quality of life.
- Symptoms at the knee and hip are linked to reduced time to all-cause mortality, but the role of foot symptoms in mortality has not been established.
- The present study provided novel results demonstrating that moderate to severe foot symptoms were related to a higher hazard of all-cause mortality, and this relationship was independent of walking speed, sex, race, obesity status, or diabetes status.
- Strategies to prevent and treat moderate to severe foot symptoms may improve physical function and activity, thus affecting mortality outcomes.

Nüesch et al reported excess all-cause mortality [standardized mortality ratio 1.55, 95% confidence interval (CI) 1.41–1.70] (10). Several other studies have also shown associations between knee symptom presence and 35% to 37% reduced time to mortality (11). Additionally, Cleveland et al reported an increased hazard of 1.3 for all-cause mortality in participants with hip symptoms without radiographic OA compared with participants with neither radiographic OA nor symptoms at the hip [adjusted hazard ratio (aHR) = 1.28, 95% CI 1.13–1.46] (12).

Walking speed is an important metric for evaluating current health and future health outcomes (13). Master et al found that adults with symptomatic knee radiographic OA (Kellgren-Lawrence grade ≥2) had increased risk of mortality and that walking speed modified this relationship (14). Specifically, slower walking speeds measured at both short (2.4-meter) and standard (20-meter) distances were associated with higher mortality [aHR (95%CI) 1.23 (1.10–1.39) and 1.25 (1.09–1.43), respectively] (14). This suggests that impaired mobility resulting from symptoms may be a cause of excess mortality independent of age and comorbid conditions. Although these studies highlight the potential association of knee and/or hip symptoms with mortality, the association of foot symptoms with mortality has not been explored in detail.

Foot symptoms may contribute to less physical activity and loss of physical function, which over time could lead to factors that impact mortality, including comorbid conditions from increasing body mass or falls from muscle weakness or impaired balance. One prior study has evaluated the association of foot symptoms with self-reported and performance-based measures of physical function in the JoCoOA cohort (1), finding that the presence of foot symptoms was significantly associated with worsened mobility (slower 8-foot walk time) irrespective of knee and hip symptoms and OA. Because of the significant impact of foot symptoms on mobility, it is important to consider whether the presence of foot symptoms alone leads to reduced time to mortality. Further, there may be individuals with foot symptoms who are at a higher risk of mortality, such as those who walk at slower speeds (13), and assessment of effect modification could assist with identifying these subgroups.

Thus, the purpose of this study was to determine whether foot symptoms are associated with reduced time to all-cause mortality, and if so, whether this association is modified by reduced walking speed using the JoCoOA cohort. Our hypotheses were that foot symptoms would be linked to reduced time to mortality and that slower walking speed would amplify the magnitude of this association.

METHODS

Study participants

Participants were from the JoCoOA cohort, a longitudinal, community-based study of the occurrence of OA in Black and White civilian, noninstitutionalized adults 45 years or older who resided in a mostly rural county in North Carolina, USA. Detailed descriptions of JoCoOA eligibility criteria have been published in other literature (15). Briefly, participants in the Original Cohort were recruited during 1991–1997, and they completed follow-up visits in 1999–2004, 2006–2010, and 2013–2015. Additional participants were enrolled during 2003–2004 (Enrichment Cohort), and they completed follow-up visits in 2006–2010 and 2013–2015.

Questions about foot symptoms were not added to JoCoOA until 1999. Thus, the present study only included those Original Cohort and Enrichment Cohort participants who attended a clinic visit during 1999–2004 and had available foot symptoms data. First and second follow-up visits for the present study were defined as 2006–2010 and 2013–2015, respectively. The JoCoOA has been continuously approved by the University of North Carolina Institutional Review Board (#92-0583).

Study outcome

All-cause mortality: Excluding those who had died prior to baseline, time to all-cause mortality was quantified from the date of the initial foot examination visit (1999–2004) to the date of death. All participants had vital status assessed at each follow-up time point. Deaths were primarily found through the National Death Index (NDI) records, although some known deaths, if not found in the NDI, were confirmed through the Johnston County Register of Deeds office.

Study exposure

Foot symptoms. Foot symptoms were considered present if a participant responded "yes" to the question: "On most days,

do you have pain, aching, or stiffness in your left/right foot?" The time framing of this question varied. At baseline participants were asked to recall over the past month, at first follow-up (T2) over the past year, and at second follow-up (T3) in any one month of the past year. These symptoms were queried by presence in either foot (ie, yes or no), and also by any foot symptom laterality (ie, none, unilateral, or bilateral). Finally, symptom severity in either foot was assessed on the following scale: 0 = none, 1 = mild, 2 = moderate, and 3 = severe. The scale was used to create a severity count summed across both feet (range 0–6).

Potential confounders

Short distance walk test. At the baseline examination, walking speed was assessed from an 8-foot walk test at usual pace over an unobstructed course. The same testing was used at follow-up exams. At each examination, time was measured using a digital stopwatch and recorded to the nearest tenth of second in two trials over the 8-foot distance. Walking speed was calculated as the total distance 8 feet (2.4 meters) divided by the time to complete the walk test; the average walking speed of the two trials was calculated and used in analyses. Among older adults, the 8-foot walk test has been shown to have fair to good retest reliability (intraclass correlation coefficients >0.5) for assessing walking speed (16,17). Walking speed was grouped into categories of 1.0 m/s or faster, 0.8 to less than 1 m/s, 0.4 to less than 0.8 m/s, and less than 0.4 m/s or unable, a scale adapted from (18).

We considered the following static variables to be potential confounders because they may be associated with foot conditions and with mortality: self-reported race (Black or White), sex (female vs. male), age (in years), and education (categorized as <12 years vs. ≥12 years of school). Additional time-varying covariables included clinically measured body mass index (BMI) (in kg/m^2), ever smoker (self-reported, yes/no), any alcohol use (self-reported, yes/no), meeting the US Department of Health and Human Services guidelines for moderate/vigorous physical activity (MVPA) equal to or more than 150 minutes/week (self-reported, yes/no), current nonsteroidal anti-inflammatory drug (NSAID) use (presentation of medication container by participant or self-report, yes/no), and any report of symptoms in the hips or knees. Comorbidities assessed by self-report included history of cardiovascular disease (CVD), hypertension (HTN), liver disease, depression, and cancer, analyzed as reporting at least one of five of the listed comorbidities. Diabetes mellitus was separated from the overall comorbidity count because of its effects on foot symptoms (eg, peripheral neuropathy). Because calendar effects could impact the outcome of mortality and other risk factors in these analyses, models were stratified by decade of birth cohort.

The following variables were considered to be potential effect modifiers because of their differences affecting mortality or foot symptoms. Sex was tested for effect modification because previous studies have shown an increased incidence of foot symptoms in females compared with males (2,3). Race was tested because Black participants of JoCoOA have been shown to have higher rates of foot symptoms (19). Obesity was also considered because it has been associated with presence of nonspecific foot pain in the general population (20). Diabetes is not only implicated in foot symptoms through peripheral neuropathy but also has been shown to be a predictor of excess mortality (10,21). Walking speed was the main effect modifier we were interested in investigating because more severe walking disability has been associated with higher risk of mortality in previous studies (10).

Statistical analysis

At baseline and at follow-up time points, descriptive statistics were calculated. Continuous variables were described using means and SDs (±SD), and categorical variables were presented as frequencies and percentages. All tests were two-sided, and statistical significance was set at the 0.05 level. All analyses were conducted using SAS software version 9.4 (Cary, NC).

Participant information from baseline, first follow-up, and second follow-up was included for the above-described covariables and foot symptom definitions. Follow-up time was calculated as the time difference between baseline and confirmed death or censoring (ie, loss to follow-up or end of study on December 31, 2015).

Survival curves. We used Kaplan-Meier methods to generate nonparametric survival curves by baseline foot symptoms status strata, and the log-rank test was used to test difference by strata. We examined the association of foot symptoms over time with all-cause mortality by calculating aHRs and 95% CIs using time-dependent Cox proportional hazards regression employing the counting process to include time-varying covariables. Each model was adjusted for potential confounders.

Models. Models are presented for each of the six aspects of our study foot symptom definitions, along with three aspects of model building with covariables. Model 1 adjusted for demographics including birth cohort (as a stratum by decade), enrollment wave (original or enrichment), age, sex, race and ethnicity, and education. Model 2 adjusted for covariables in model 1 as well as comorbidities and relevant clinical risk factors including NSAIDs, smoking, alcohol use, meeting MVPA guidelines, BMI, diabetes, at least one comorbidity out of five comorbidities, knee symptoms, and hip symptoms. Model 3 adjusted for covariables in model 2 and walking speed categories.

Missing data. Covariable information for at least one measure was missing for 6.2% of participants (Figure 1). Multiple imputation was used to impute missing values with missing information assumed to be missing at random. Logistic and linear regression models were used to impute binary and continuous covariable

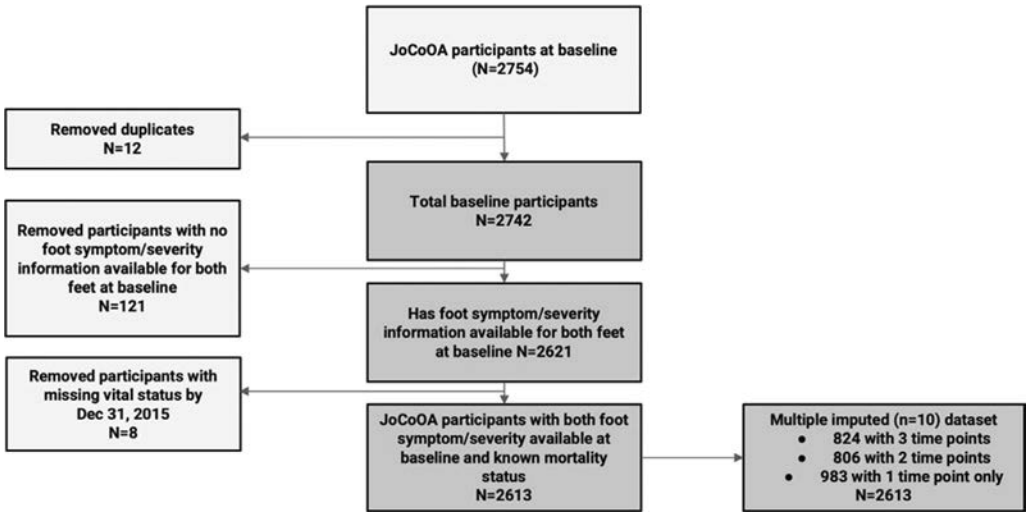


Figure 1. Analytic sample size for JoCoOA participants with baseline foot symptoms data and known mortality status. Abbreviation: JoCoOA = Johnston County Osteoarthritis Project.

Table 1. Baseline demographic characteristics, the Johnston County Osteoarthritis Project*

	Baseline (1999–2004; n = 2,613)	First follow-up (2006–2010; n = 1,604)	Second follow-up (2013–2015; n = 850)
Demographics at baseline			
Enrichment 2003–2004 cohort, n (%)	999 (38.2)	572 (35.7)	320 (37.6)
Age, years, mean ± SD	63.4 ± 10.5	68.4 ± 9.1	71.4 ± 7.7
Female sex, n (%)	1,704 (65.2)	1,069 (66.6)	574 (67.5)
Black, n (%)	856 (32.8)	492 (30.7)	280 (32.9)
<12 years education, n (%) (missing n = 12)	716 (27.4)	341 (21.3)	119 (14.0)
Time-varying variables			
BMI (kg/m^2), mean ± SD (missing n = 4)	30.6 ± 6.7	31.3 ± 7.2	31.0 ± 6.6
Ever smoker, n (%) (missing n = 121)	1,058 (40.5)	865 (53.9)	462 (54.4)
Any alcohol use, n (%) (missing n = 95)	957 (36.6)	658 (41.0)	348 (40.9)
≥150 MVPA min/week, n (%) (missing n = 3)	791 (30.3)	341 (21.3)	145 (17.1)
NSAID use, n (%) (missing n = 1)	1,259 (48.2)	1,082 (67.5)	584 (68.7)
HTN, n (%)	1,265 (48.4)	1,087 (67.8)	690 (81.2)
CVD, n (%) (missing = 1)	575 (22.0)	566 (35.3)	364 (42.8)
Diabetes, n (%)	425 (16.3)	383 (23.9)	253 (29.8)
Depression, n (%)	344 (13.2)	187 (11.7)	101 (11.9)
Liver disease, n (%)	36 (1.4)	32 (2.0)	22 (2.6)
Cancer, n (%)	28 (1.1)	41 (2.6)	81 (9.5)
Five comorbidity count: HTN, CVD, depression, liver disease, cancer, n (%) (missing n = 1)			
0	985 (37.7)	331 (20.6)	101 (11.9)
1	1,100 (42.1)	734 (45.8)	337 (39.6)
2	442 (16.9)	446 (27.8)	325 (38.2)
3	78 (3.0)	86 (5.4)	77 (9.1)
4–5	7 (0.2)	7 (0.5)	10 (1.2)
Any knee symptoms, n (%) (missing n = 2)	1,321 (50.6)	608 (37.9)	325 (38.2)
Any hip symptoms, n (%) (missing n = 3)	1,005 (38.5)	474 (29.6)	288 (33.9)
Walking speed (m/s), mean ± SD (missing n = 10)	0.7 ± 0.3	0.7 ± 0.2	0.9 ± 0.3
Walking speed groups, n (%)			
1.0 m/s or better	319 (12.2)	188 (11.7)	226 (26.6)
0.8 to <1 m/s	633 (24.2)	361 (22.5)	276 (32.7)
0.4 to <0.8 m/s	1,430 (54.7)	874 (54.6)	305 (35.9)
<0.4 m/s or unable	221 (8.5)	177 (11.1)	41 (4.8)

* Abbreviations: BMI = body mass index; CVD = cardiovascular disease; HTN = hypertension; MVPA = moderate/vigorous physical activity (yes/no); NSAID = nonsteroidal antiinflammatory drugs (yes/no).

information, respectively, by fully conditional specification methods. These methods are optimal when data are missing at random and the proportion missing is less than 50%. Ten imputed data sets were generated so that the number of imputations were similar to the percentage of data missing one or more covariables. Separate analyses were carried out in each of the 10 imputed data sets; then estimated parameters from all imputed data sets were pooled to generate a single estimate according to Rubin's rules (22).

For model 3, effect modification of the associations of foot symptoms with mortality was considered for walking speed, in addition to sex, race, obesity, and diabetes. An interaction term was used between each foot symptom definition and each potential effect modifier and considered significant if *P* was less than 0.10. If statistically significant, the HR between foot symptom and mortality was shown by level of effect modifier.

For model 3, we also generated adjusted survival curves for foot symptoms definitions that were statistically significantly associated with mortality showing the survival experience of an average participant in the population from which JoCoOA was sampled.

RESULTS

Figure 1 provides a summary of the final analytic sample for the current analyses. Initially, 2,754 participants were identified from baseline; after removal of duplicate IDs (n = 12), missing foot symptom/severity information for both feet (n = 121), and missing vital status (n = 8), 2,613 participants were available for analysis. Those with missing baseline covariables (n = 162), were not excluded. Three time points of data were available for 824 participants, two time points for 806 participants, and one time point for 983 participants (Figure 1).

At baseline, of the 2,613 participants, the mean age was 63 (range 45–102) years, the mean BMI was approximately 31 kg/m^2, and nearly half of the participants (42%) had at least one comorbidity (Table 1). Of these participants, approximately one third were Black and two thirds were women. Over half of participants (55%) had walking speeds between 0.4 and less than 0.8 m/s.

At first follow-up (n = 1,604) and at the second follow-up (n = 850), BMI remained around 31 kg/m^2. Most participants at the first follow-up continued to have one comorbidity count (46%), whereas those at the second follow-up had one to two comorbidities (40% and 38%, respectively), with the presence of HTN and CVD increasing across time. At the second follow-up, remaining participants had the fastest walking speeds with 27% at 1.0 or better, 32% at 0.8 to less than 1 m/s, and 36% at 0.4 to less than 0.8 m/s.

At baseline, 37% of participants had any foot symptoms, whereas 25% and 21% had any foot symptoms at first and second follow-ups, respectively (Table 2). This trend continues with

Table 2. Baseline foot symptom/severity categories, the Johnston County Osteoarthritis Project

	Baseline (1999–2004; n = 2,613)	First follow-up (2006–2010; n = 1,604)	Second follow-up (2013–2015; n = 850)
Foot symptoms/severity variables			
Any foot symptoms, n (%)	969 (37.1)	401 (25.0)	175 (20.6)
Any foot symptoms laterality, n (%)			
none	1644 (62.9)	1203 (75.0)	675 (79.4)
unilateral	198 (7.6)	111 (6.9)	44 (5.2)
bilateral	771 (29.5)	290 (18.1)	131 (15.4)
At least moderate severity foot symptoms, n (%)	637 (24.4)	255 (15.9)	121 (14.2)
At least moderate severity foot symptoms laterality, n (%)			
none	1976 (75.6)	1349 (84.1)	729 (85.8)
unilateral	167 (6.4)	84 (5.2)	45 (5.3)
bilateral	470 (18.0)	171 (10.7)	76 (8.9)
Foot symptoms' worst severity from either foot, n (%)			
none (0)	1644 (62.9)	1207 (75.2)	675 (79.4)
mild (1)	332 (12.7)	142 (8.9)	54 (6.4)
moderate (2)	409 (15.7)	181 (11.3)	45 (5.3)
severe (3)	228 (8.7)	74 (4.6)	76 (8.9)
Foot symptoms' severity count for both feet, n (%)			
0	1644 (62.9)	1207 (75.2)	675 (79.4)
1	68 (2.6)	53 (3.3)	18 (2.1)
2	360 (13.8)	128 (8.0)	42 (4.9)
3	61 (2.3)	35 (2.2)	35 (4.1)
4	296 (11.3)	132 (8.2)	28 (3.3)
5	21 (0.8)	9 (0.6)	8 (0.9)
6	163 (6.2)	40 (2.5)	44 (5.2)

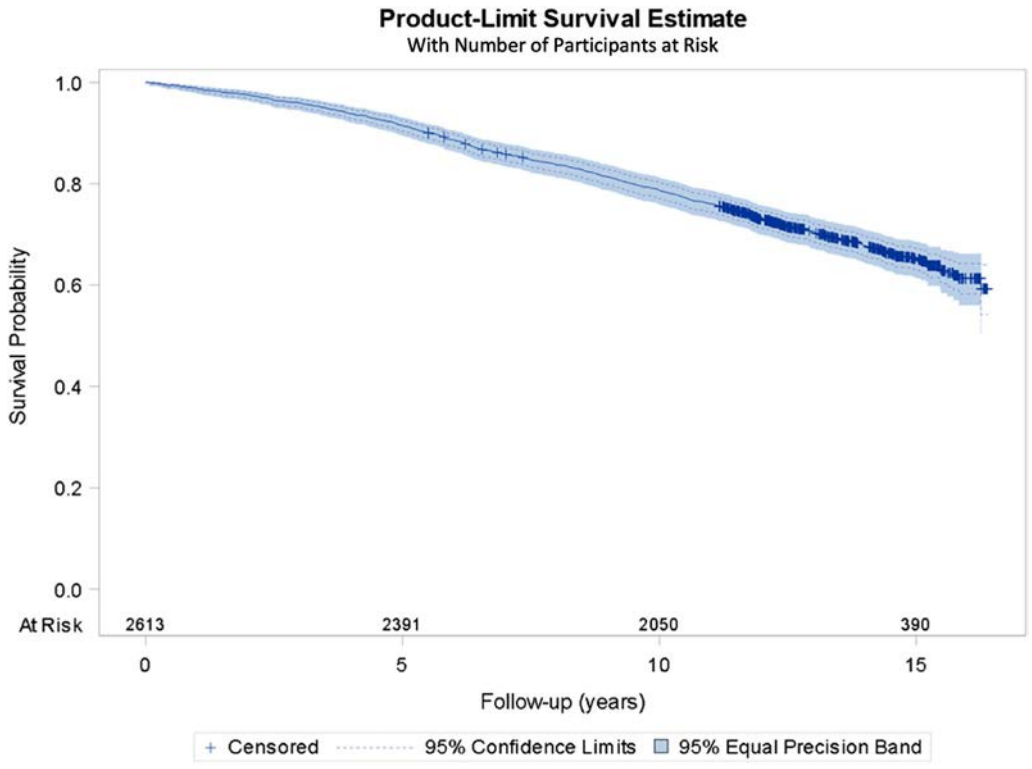
moderate foot symptoms. By first and second follow-ups, only 16% and 14% of participants remained with moderate foot symptoms compared with 24% at baseline. We also observed decreases in foot symptom severity over time, with 16% and 9% of participants experiencing moderate or severe foot symptoms, respectively at baseline. This is compared with 11% and 5% at first follow-up and 5% and 9% at second follow-up. Finally, we observed this same trend among those with bilateral moderate foot symptoms but not in those with unilateral moderate foot symptoms. At baseline, 18% of participants with moderate foot symptom severity experienced symptoms across both feet; at first and second follow-ups, this decreased to 11% and 9%, respectively.

An unadjusted Kaplan-Meier survival curve is presented in Figure 2. Although we would typically report median survival time, 69% of participants were censored, so we reported the quartile time to death instead (ie, the time when the first 25% of the sample had died). The curve shows those with foot symptoms had quartile time to death (0.75) of 11.4 years (95% CI 10.7–11.9) in which we observed 818 deaths overall (31% of JoCoOA).

In Figure 3, a Kaplan-Meier plot shows baseline moderate pain, aching, or stiffness (2) having a statistically significant difference from no or mild foot symptoms (1) ($P = 0.0002$). Moderate to severe foot symptoms show worse survival compared with mild or no symptoms.

Table 3 shows the aHRs from the time-dependent Cox proportional hazards models for the association of each of the six foot symptom definitions with time to death. For most of our foot symptom definitions, the effect of foot symptoms on hazard of death was independent of demographics (model 1), of risk factors and comorbidities (model 2), and of walking speed (model 3). The additional covariables did attenuate the HRs somewhat, but most definitions remained statistically significant between foot symptoms and death over follow-up—particularly those involving severity.

We examined effect modification with the main hypothesis that slower walking speed would modify the association seen between foot symptoms and time to death. Although slower walking speed significantly increased the hazard of mortality



Summary of the Number of Censored and Uncensored Values

Total	Failed	Censored	Percent Censored
2613	818	1795	68.69

Quarter time to death= 11.4 (95% CI 10.7, 11.9) years

Figure 2. Unadjusted Kaplan-Meier survival curve reporting quartile time to death. Abbreviation: CI = confidence interval.

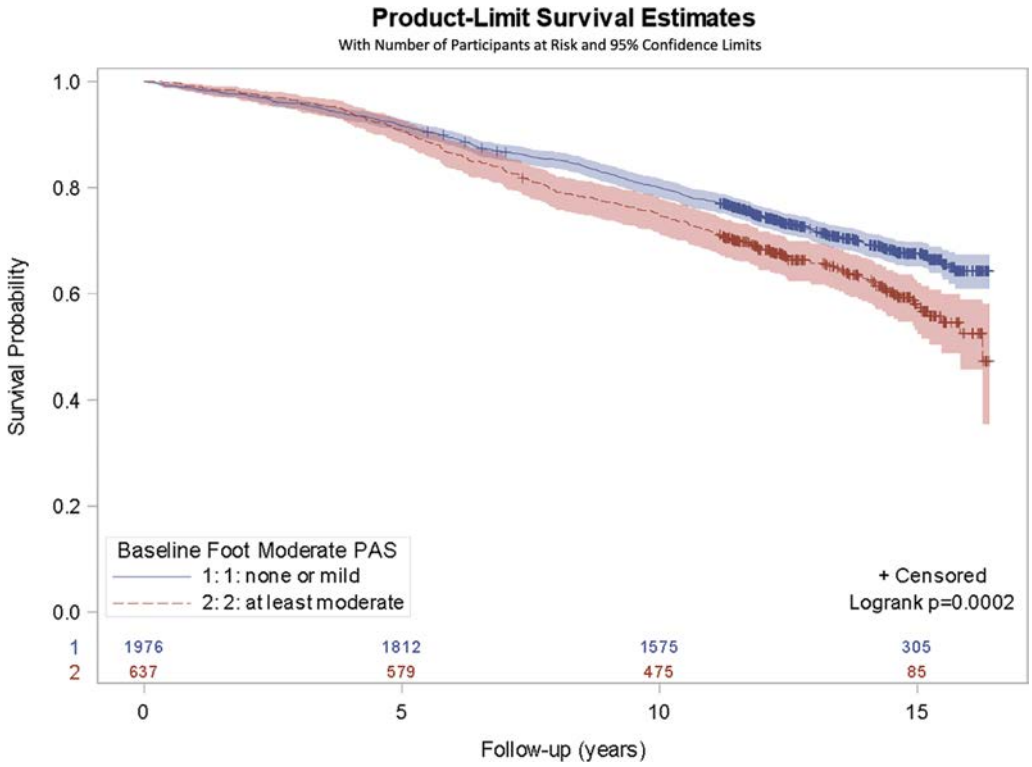


Figure 3. Kaplan-Meier plot showing baseline moderate PAS (2) and no or mild foot PAS (1). PAS = pain, aching, stiffness. Color figure can be viewed in the online issue, which is available at http://onlinelibrary.wiley.com/doi/10.1002/acr.25186/abstract.

Table 3. Adjusted HRs and 95% CIs for the association between foot symptoms and all-cause mortality over follow-up*

	Model by included covariables		
	Model 1	Model 2	Model 3
Model by foot symptom definition	HR (95% CI)	HR (95% CI)	HR (95% CI)
Any foot symptoms			
No foot symptoms	ref	ref	ref
Any foot symptoms	1.30 (1.12–1.50)	1.15 (0.99–1.35)	1.11 (0.95–1.29)
Laterality of foot symptoms			
No foot symptoms	ref	ref	ref
Any foot symptoms unilaterally	1.15 (0.89–1.51)	1.08 (0.83–1.41)	1.04 (0.79–1.36)
Any foot symptoms bilaterally	1.34 (1.15–1.57)	1.18 (1.00–1.40)	1.13 (0.95–1.34)
Severity of foot symptoms (grouped)			
None or mild severity foot symptoms	ref	ref	ref
At least moderate severity foot symptoms	1.52 (1.29–1.78)	1.34 (1.13–1.59)	1.30 (1.09–1.54)
Laterality of foot symptom severity			
None or mild severity foot symptoms	ref	ref	ref
At least moderate severity foot symptoms unilaterally	1.43 (1.09–1.89)	1.30 (0.98–1.72)	1.29 (0.97–1.71)
At least moderate severity foot symptoms bilaterally	1.55 (1.30–1.85)	1.36 (1.12–1.64)	1.30 (1.07–1.57)
Severity of foot symptoms			
No foot symptom severity	ref	ref	ref
Mild foot symptom severity	0.96 (0.75–1.22)	0.90 (0.71–1.15)	0.86 (0.68–1.10)
Moderate foot symptom severity	1.35 (1.11–1.64)	1.20 (0.98–1.47)	1.16 (0.95–1.43)
Severe foot symptom severity	1.86 (1.47–2.35)	1.58 (1.23–2.03)	1.48 (1.15–1.90)
Summed severity count of foot symptoms			
Symptom severity count for both feet (1 unit increase)	1.10 (1.06–1.15)	1.07 (1.03–1.12)	1.06 (1.02–1.10)

* Data used are multiply imputed (n = 10). Time-dependent Cox proportional hazard modeling time to death using the counting process for time-varying covariables. Model 1 (demographics): adjusted for birth cohort (strata), enrollment wave, age, sex, race, and education. Model 2 (comorbidities and relevant clinical risk factors): adjusted for model 1 + NSAIDs, smoking, alcohol use, MVPA, BMI, diabetes, five comorbidity count, knee PAS, and hip PAS. Model 3 (gait speed): adjusted for model 2 + gait speed categories.
Abbreviations: BMI = body mass index; CI = confidence interval; HR = hazard ratio; MVPA = moderate/vigorous physical activity; NSAID = nonsteroidal antiinflammatory drug; PAS = pain, aching, stiffness.

(compared with 1 m/s or better; HR, 95% CI for 0.8 m/s to <1 m/s: HR = 1.35, 0.88–2.07; 0.4 m/s to <0.8 m/s: HR = 2.60, 1.75–3.85; <0.4 m/s or unable: HR = 3.58, 2.33–5.48), there was no evidence of effect modification (interactions $P > 0.1$) between any of the six definitions of foot symptoms and walking speed. We also did not observe any effect modification between foot symptoms and the other variables considered (sex, race and ethnicity, obesity status, or diabetes status). In the fully adjusted model (model 3), no statistically significant associations were observed between mild or moderate foot symptom severity and mortality (HR = 0.86, 95% CI 0.68–1.10 and HR = 1.16, 95% CI 0.95–1.43, respectively). However, there was a significant association with severe foot symptom severity and mortality (HR = 1.48, 95% CI 1.15–1.90).

DISCUSSION

This community-based study found that the presence of at least moderate foot symptoms, independent of walking speed and many other potential confounding factors, was associated with a 30% to 48% increased hazard of reduced time to all-cause mortality in a large cohort of Black and White men and women. These effects were larger in those with greater symptom severity. Several definitions of foot symptoms were investigated, including any presence of pain, aching, or stiffness, foot symptom laterality, and measures of severity. A severity count across both feet was calculated because the presence of more severe symptoms or bilateral foot symptoms would likely have larger impacts on decreased walking speed compared with mild foot symptoms or unilateral foot symptoms. Although slower walking speeds continued to predict worse survival ($P < 0.0001$; data not shown), we did not find evidence that walking speed modified the effect of foot symptoms on mortality.

Previous studies have focused on hip or knee OA and pain in relation to mortality and may have been limited in the potential confounding factors considered in their analyses of OA to mortality (8–12). Our study is novel, in that to our knowledge, no previous studies have investigated the relationship of foot symptoms to mortality, and additionally, this study has adjusted for a wide variety of confounding factors, which may mediate effects on time to mortality. Even after adjustment for demographics including age, sex, race and ethnicity, and education, the effect of foot symptoms continued to predict mortality. Even further adjustment for other factors, including physical activity level, BMI, or comorbidities, as well as pain, aching, or stiffness at the knee or hip, did not explain away the association between foot symptoms on time to mortality in our study.

The effect of foot symptoms on mortality was independent of walking speed in this study. Master et al previously used data from JoCoOA with recorded 8-foot walk times (2.4 meters) to examine the association of walking speed with mortality risk over 9 years. They found a 23% higher hazard of mortality in those with symptomatic knee radiographic OA and proposed that walking difficulty may modify this relation (14). We did not find walking speed to be a moderator of the effect of foot symptoms on mortality, suggesting that poor mobility is not the explanation for the observed association.

Our results show a statistically significant and persistent relationship between foot symptoms and time to all-cause mortality; however, the underlying cause of this relationship remains unclear. Previous studies have shown that chronic pain, which is most often musculoskeletal in etiology (23), was significantly associated with mortality (24,25), but these studies have been inconsistent in analyzing for sociodemographic factors and differentiating arthritis from other causes of chronic pain. Torrance et al found that severe chronic pain was significantly associated with all-cause mortality (HR = 1.49, 99% CI 1.21–1.84) independent of sociodemographic factors including age, sex, marital status, education, and housing (25). However, almost half of the participants had chronic pain due to arthritis, and among those with arthritis, there was no significant association found between chronic pain and overall mortality. Self-reported chronic musculoskeletal pain in a large prospective population-based study of middle-aged women was associated with increased risk of mortality (HR = 2.1, 95% CI 1.1–4.2) (24). Notably, there was no difference in death from CVD or cancer between pain-free individuals compared with those with chronic pain. Torrance et al did not account for other confounding factors, including depression, comorbidity, lifestyle factors, social factors, and the duration of pain at baseline. Additionally, they stated that individuals with chronic pain have a less healthy lifestyle, are less physically active, smoke more, and belong to a lower social class, which they believe may be important confounders (24). Our study was able to include these potential confounders, and we did not observe that the effect of foot symptoms on mortality was altered by these factors. Additional studies in other populations should examine the presence of chronic foot pain and mortality to determine whether findings are consistent or differ. Moreover, the presence of anxiety and depression has been observed in around a third of patients with chronic foot and ankle diseases (26), and individuals with foot symptoms are more likely to report depressive symptoms (27). Although we accounted for depression as a confounder, future studies may examine this relationship more closely in relation to mortality.

It is likely that the presence of foot symptoms and/or OA may impair balance and muscle strength within the lower extremities, similar to knee OA. Previous research has investigated the presence of symptomatic radiographic hip and/or knee OA characterized by pain, aching, or stiffness on most days in relation to fall risk (28). This study found an increased incidence of falls with increasing number of knee and/or hip joints with symptomatic OA. Examination of foot OA symptoms within these analyses would be an important next step because falls are a leading cause of morbidity and mortality in older adults and are associated with foot pain (21).

Our study had several limitations. Regarding participant data, duration of foot symptoms prior to study entry was not assessed. In future studies, information about foot symptoms prior to study entry would expand the knowledge of the observed association between foot symptoms and mortality. Because of our cohort study design, we were unable to capture incident foot symptoms as they occur because the study examinations at which participants provided foot symptoms information occurred at ∼5-year cycles. Additionally, we did not include radiographic OA within our study because foot radiographs were not obtained for our cohort until 2013, significantly decreasing any possible follow-up time for assessing mortality. Additional investigations of the association of foot OA to mortality should include radiographs to compare symptomatic radiographic OA to asymptomatic radiographic OA. Another limitation may be our use of the 8-foot walk because in our study, participants were unable to decelerate past the 8-foot mark because of space limitations within the room. For future studies, we suggest using a larger space and having participants complete a short walk (eg, 8 feet) and a longer walk (eg, greater than 60 feet) to assess mobility and its relationship to foot symptoms and OA and mortality more accurately.

Finally, although we produced estimates of the direct effects of foot symptom severity on mortality and considered effect modifiers of this association, future studies could assess whether modifiable mediators (for example, weight, depression, physical function, sleep) of this association exist and estimate both direct and indirect effects of this association.

The major strength of our study was use of a large group of community-dwelling Black and White men and women with a long follow-up time and a wealth of well-characterized covariables, including walking speed. We were able to assess models controlling for several comorbid conditions as well as lifestyle factors and risk factors for OA in our analyses. Finally, we were able to use the data from multiple time points to analyze foot symptoms over time. Although foot symptoms were not collected prior to study entry, we did have a long-recorded period of foot symptoms, which is not seen in many previous studies.

In conclusion, foot symptoms, after adjustment for walking speed, sex, race and ethnicity, obesity status, or diabetes status, may signify a higher hazard of all-cause mortality in older adults. This is especially an issue for those with at least moderate to severe foot pain, aching, or stiffness. Although we continued to see the effects of foot symptoms on increased hazard of mortality with slower walking speed, we did not find any evidence of effect modification of this association by sex, race, obesity, diabetes, or walking speed. Health professionals may consider therapeutic management of foot pain, perhaps with a view towards chronic pain management in adults with foot symptoms to manage risks in these individuals. Furthermore, this study highlights the need for future investigations of modifiable factors that alleviate foot symptoms.

ACKNOWLEDGMENTS

The authors acknowledge the staff and participants in the JoCoOA, without whom this work would not be possible.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Harmon, Alvarez, Hannan, Callahan, Gates, Bowen, Menz, Nelson, Golightly.

Acquisition of data. Callahan, Nelson, Golightly.

Analysis and interpretation of data. Harmon, Alvarez, Hannan, Callahan, Gates, Bowen, Menz, Nelson, Golightly.

ADDITIONAL DISCLOSURES

Author Hannan received funding to her institution from Amgen, Inc. and serves on the Board of Directors for the Rheumatology Research Foundation; neither of these activities are related to the work reported in this project.

REFERENCES

1. Golightly YM, Hannan MT, Shi XA, et al. Association of foot symptoms with self-reported and performance-based measures of physical function: the Johnston County osteoarthritis project. Arthritis Care Res (Hoboken) 2011;63:65–9.
2. Gates LS, Arden NK, Hannan MT, et al. Prevalence of foot pain across an international consortium of population-based cohorts. Arthritis Care Res (Hoboken) 2019;71:661–70.
3. International FaAOC, Arnold JB, Bowen CJ, et al. International Foot and Ankle Osteoarthritis Consortium review and research agenda for diagnosis, epidemiology, burden, outcome assessment and treatment. Osteoarthritis and cartilage / OARS, Osteoarthritis Research Society 2022;30:945–55.
4. Benvenuti F, Ferrucci L, Guralnik JM, et al. Foot pain and disability in older persons: an epidemiologic survey. J Am Geriatr Soc 1995;43:479–84.
5. Keysor JJ, Dunn JE, Link CL, et al. Are foot disorders associated with functional limitation and disability among community-dwelling older adults? J Aging Health 2005;17:734–52.
6. Menz HB, Tiedemann A, Kwan MM, et al. Foot pain in community-dwelling older people: an evaluation of the Manchester Foot Pain and Disability Index. Rheumatology (Oxford) 2006;45:863–7.
7. Roddy E, Thomas MJ, Marshall M, et al. The population prevalence of symptomatic radiographic foot osteoarthritis in community-dwelling older adults: cross-sectional findings from the clinical assessment study of the foot. Ann Rheum Dis 2015;74:156–63.
8. Golightly YM, Gates LS. Foot osteoarthritis: addressing an overlooked global public health problem. Arthritis Care Research (Hoboken) 2021;73:767–9.
9. Cleveland RJ, Alvarez C, Schwartz TA, et al. The impact of painful knee osteoarthritis on mortality: a community-based cohort study with over 24 years of follow-up. Osteoarthritis and cartilage / OARS, Osteoarthritis Res Soc 2019;27:593–602.
10. Nüesch E, Dieppe P, Reichenbach S, et al. All cause and disease specific mortality in patients with knee or hip osteoarthritis: population based cohort study. BMJ 2011;342:d1165.

11. Leyland KM, Gates LS, Sanchez-Santos MT, et al. Knee osteoarthritis and time-to all-cause mortality in six community-based cohorts: an international meta-analysis of individual participant-level data. Aging Clin Experimental Res 2021;33:529–45.
12. Cleveland RJ, Alvarez C, Nelson AE, et al. Hip symptoms are associated with premature mortality: the Johnston County Osteoarthritis Project. Osteoarthritis and cartilage / OARS, Osteoarthritis Res Soc 2020;28:1330–40.
13. Margulis H. Gait and walking speed as a predictor of health Hebrew SeniorLife 2021 URL: https://www.hebrewseniorlife.org/blog/gait-and-walking-speed-predictor-health.
14. Master H, Neogi T, Callahan LF, et al. The association between walking speed from short- and standard-distance tests with the risk of all-cause mortality among adults with radiographic knee osteoarthritis: data from three large United States cohort studies. Osteoarthritis Cartilage / OARS, Osteoarthritis Res Soc 2020;28:1551–8.
15. Jordan JM, Helmick CG, Renner JB, et al. Prevalence of knee symptoms and radiographic and symptomatic knee osteoarthritis in African Americans and Caucasians: the Johnston County Osteoarthritis Project. J Rheumatol 2007;34:172–80.
16. Jette AM, Jette DU, Ng J, et al. Are performance-based measures sufficiently reliable for use in multicenter trials? Musculoskeletal Impairment (MSI) Study Group. J Gerontol A Biol Sci Med Sci 1999;54:M3–6.
17. Ostchega Y, Harris TB, Hirsch R, et al. Reliability and prevalence of physical performance examination assessing mobility and balance in older persons in the US: data from the Third National Health and Nutrition Examination Survey. J Am Geriatr Soc 2000;48:1136–41.
18. Middleton A, Fritz SL, Lusardi M. Walking speed: the functional vital sign. J Aging Phys Act 2015;23:314–22.
19. Flowers P, Nelson AE, Hannan MT, et al. Foot osteoarthritis frequency and associated factors in a community-based cross-sectional study of white and African American adults. Arthritis Care Res (Hoboken) 2021;73:1784–8.
20. Butterworth PA, Landorf KB, Smith SE, et al. The association between body mass index and musculoskeletal foot disorders: a systematic review. Obesity Rev 2012;13:630–42.
21. Menz HB, Auhl M, Spink MJ. Foot problems as a risk factor for falls in community-dwelling older people: a systematic review and meta-analysis. Maturitas 2018;118:7–14.
22. Rubin DB, Schenker N. Multiple imputation for interval estimation from simple random samples with ignorable nonresponse. J Am Stat Assoc 1986;81:366–74.
23. Ali A, Arif AW, Bhan C, et al. managing chronic pain in the elderly: an overview of the recent therapeutic advancements. Cureus 2018;10:e3293.
24. Nitter AK, Forseth KO. Mortality rate and causes of death in women with self-reported musculoskeletal pain: results from a 17-year follow-up study. Scand J Pain 2013;4:86–92.
25. Torrance N, Elliott AM, Lee AJ, et al. Severe chronic pain is associated with increased 10 year mortality. A cohort record linkage study. Eur J Pain 2010;14:380–6.
26. Nakagawa R, Yamaguchi S, Kimura S, et al. Association of anxiety and depression with pain and quality of life in patients with chronic foot and ankle diseases. Foot Ankle International. 2017;38:1192–8.
27. Awale A, Dufour AB, Katz P, et al. Link between foot pain severity and prevalence of depressive symptoms. Arthritis Care Res (Hoboken) 2016;68:871–6.
28. Dore AL, Golightly YM, Mercer VS, et al. Lower-extremity osteoarthritis and the risk of falls in a community-based longitudinal study of adults with and without osteoarthritis. Arthritis Care Res (Hoboken) 2015;67:633–9.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 77, No. 3, March 2025, pp 376–384
DOI 10.1002/acr.25436

Low Socioeconomic Status and Female Sex are Associated With Worse Functional Status in Axial Spondyloarthritis

Rachael Stovall,[1] Jing Li,[2] Jessica Fitzpatrick,[2] Eric Roberts,[2] Andriko Palmowski,[3] Christine Anastasiou,[2] Zara Izadi,[2] Janna Friedly,[1] Namrata Singh,[1] Lianne S. Gensler,[2] Gabriela Schmajuk,[4] and Jinoos Yazdany[2]

Objective. We determined whether socioeconomic status (SES) and sex are associated with functional status (FS) in axial spondyloarthritis (axSpA).

Methods. We conducted a cohort study of patients with axSpA in the Rheumatology Informatics System for Effectiveness registry. We performed cross-sectional and longitudinal analyses of FS through the Multidimensional Health Assessment Questionnaire (MDHAQ) using generalized estimating equation models. Area Deprivation Index (ADI) was used as an SES proxy. The cross-sectional analysis tested for a linear trend across ADI quintiles for MDHAQ. The longitudinal analysis' outcome was functional decline. We reported predictive margins and assessed for interaction with sex. In the longitudinal analysis, we reported odds of functional decline.

Results. In the cross-sectional analysis (N = 5,658), the mean ± SD age was 53.8 ± 15.2 years, 55.8% were female, and 71.4% were non-Hispanic White. The mean ± SD MDHAQ scores were 1.6 ± 2.0 in men versus 2.1 ± 2.2 in women. Predicted mean MDHAQ scores were 2.2 (95% confidence interval [CI] 1.8–2.7) for the lowest ADI quintile and 1.8 (95% CI 1.4–2.1) for the highest. Women had lower FSs compared to men across quintiles. In the longitudinal analysis (n = 2,341), the proportion with FS decline was 14.3% (95% CI 7.6–25.5%) for the lowest SES quintile compared to 9.6% (95% CI 5.2–17.1%) for the highest. Women had 1.7 (95% CI 1.3–2.2) times higher odds of functional decline compared to men. There was no interaction with sex.

Conclusion. In this large sample of patients with axSpA, those with lower SES had worse FS and functional decline. Women had worse FS than men, initially and over time.

INTRODUCTION

Axial spondyloarthritis (axSpA) is a potentially disabling condition that primarily affects the axial skeleton but can involve peripheral joints and have extramusculoskeletal manifestations, as well. Some studies suggest that lower socioeconomic status (SES) is associated with worse functional status (FS) in axSpA, but little is known on this topic, and studies have been small and from single centers.[1] Similar findings have also been observed in other rheumatic diseases, such as rheumatoid arthritis[2] and

Data collection was supported by the American College of Rheumatology's (ACR) Rheumatology Informatics System for Effectiveness (RISE) Registry. The views expressed represent those of the authors, not necessarily those of ACR.

Supported by the National Center for Advancing Translational Sciences, NIH, University of California San Francisco (UCSF) Clinical & Translational Science Institute (TL1 TR001871), PREMIER, a NIH/National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) P30 Center for the Advancement of Precision Medicine in Rheumatology at UCSF (grant P30-AR-070155), and NIH/NIAMS (grant K24-AR-074534). Dr Palmowski's work was supported by the Deutsche Autoimmun-Stiftung ("German Autoimmune Foundation"). Dr Anastasiou's work was supported by the Rheumatology Research Foundation Scientist Development Award. Dr Singh's work was supported by NIH/NIAMS (grant K23-AR-079588) and by the NIH/National Institute on Aging (grant R03-AG-082857).

[1]Rachael Stovall, MD, MAS, Janna Friedly, MD, MPH, Namrata Singh, MD, MSc: University of Washington, Seattle; [2]Jing Li, MPH, Jessica Fitzpatrick, PhD, MPH, Eric Roberts, PhD, MPH, Christine Anastasiou, MD, MAS, Zara Izadi, PharmD, PhD, Lianne S. Gensler, MD, Jinoos Yazdany, MD, MPH: University of California San Francisco; [3]Andriko Palmowski, MD: University of California San Francisco, Charité – Universitätsmedizin Berlin, Berlin, Germany, and University of Copenhagen, Copenhagen, Denmark; [4]Gabriela Schmajuk, MD, MS: University of California San Francisco, San Francisco VA Healthcare System, San Francisco, California, and University of California San Francisco Institute for Health Policy Research.

The ACR owns the data in the RISE registry, and UCSF, as a Data Analytic Center for the ACR, has access to the data for specific research projects, including this one, but is contractually obligated through data use agreements with practices to not share this data, even in a deidentified state. Researchers interested in performing additional analyses from this data are invited to submit proposals to the ACR RISE registry (https://rheumatology.org/request-rise-data).

Additional supplementary information cited in this article can be found online in the Supporting Information section (http://onlinelibrary.wiley.com/doi/10.1002/acr.25436).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25436.

Address correspondence via email to Rachael Stovall, MD, MAS, at stovall@uw.edu.

Submitted for publication March 12, 2024; accepted in revised form September 16, 2024.

SIGNIFICANCE & INNOVATIONS

- This research is one of the first to evaluate functional status (FS) in axial spondyloarthritis (axSpA) by socioeconomic status (SES) and sex in a large sample of US patients.
- We found that those with lower SES had worse FS and greater functional decline over time. Women also had worse FS initially and over time compared to men.
- Individuals with low SES and women should be prioritized for interventions to reduce functional decline and disability in axSpA. Future work should assess the mechanisms of SES and sex-based FS differences.

lupus.[3] Researchers have hypothesized that social determinants of health, such as access to health care and drug therapy, environmental exposures, and psychosocial stressors are contributing to these disparate outcomes.[4]

Although men and women appear to have an equal prevalence of axSpA, there are well-known sex differences in this condition, including higher disease activity and worse patient-reported outcomes among women.[5,6] Regarding FS, there are conflicting studies as to whether sex differences are present. Brazilian, British, and French studies found that women had worse FS compared to men via the Bath Ankylosing Spondylitis Functional Index (BASFI).[7–9] A Danish study found no difference in FS by sex in radiographic axSpA; however, with nonradiographic axSpA (nr-axSpA), there appeared to be worse FS among women compared to men.[10] A US study found no difference in FS by sex overall; however, when adjusting for radiographic axial damage, women had worse FS for their level of damage compared to men.[11] Meanwhile, studies from Iran, South America, and Turkey did not find any sex differences in FS among patients with axSpA.[1,12,13]

Determining whether people of low SES and women have worse FS and/or more functional decline in the United States is the first step before delving into reasons why this may be and interventions to address it. Furthermore, it is possible that SES and sex may jointly affect FS in axSpA. For example, women of lower SES tend to handle more household chores and childcare responsibilities, even when working full time[14–16]; female sex and low SES together could multiplicatively lower FS.

To address these concerns, we conducted a large national study to examine the relationships between both SES and sex with FS and decline in people with axSpA in the United States. We were interested in both SES and sex individually as risk factors, as well as the interaction between these factors. We therefore evaluated FS by SES and sex in axSpA using the Rheumatology Informatics System for Effectiveness (RISE) registry, a national electronic health record (EHR)–based registry of US individuals with rheumatic disease. We hypothesized that patients with low SES and women would have worse functional outcomes compared to those with high SES and men.

MATERIALS AND METHODS

Data source. We conducted a retrospective cohort study using data from the RISE registry from 2016 to 2022. RISE contains EHR data recorded during routine outpatient clinical care in participating rheumatology practices across the United States.[17] As of March 2022, RISE includes information from the EHRs of 238 US rheumatology practices and >1,000 clinicians, which represents about 30% of the US clinical rheumatology workforce. Since its inception in 2014, RISE has collected information on >2.5 million patients with more than 20 million encounters. RISE is designed to minimize the impact on practice workflow. No data entry into a separate database is required; instead, RISE collects data that are entered during routine care into the EHR, including structured data, such as diagnoses, medications, laboratory tests, and patient-reported outcomes, as well as clinical notes. RISE can connect to most certified EHR systems; NextGen, eClinicalWorks, AllScripts, and Practice Fusion are the most common systems.

Study sample. To identify individuals with prevalent axSpA, we used *International Classification of Diseases*, *Ninth Revision* (ICD-9) and *Tenth Revision* (ICD-10) codes for nr-axSpA and ankylosing spondylitis (AS) (720.0, M45.x, and M46.8) in the RISE registry from January 1st, 2016, to March[18–20] 31st, 2022 (Supplemental Table 1). In 2019, ICD-10 codes expanded to include nr-axSpA as its own separate ICD code, M45.A. Because we included these codes in our study, we considered this population of patients to have axSpA rather than only AS. Individuals were eligible for cohort inclusion if they were 18 years or older, had two or more axSpA ICD codes ≥30 days apart in RISE, and had one or more FS score (Multidimensional Health Assessment Questionnaire [MDHAQ]) documented during 2016 to 2022. We excluded patients who had a history of HIV or cancer (ICD-9, *Clinical Modification* [CM] codes 140.x-238.x; ICD-10-CM codes C00.x-D47.x) because these comorbidities can be associated with FS decline and limit therapeutic options for axSpA.

Validation of administrative definition in RISE data. In our previous work studying AS in RISE, we performed a detailed review of a sample of available RISE clinical notes for documentation of an AS diagnosis (gold standard) in those with two or more ICD-9 or ICD-10 codes (the definition used in the current study) and found a positive predictive value[21] of 88%. In this review, 3 of 50 (6%) had an alternative diagnosis in RISE, yielding a sensitivity of 88%.

Outcomes. We first performed a cross-sectional analysis in which the outcome was FS, as measured by the most recent

MDHAQ score (0–10 scale) during 2019 to 2022. The MDHAQ FS score is part of the Routine Assessment of Patient Index Data 3 (RAPID3), which has been validated in AS, and is a validated measure of FS in general.[22,23] Rheumatology practices using 1 to 30–point or 0 to 3–point scales for MDHAQ were all converted to 0 to 10–point scales, by dividing the scores by 10 or multiplying by 3.3333, respectively. We then performed a longitudinal analysis in which the outcome was functional decline (yes or no) defined as a >1.2 point difference in MDHAQ at two time points at least one year apart (most recent MDHAQ minus the next most recent MDHAQ ≥12 months before); this is the minimally clinically important difference (MCID) for the MDHAQ.[2,23]

Study design. In the cross-sectional analysis, variables were measured within 12 months of the latest FS measurement. In the longitudinal analysis, variables were measured 12 months before the initial FS measurement.

Covariates. Our primary covariates of interest were sex and Area Deprivation Index (ADI). Given that RISE includes multisite EHR data, we assessed sex as female or male only based on data availability, although we acknowledge that sex is not binary.[24] EHR data in the RISE registry did not allow for analysis of gender. We used ADI, a zip code–based measure for neighborhood poverty, as a proxy for SES, as used in other studies.[2,25] ADI was categorized into quintiles with quintile 1 representing the highest SES category and quintile 5 the lowest SES category.

Additional covariates included age, race, ethnicity, body mass index (BMI) (<18, 18–24.9, 25–29.9, and 30+), Charlson comorbidity index (≥2; yes or no),[26] smoking status (ever or never), administration of biologic disease modifying antirheumatic drugs (DMARDs) (tumor necrosis factor inhibitors, interleukin 17 inhibitors; yes or no), and administration of targeted synthetic DMARDs (JAK inhibitors; yes or no). The initial MDHAQ was used as an adjustment variable in the longitudinal analysis.

To further characterize our study population, we reported information on insurance type (private, Medicare, Medicaid, other, and unknown) and practice characteristics including practice type (single specialty group practice, solo practitioner, multispecialty group practice), practice location (according to the nine US geographic divisions), and EHR vendor (NextGen, eClinicalWorks, Amazing Charts, GE Centricity, or other). We calculated the median number of patient visits during the study period. Lastly, we described oral glucocorticoid and opioid administration.

Statistical analysis. First, we generated descriptive statistics for patient demographics, ADI, Charlson comorbidity index, smoking status, and biologic or targeted synthetic DMARD (b/tsDMARD) administration. We then constructed cross-sectional and longitudinal models using generalized estimating equations (GEEs), in which the outcomes were MDHAQ score and functional decline, respectively, adjusting for covariates and accounting for clustering of patients within rheumatology practices. In the cross-sectional analysis, we first assessed for associations between (1) SES (ADI quintile) and (2) sex with the most recent MDHAQ score during the study period. We specifically examined whether the impact of SES on FS varied by participant sex by testing for an interaction between the two variables.

In the longitudinal analysis, we evaluated for evidence of functional decline. Then we tested for a linear trend across ADI quintiles for decline in MDHAQ, with and without interaction for sex. This is because we hypothesized that women would have worse functional decline beyond what would be expected by ADI quintiles. In models in which the outcome was FS decline, we adjusted for BMI, Charlson comorbidity index, smoking, and initial MDHAQ scores. Predictive margins were generated in both analyses to facilitate model interpretation. We also calculated the odds of functional decline in the longitudinal study.

Of the variables included in the models, we addressed missing data for race and ethnicity, BMI, and smoking status. Participants were categorized as "unknown" if their data were missing for those variables; no observations were dropped from the models.

The characteristics of patients included in the cross-sectional analysis were compared to those excluded from the study (because of missingness of MDHAQ) using a logistic regression model with a binary indicator for outcome missingness. Little's test was used to assess whether the data were missing completely at random.

Exploratory analyses. We repeated our main analysis using initial MDHAQ scores represented as quartiles, treating them as a categorical covariate instead of a continuous variable. We evaluated for potential nonlinear relationships between variables, including age and initial MDHAQ and age and functional decline. The Western Institutional Review Board and University of California San Francisco Committee on Human Research approved this study. We adhered to the Strengthening the Reporting of Observational Studies in Epidemiology guidelines to improve the reporting of this observational study (Supplementary File).[27]

RESULTS

We identified 5,658 adults with axSpA for the cross-sectional analysis, and 2,357 of those were included in the longitudinal cohort. In the cross-sectional analysis, the mean ± SD age was 53.8 ± 15.2 years, 55.8% were female, and 71.4% were non-Hispanic White (Table 1). A majority of individuals (80.1%) were seen in single specialty group practices. Nearly half (46.3%) resided in the South Atlantic, followed by 17.3% in East South Central United States. Nextgen was the most common EHR used (81.7%), followed by eClinicalWorks (5.2%). The median number of visits per patient per study period (interquartile range [IQR])

Table 1. Baseline characteristics*

Characteristic	Total (N = 5,658)	Longitudinal (n = 2,357)[a]
Age, mean (SD), y	53.8 (15.2)	54.8 (15.0)
<25	148 (2.6)	40 (1.7)
25–50	2,185 (38.6)	875 (37.1)
51–75	2,886 (51.0)	1,249 (53.0)
>75	439 (7.8)	193 (8.2)
Sex, n (%)		
Female	3,155 (55.8)	1,329 (56.4)
Male	2,503 (44.2)	1,028 (43.6)
Race and ethnicity, n (%)		
White	4,041 (71.4)	1,775 (75.3)
African American	252 (4.5)	105 (4.5)
Hispanic	222 (3.9)	79 (3.4)
Asian	128 (2.3)	47 (2.0)
Other or mixed	34 (0.6)	16 (0.7)
Unknown	981 (17.3)	335 (14.2)
Insurance, n (%)		
Private	3,339 (59.0)	1,371 (58.2)
Medicare	1,476 (26.1)	685 (29.1)
Medicaid	241 (4.3)	100 (4.2)
Other	113 (2.0)	43 (1.8)
Unknown	489 (8.6)	158 (6.7)
National ADI, median (IQR)	41 (21–64)	41 (20–63)
Practice type, n (%)		
Single specialty group practice	4,201 (80.1)	1,917 (86.2)
Solo practitioner	523 (10.0)	121 (5.4)
Multispecialty group practice	519 (9.9)	185 (8.3)
Division, n (%)		
East North Central	386 (6.8)	162 (6.9)
East South Central	981 (17.3)	488 (20.7)
Mid-Atlantic	229 (4.0)	124 (5.3)
Mountain	99 (1.7)	22 (0.9)
Pacific	143 (2.5)	40 (1.7)
South Atlantic	2,621 (46.3)	971 (41.2)
West North Central	688 (12.2)	287 (12.2)
West South Central	511 (9.0)	263 (11.2)
Electronic health record, n (%)		
Nextgen	4,621 (81.7)	2,050 (87.0)
eClinicalWorks	293 (5.2)	61 (2.6)
GE Centricity	119 (2.1)	21 (0.9)
Multiple	595 (10.5)	211 (9.0)
Other	30 (0.5)	14 (0.5)
Visits per patient during the study period, median (IQR)	8 (5–14)	11 (8 –19)
Body mass index, n (%)		
Underweight (<18.5)	68 (1.2)	33 (1.4)
Normal weight (18.5–24.9)	999 (17.7)	447 (19.0)
Overweight (25–29.9)	1,876 (33.1)	922 (39.1)
Obesity (30+)	2,035 (36.0)	916 (38.9)
Unknown	680 (12.0)	39 (1.7)
Charlson Comorbidity Index ≥2, n (%)	333 (5.9)	165 (7.0)
Smoking status, n (%)		
Ever smoker	1,747 (30.9)	747 (31.7)
Nonsmoker	3,709 (65.6)	1,528 (64.8)
Unknown	202 (3.6)	82 (3.5)
Medication administration, n (%)		
bDMARDs (TNFi, IL-17i)	4,064 (71.8)	1,779 (75.5)
tsDMARDs (JAKi)	75 (1.3)	30 (1.3)
Glucocorticoids	3,143 (55.5)	1,446 (61.3)
Opioids	1,376 (24.3)	666 (28.3)

* ADI, Area Deprivation Index; bDMARD, biologic disease modifying antirheumatic drug; IL-17i, interleukin-17 inhibitor; IQR, interquartile range; JAKi, JAK inhibitor; TNFi, tumor necrosis factor inhibitor; tsDMARD, targeted synthetic disease modifying antirheumatic drug.

[a] All participants in the longitudinal analysis were included in the cross-sectional analysis.

Table 2. Predictive margins on ADI quintiles by most recent MDHAQ (cross-sectional) and change in MDHAQ (longitudinal) during the study period*

	Cross-sectional analysis (N = 5,658)		Longitudinal analysis (n = 2,341)[a]	
ADI quintile	Unadjusted, mean MDHAQ score (95% CI)	Adjusted GEE model,[b] predicted mean (95% CI)	Unadjusted, proportion of patients with declines in MDHAQ score (95% CI)	Adjusted GEE model,[c] predicted proportion (95% CI)
First ADI quintile (highest SES)	1.6 (1.3–2.0)	1.8 (1.4–2.1)	9.2 (5.7–14.6)	9.6 (5.2–17.1)
Second ADI quintile	1.9 (1.6–2.2)	2.0 (1.6–2.3)	10.4 (6.6–15.9)	10.7 (5.3–20.5)
Third ADI quintile	1.9 (1.6–2.2)	2.0 (1.6–2.3)	12.8 (8.8–18.2)	12.8 (7.5–21.0)
Fourth ADI quintile	2.1 (1.7–2.5)	2.1 (1.7–2.5)	12.8 (9.0–17.9)	12.8 (7.1–22.0)
Fifth ADI quintile (lowest SES)	2.3 (1.9–2.6)	2.2 (1.8–2.7)	14.2 (9.0–21.8)	14.3 (7.6–25.5)

* ADI, Area Deprivation Index; CI, confidence interval; GEE, generalized estimating equation; MDHAQ, Multidimensional Health Assessment Questionnaire; SES, socioeconomic status.
[a] Longitudinal model additionally adjusted for baseline MDHAQ score.
[b] Adjusted for age, sex, race, ethnicity, smoking status, Charlson Comorbidity Index ≥2, and biologic or targeted synthetic disease modifying antirheumatic drug administration.
[c] Patients of other or mixed race were excluded because zero had a decline (n = 16).

was 8 (5–14). The mean ± SD MDHAQ scores for women and men were 2.1 ± 2.2 and 1.6 ± 2.0, respectively. Glucocorticoids were administered in 55.5% of patients, and opioids were administered in 24.3% of patients. There were missing data for race and ethnicity (17.3%), BMI (12.0%), and smoking status (3.6%). Characteristics were similar in the longitudinal cohort, although there was a higher median number of patient visits (11 [IQR 8–19]), slightly higher administration of glucocorticoids (61.3%) and opioids (28.3%), and less missingness for race and ethnicity (14.2%) and BMI (1.7%).

The data for MDHAQ in this study were not missing completely at random, and patient level characteristics of those included in the analysis were somewhat different compared to those excluded from the study: included patients were more likely to be female, nonsmokers, and taking b/tsDMARDs. Additionally, excluded patients were more likely to receive care from a solo practitioner, and there was variability in outcome completeness based on practice geographic location.

Cross-sectional study. In the cross-sectional analysis, the unadjusted predicted mean MDHAQ score was 2.3 (95% confidence interval [CI] 1.9–2.6) for the fifth ADI quintile (lowest SES) compared to 1.6 (95% CI 1.3–2.0) for the first ADI quintile (highest SES; Table 2). The adjusted predicted mean MDHAQ scores were 2.2 (95% CI 1.8–2.7) for the fifth ADI quintile (lowest SES) compared to 1.8 (95% CI 1.4–2.1) for the first ADI quintile. In a GEE model, there was a statistically significant association between MDHAQ and female sex (adjusted β 0.4, 95% CI 0.2–0.5), Hispanic ethnicity (adjusted β 0.4, 95% CI 0.1–0.6), and BMI ≥30 (adjusted β 0.4, 95% CI 0.2–0.6) (Table 3). We did not see evidence of an interaction in the relationship between functional status and SES by sex, although women had consistently lower FSs compared to men across all ADI quintiles in the cross-sectional analysis (Supplemental Table 2).

Longitudinal study. In the longitudinal analysis, the unadjusted predicted proportion of patients with FS decline was 14.2% (95% CI 9.0–21.8%) in the fifth ADI quintile compared to

Table 3. Cross-sectional study: assessing adjusted associations with MDHAQ score during the study period (N = 5,658)*

	Unadjusted β (95% CI)	Adjusted β (95% CI)
Age	0.0 (0.0 to 0.0)	0.0 (0.0 to 0.0)
Female sex	**0.4 (0.2 to 0.5)**[a]	**0.4 (0.2 to 0.5)**[a]
Race and ethnicity		
White	Ref	Ref
African American	0.2 (0.0 to 0.5)	0.1 (−0.2 to 0.4)
Hispanic	**0.3 (0.1 to 0.6)**[a]	**0.4 (0.1 to 0.6)**[a]
Asian	**−0.3 (−0.6 to −0.1)**[a]	−0.2 (−0.4 to 0.0)
Other or mixed	−0.2 (−0.9 to 0.5)	−0.2 (−0.9 to 0.6)
Unknown	0.0 (−0.1 to 0.2)	0.0 (−0.1 to 0.2)
ADI quintile		
1 (highest SES)	Ref	Ref
2	**0.3 (0.0 to 0.5)**[a]	0.2 (0.0 to 0.4)
3	**0.3 (0.1 to 0.5)**[a]	0.2 (0.0 to 0.4)
4	**0.5 (0.2 to 0.8)**[a]	**0.4 (0.1 to 0.6)**[a]
5 (lowest SES)	**0.7 (0.4 to 0.9)**[a]	**0.5 (0.2 to 0.7)**[a]
BMI		
Underweight (<18.5)	0.2 (−0.2 to 0.6)	0.2 (−0.2 to 0.6)
Normal weight (18.5–24.9)	Ref	Ref
Overweight (25–29.9)	0.1 (0.0 to 0.3)	0.1 (0.0 to 0.3)
Obesity (30+)	**0.5 (0.3 to 0.7)**[a]	**0.4 (0.2 to 0.6)**[a]
Unknown	0.4 (0.0 to 0.5)	0.2 (−0.1 to 0.5)
Charlson comorbidity index ≥2	**0.2 (0.1 to 0.4)**[a]	0.1 (−0.1 to 0.3)
Smoking status		
Ever smoke	Ref	Ref
Nonsmoker	**−0.2 (−0.3 to −0.1)**[a]	**−0.2 (−0.3 to 0.0)**[a]
Unknown	0.0 (−0.3 to 0.3)	0.00 (−0.3 to 0.3)
Taking b/tsDMARDs	0.1 (0.0 to 0.2)	0.1 (0.0 to 0.2)

* Bold values indicate statistical significance. *P* trend for ADI quintile <0.001. Adjusted for age, sex, race, ethnicity, smoking status, and b/tsDMARD administration. ADI, Area Deprivation Index; b/tsDMARD, biologic or targeted synthetic disease modifying antirheumatic drug; BMI, body mass index; CI, confidence interval; MDHAQ, Multidimensional Health Assessment Questionnaire; Ref, reference; SES, socioeconomic status.
[a] Indicates $P < 0.05$.

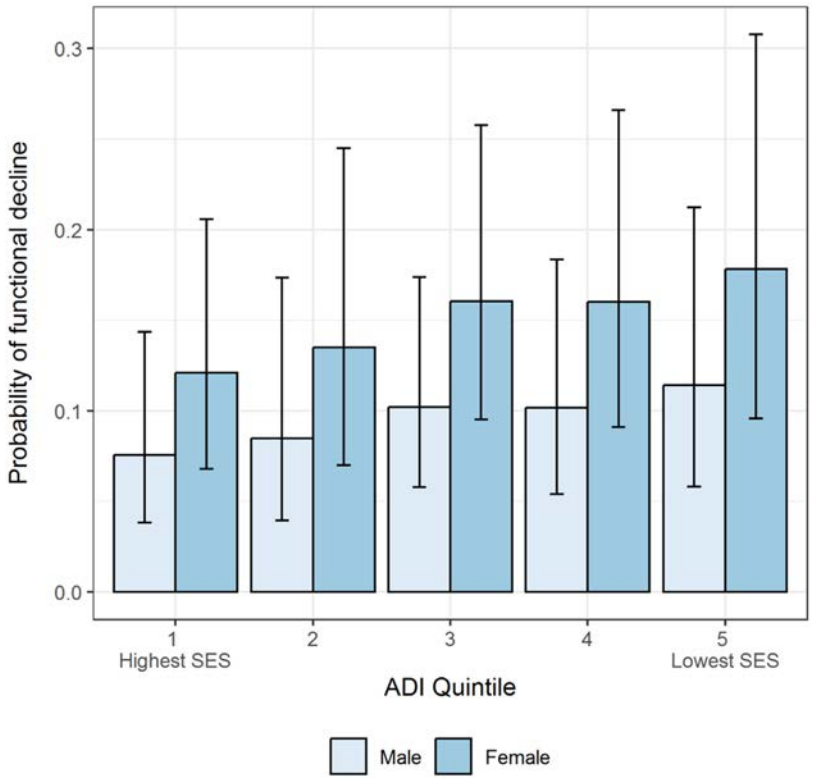


Figure 1. Primary outcome for the longitudinal analysis: predicted proportion of a decline in Multidimensional Health Assessment Questionnaire score of >1.2 by sex and ADI quintiles. ADI, Area Deprivation Index; SES, socioeconomic status.

9.2% (95% CI 5.7–14.6%) in the first ADI quintile (Table 2). The adjusted predicted proportion of patients with FS decline was 14.3% (95% CI 7.6–25.5%) in the fifth ADI quintile compared to 9.6% (95% CI 5.2–17.1%) in the first ADI quintile. Women had more functional decline compared to men across all ADI quintiles (Figure 1). We did not see evidence of an interaction the relationship between FS and SES by sex, although women had 1.7 (95% CI 1.3–2.2) times higher odds of functional decline compared to men (Table 4). As expected, a Charlson comorbidity index ≥2 was associated with 2.0 (95% CI 1.5–2.7) times the odds of functional decline. Age was not associated with functional decline (odds ratio [OR] 1.0 [95% CI 1.0–1.0]), and neither was b/tsDMARD treatment over administration (OR 0.9 [95% CI 0.7–1.1]).

Exploratory analyses. When modeling initial MDHAQ quartiles, instead of MDHAQ as a continuous variable, we found the results were similar to our primary analysis (data not shown). Results were also similar when examining age and initial MDHAQ deciles. When we stratified by age deciles, the proportion of patients experiencing functional decline remained consistent across the different age groups (Supplemental Figure).

DISCUSSION

In this national study, we examined the relationships between SES and sex with FS in people with axSpA. We found that both SES and sex were associated with worse FS and greater functional decline over time. We also found that sex does not interact with SES on functional decline. These findings are important insofar as they highlight the need to address both the socioeconomic and sex-related factors that contribute to poorer FS among individuals with axSpA in the United States.

The relationship between SES and physical function in axSpA is consistent with the growing literature suggesting that social determinants of health influence outcomes across a wide variety of diseases. We suspect that SES in this study could impact FS through a variety of mediators, including lack of access to medical care, higher chronic disease burden, or environmental or occupational conditions that increase the risk of injuries, respiratory problems, or psychosocial stress. Although imperfect, as a proxy for health care access we measured b/tsDMARD administration, and we did not find an association between b/tsDMARD administration and FS. This reinforces the point that there are many factors associated with worse FS outside of medical therapy.

Our findings are consistent with an international, cross-sectional study of patients with spondyloarthritis (including axSpA, psoriatic arthritis, and peripheral spondyloarthritis) that found worse functional outcomes (using the BASFI) among patients with lower education (a proxy for SES).[28] In this study, female sex was also associated with lower FS. A different

Table 4. Longitudinal analysis: assessing odds of functional decline in MDHAQ (increase in MDHAQ score from baseline >1.2) during the study period (n = 2,341)*

	Unadjusted odds ratio (95% CI)	Adjusted odds ratio (95% CI)
Age	1.0 (1.0–1.0)	1.0 (1.0–1.0)
Female sex	**1.5 (1.1–1.9)**[a]	**1.7 (1.3–2.2)**[a]
Race and ethnicity		
White	Ref	Ref
African American	1.5 (0.9–2.8)	1.6 (0.9–2.7)
Asian	1.7 (0.9–3.4)	1.6 (0.8–3.1)
Hispanic	0.9 (0.4–2.3)	1.0 (0.4–2.5)
Unknown	1.0 (0.7–1.4)	0.9 (0.6–1.3)
ADI quintile		
1 (highest SES)	Ref	Ref
2	1.1 (0.6–2.3)	1.1 (0.6–2.2)
3	1.4 (0.8–2.3)	1.4 (0.8–2.5)
4	1.4 (0.8–2.7)	1.4 (0.8–2.6)
5	1.6 (0.8–3.3)	1.6 (0.8–3.1)
BMI		
Underweight (<18.5)	0.3 (0.0–1.9)	0.3 (0.1–1.4)
Normal weight (18.5–24.9)	Ref	Ref
Overweight (25–29.9)	1.0 (0.8–1.2)	1.0 (0.8–1.3)
Obesity: 30+	1.1 (0.8–1.6)	1.2 (0.8–1.6)
Unknown	**2.2 (0.8–1.6)**[a]	**3.0 (1.3–7.2)**[a]
Charlson comorbidity index ≥2	**1.7 (1.2–2.2)**[a]	**2.0 (1.5–2.7)**[a]
Smoking status		
Ever smoker	Ref	Ref
Nonsmoker	1.1 (0.8–1.6)	1.1 (0.8–1.5)
Unknown	**0.5 (0.3–0.9)**[a]	**0.5 (0.3–0.8)**[a]
On b/ts DMARD	0.9 (0.7–1.2)	0.9 (0.7–1.1)
Baseline MDHAQ	**0.8 (0.6–0.9)**[a]	**0.7 (0.6–0.9)**[a]

* Bold values indicate statistical significance. Patients of other or mixed race were excluded because zero had a decline (n = 16). *P* trend for ADI quintile = 0.13. ADI, Area Deprivation Index; BMI, body mass index; b/tsDMARD, biologic or targeted synthetic disease modifying antirheumatic drug; CI, confidence interval; MDHAQ, Multidimensional Health Assessment Questionnaire; Ref, reference; SES, socioeconomic status.
[a] Indicates $P < 0.05$.

international cross-sectional study among patients with spondyloarthritis (including axial and peripheral spondyloarthritis) found similar results.[29]

Several explanations have been proposed for disparate FS in women compared to men. First, women with axSpA tend to have more peripheral arthritis and enthesitis, whereas men tend to have more axial symptoms.[7,9,12,13,30] This could hinder women from performing activities of daily living, thus resulting in worse FS. Despite women making up half of the workforce in the United States, there is still concern that women take up more household chores and childcare responsibilities compared to men.[14,15] This is especially true in lower SES in households[15] and could potentially impact women's reporting of their FS.[16]

Second, patients with axSpA can report symptoms of fatigue, pain, and stiffness that may be related to their disease or another medical problem. Patient-reported outcomes and FS measures must be interpreted with this in mind. Outside of axSpA disease activity, patients can report pain due to central sensitization, an alteration of the ways in which the central nervous system processes pain and other stimuli.[31] Some literature suggests that fibromyalgia, which involves the pathogenic mechanism of central sensitization,[32] is more common in women than men with axSpA.[33] A remote study from the Prospective Study of Outcomes in AS cohort found that women and men with axSpA had similar BASFI scores.[11] However, women reported worse FS for the level of radiographic spinal damage observed compared to men. Women may have more burden of disease outside the axial skeleton and/or may be experiencing more pain sensitization.

We were interested in exploring whether SES and sex interacted to jointly impact physical functioning in people with axSpA. Interestingly, we found no evidence of an interaction. This implies that both SES and sex are independently associated with low FS, but their combined effect is not significantly different than the sum of their individual effects.

Our study found that Hispanic ethnicity was associated with worse FS. In rheumatoid arthritis, Hispanic ethnicity has been associated with worse FS.[34] In a US-based spondyloarthritis study, African American patients had the highest BASFI scores (lowest FS), followed by Hispanic patients and then White patients.[35] Higher BMI was also associated with worse FS in our study, and this has been observed in other studies, as well.[29] This further supports recommendations for patients to engage in regular physical activity and for patients of elevated BMI to engage in weight loss.

Age was not associated with FS or functional decline. This may be due to the short time period over which we measured functional decline (at least one year), and perhaps if we studied patients over a longer period for evidence of functional decline, we would have seen a decline with age.

Although the patient characteristics in the cross-sectional and longitudinal studies were mostly similar, the longitudinal study group had a higher administration of glucocorticoids and opioids, medications that are more commonly administered in patients experiencing functional limitations. This implies that this group could have worse FS than those in the cross-sectional group, although the difference in administration of these medications was small.

This study has several strengths. First, using multisite EHR data from the RISE registry with FS patient-reported outcomes allowed us to study a large number of patients with axSpA from diverse geographic locations across the United States. Second, because all patients were seen in a rheumatology clinic, we were able to obtain longitudinal data on FS scores (MDHAQ) to evaluate functional decline over time. Third, our cohort is enriched with women, who made up more than half of our study sample, making RISE an important data source to study sex differences in axSpA.

Our study also has some limitations. Not all patients had an FS score available; this study only included 20.4% of patients with axSpA in RISE (Supplemental Table 1) and therefore may not be generalizable to all patients with axSpA in the United States.

However, this study was still able to capture a substantial number of patients with axSpA and is significantly larger than other studies measuring FS in axSpA. There could be concern that patients with poor FSs are more likely to fill out an MDHAQ at their rheumatology visit, biasing our results. However, previous analyses from RISE suggest that questionnaires are typically distributed to all patients in a clinic across rheumatic conditions and despite concern for assumed FS[36,37]; in other words, we find that missingness is not completely at random, but major differences are related to practices and not individuals. We cannot rule out the possibility, however, that some patients may not complete questionnaires they are handed in the clinic and that differential missingness could bias the results. RISE contains many community practices and few academic centers. If patients seeking care in academic centers have worse FS, studying patients in RISE could bias our results toward the null. This is especially true because RISE also includes more privately insured patients who are of higher SES, which is supported with the ADI distribution in Table 1. We chose ADI because of its emphasis on socioeconomic deprivation, although we recognize that it is an imperfect measure because it does not take into account all aspects of SES, including median housing value. Although BASFI is the axSpA-specific FS measure used in axSpA registries, we chose to use the MDHAQ FS measure, which is part of the RAPID3, another validated measure for FS in axSpA (there is a high correlation[22] of BASFI and RAPID3; $\rho = 0.73$, $P < 0.0001$). In doing so, we were able to capture a larger number of patients to study, but this limits the generalizability and ability to make direct comparisons to studies using other functional score measures.

In our exploratory analyses, we did not see any differences in our results when using age and initial MDHAQ deciles. Results were also unchanged when modeling initial MDHAQ quartiles instead of MDHAQ as a continuous variable. We did not find evidence of nonlinear relationships. We are aware of the limitations of dichotomizing the outcome variable (functional decline) in our analyses. This could lead to lost statistical power to detect a relationship between the variables and outcomes or conceal nonlinear relationships between covariates and outcomes. However, we did not find any evidence of nonlinear relationships. We chose our dichotomous modeling approach for the outcome because we were specifically interested in examining functional declines that were clinically significant, something that the use of a continuous variable would not allow. This led to the clinical decision to define functional decline by its MCID. Lastly, we cannot rule out unmeasured confounders, such as depression,[38] which could have driven some of the relationships we are seeing between FS and sex and SES.

In this large US sample of adults with axSpA, we found those with lower SES had worse FS and worse functional decline and there was no interaction with sex. Women had worse FS than men, initially and over time. Individuals with low SES and women should be prioritized for interventions to reduce functional decline and disability in axSpA. Future work should assess the mechanisms of SES and sex-based FS differences.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Stovall, Li, Roberts, Palmowski, Anastasiou, Izadi, Friedly, Singh, Gensler, Schmajuk, Yazdany.

Acquisition of data. Li, Fitzpatrick.

Analysis and interpretation of data. Stovall, Li, Fitzpatrick, Roberts, Palmowski, Anastasiou, Izadi, Friedly, Singh, Gensler, Schmajuk, Yazdany.

REFERENCES

1. Cansu DU, Calışır C, Savaş Yavaş U, et al. Predictors of radiographic severity and functional disability in Turkish patients with ankylosing spondylitis. Clin Rheumatol 2011;30(4):557–562.
2. Izadi Z, Li J, Evans M, et al. Socioeconomic disparities in functional status in a national sample of patients with rheumatoid arthritis. JAMA Netw Open 2021;4(8):e2119400.
3. Lotstein DS, Ward MM, Bush TM, et al. Socioeconomic status and health in women with systemic lupus erythematosus. J Rheumatol 1998;25(9):1720–1729.
4. Bonnell LN, Crocker AM, Kemp K, et al. The relationship between social determinants of health and functional capacity in adult primary care patients with multiple chronic conditions. J Am Board Fam Med 2021;34(4):688–697.
5. Stovall R, van der Horst-Bruinsma IE, Liu SH, et al. Sexual dimorphism in the prevalence, manifestation and outcomes of axial spondyloarthritis. Nat Rev Rheumatol 2022;18(11):657–669.
6. Rusman T, van Bentum RE, van der Horst-Bruinsma IE. Sex and gender differences in axial spondyloarthritis: myths and truths. Rheumatology (Oxford) 2020;59(suppl 4):iv38–iv46.
7. de Carvalho HM, Bortoluzzo AB, Gonçalves CR, et al; Brazilian Registry on Spondyloarthritis. Gender characterization in a large series of Brazilian patients with spondyloarthritis. Clin Rheumatol 2012;31(4):687–695.
8. Roussou E, Sultana S. Spondyloarthritis in women: differences in disease onset, clinical presentation, and Bath Ankylosing Spondylitis Disease Activity and Functional indices (BASDAI and BASFI) between men and women with spondyloarthritides. Clin Rheumatol 2011;30(1):121–127.
9. Tournadre A, Pereira B, Lhoste A, et al. Differences between women and men with recent-onset axial spondyloarthritis: results from a prospective multicenter French cohort. Arthritis Care Res (Hoboken) 2013;65(9):1482–1489.
10. Andreasen RA, Kristensen LE, Egstrup K, et al. The impact of sex and disease classification on patient-reported outcome measures in axial spondyloarthritis: a descriptive prospective cross-sectional study. Arthritis Res Ther 2019;21(1):221.
11. Lee W, Reveille JD, Davis JC Jr, et al. Are there gender differences in severity of ankylosing spondylitis? Results from the PSOAS cohort. Ann Rheum Dis 2007;66(5):633–638.
12. Shahlaee A, Mahmoudi M, Nicknam MH, et al. Gender differences in Iranian patients with ankylosing spondylitis. Clin Rheumatol 2015;34(2):285–293.

13. Landi M, Maldonado-Ficco H, Perez-Alamino R, et al; RESPONDIA Group. Fundación Reumatológica Argentina "Dr. Osvaldo García Morteo. Gender differences among patients with primary ankylosing spondylitis and spondylitis associated with psoriasis and inflammatory bowel disease in an iberoamerican spondyloarthritis cohort. Medicine (Baltimore) 2016;95(51):e5652.
14. Parker K, Wang W. Modern Parenthood: Roles of Moms and Dads Converge as They Balance Work and Family. Pew Research Center; 2013. https://www.pewresearch.org/social-trends/wp-content/uploads/sites/3/2013/03/FINAL_modern_parenthood_03-2013.pdf.
15. Heisig JP. Who does more housework: rich or poor?: a comparison of 33 countries. Am Sociol Rev 2011;76(1):74–99.
16. Theis KA, Brady TJ, Helmick CG, et al. Associations of arthritis-attributable interference with routine life activities: a modifiable source of compromised quality-of-life. ACR Open Rheumatol 2019;1(7):412–423.
17. Yazdany J, Bansback N, Clowse M, et al. Rheumatology Informatics System for Effectiveness: a national informatics-enabled registry for quality improvement. Arthritis Care Res (Hoboken) 2016;68(12):1866–1873.
18. Lindström U, Exarchou S, Sigurdardottir V, et al. Validity of ankylosing spondylitis and undifferentiated spondyloarthritis diagnoses in the Swedish National Patient Register. Scand J Rheumatol 2015;44(5):369–376.
19. Singh JA, Holmgren AR, Krug H, et al. Accuracy of the diagnoses of spondylarthritides in veterans affairs medical center databases. Arthritis Rheum 2007;57(4):648–655.
20. Dubreuil M, Peloquin C, Zhang Y, et al. Validity of ankylosing spondylitis diagnoses in the Health Improvement Network. Pharmacoepidemiol Drug Saf 2016;25(4):399–404.
21. Stovall R, Kersey E, Li J, et al. Incidence rate and factors associated with fractures among Medicare beneficiaries with ankylosing spondylitis in the United States. Arthritis Care Res (Hoboken) 2024;76(2):265–273.
22. García-Valle A, Andrés-de Llano JM, Fariña-González AJ, et al. Construct validity of the Routine Assessment of Patient Index Data 3 (RAPID3) in the evaluation of axial spondyloarthritis. J Rheumatol 2022;49(1):36–43.
23. Maska L, Anderson J, Michaud K. Measures of functional status and quality of life in rheumatoid arthritis: Health Assessment Questionnaire Disability Index (HAQ), Modified Health Assessment Questionnaire (MHAQ), Multidimensional Health Assessment Questionnaire (MDHAQ), Health Assessment Questionnaire II (HAQ-II), Improved Health Assessment Questionnaire (Improved HAQ), and Rheumatoid Arthritis Quality of Life (RAQoL). Arthritis Care Res (Hoboken) 2011;63(suppl 11):S4–S13.
24. Mauvais-Jarvis F, Bairey Merz N, Barnes PJ, et al. Sex and gender: modifiers of health, disease, and medicine. Lancet 2020;396(10250):565–582.
25. Kind AJ, Jencks S, Brock J, et al. Neighborhood socioeconomic disadvantage and 30-day rehospitalization: a retrospective cohort study. Ann Intern Med 2014;161(11):765–774.
26. Brusselaers N, Lagergren J. The Charlson Comorbidity Index in registry-based research. Methods Inf Med 2017;56(5):401–406.
27. Vandenbroucke JP, von Elm E, Altman DG, et al; STROBE Initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. Int J Surg 2014;12(12):1500–1524.
28. Capelusnik D, Zhao SS, Boonen A, et al. Individual-level and country-level socio-economic factors and health outcomes in spondyloarthritis: analysis of the ASAS-perSpA study. Rheumatology (Oxford) 2022;61(5):2043–2053.
29. Putrik P, Ramiro S, Moltó A, et al. Individual-level and country-level socioeconomic determinants of disease outcomes in SpA: multinational, cross-sectional study (ASAS-COMOSPA). Ann Rheum Dis 2019;78(4):486–493.
30. Lubrano E, Perrotta FM, Manara M, et al. The sex influence on response to tumor necrosis factor-α inhibitors and remission in axial spondyloarthritis. J Rheumatol 2017:jrheum.17666.
31. Volcheck MM, Graham SM, Fleming KC, et al. Central sensitization, chronic pain, and other symptoms: better understanding, better management. Cleve Clin J Med 2023;90(4):245–254.
32. Mezhov V, Guymer E, Littlejohn G. Central sensitivity and fibromyalgia. Intern Med J 2021;51(12):1990–1998.
33. Salaffi F, De Angelis R, Carotti M, et al. Fibromyalgia in patients with axial spondyloarthritis: epidemiological profile and effect on measures of disease activity. Rheumatol Int 2014;34(8):1103–1110.
34. Katz PP, Barton J, Trupin L, et al. Poverty, depression, or lost in translation? Ethnic and language variation in patient-reported outcomes in rheumatoid arthritis. Arthritis Care Res (Hoboken) 2016;68(5):621–628.
35. Jamalyaria F, Ward MM, Assassi S, et al. Ethnicity and disease severity in ankylosing spondylitis a cross-sectional analysis of three ethnic groups. Clin Rheumatol 2017;36(10):2359–2364.
36. Hammam N, Izadi Z, Li J, et al. The relationship between electronic health record system and performance on quality measures in the American College of Rheumatology's Rheumatology Informatics System for Effectiveness (RISE) registry: observational study. JMIR Med Inform 2021;9(11):e31186.
37. Nasrallah CJL, Schmajuk G, Kersey E, et al. Development of the American College of Rheumatology toolkit for implementation of rheumatoid arthritis outcome measures in clinical practice [abstract]. Arthritis Rheumatol 2023;75(suppl 9). https://acrabstracts.org/abstract/development-of-the-american-college-of-rheumatology-toolkit-for-implementation-of-rheumatoid-arthritis-outcome-measures-in-clinical-practice/.
38. Parkinson JT, Foley ÉM, Jadon DR, et al. Depression in patients with spondyloarthritis: prevalence, incidence, risk factors, mechanisms and management. Ther Adv Musculoskelet Dis 2020;12:1759720x20970028.

Arthritis Care & Research
Vol. 77, No. 3, March 2025, pp 385–392
DOI 10.1002/acr.25204

American College of Rheumatology
Empowering Rheumatology Professionals

Global Functioning in Axial Spondyloarthritis is Stronger Associated With Disease Activity and Function Than With Mobility and Radiographic Damage

David Kiefer,[1] Jürgen Braun,[1] Uta Kiltz,[1] Varvara Chatzistefanidi,[1] Daniela Adolf,[2] Ilka Schwarze,[3] Maria Kabelitz,[2] Uwe Lange,[4] Jan Brandt-Jürgens,[5] Edgar Stemmler,[6] Sabine Sartingen,[6] and Xenofon Baraliakos[1]

Objective. The Assessment of Spondyloarthritis International Society Health Index (ASAS HI) is a validated patient-reported outcome (PRO) for global functioning of patients with axial spondyloarthritis (axSpA). The Epionics SPINE (ES) is an electronic device for assessment of axial mobility that provides an objective measure of spinal mobility by assessing range of motion (RoM) and range of kinematics (RoK). The aim of this study is to investigate the relationship between global functioning and clinical measures of disease activity, physical function, spinal mobility, and radiographic damage.

Methods. In a cross-sectional study design, consecutive patients with radiographic and nonradiographic axSpA were included, and the following established tools were assessed: Bath ankylosing spondylitis (AS) disease activity index (BASDAI), Bath AS functional index (BASFI), Bath AS metrology index (BASMI), ASAS HI, and RoM and RoK using ES. Structural damage of spine and sacroiliac joints (SIJ) were assessed by counting the number of syndesmophytes and by New York grading of sacroiliitis. Kendall's tau correlation coefficients were calculated.

Results. In 103 patients with axSpA, ASAS HI scores correlated significantly with PRO scores (BASDAI, $r = 0.36$; BASFI, $r = 0.48$; and back pain, $r = 0.41$; all $P < 0.001$). In contrast, no significant correlation between ASAS HI and RoM and RoK (r between –0.08 and 0.09) and radiographic damage in SIJ and spine (all r between 0.03 and 0.004) were seen, respectively. BASMI scores correlated weakly ($r = 0.14$; $P = 0.05$).

Conclusion. This study shows that axSpA disease-specific PROs have an impact on global functioning, whereas spinal mobility scores, even if objectively assessed by the ES, have limited impact on patient reported–global functioning. The results also suggest that global functioning is, in this cohort, not much dependent on the degree of structural damage in the axial skeleton.

INTRODUCTION

Axial spondyloarthritis (axSpA) is a chronic inflammatory disease affecting the axial skeleton that usually begins in early adulthood. As a result, various health problems occur in these patients in the long course of the disease (1–4). The most characteristic and leading symptom is inflammatory back pain, but patients with axSpA may also present with many different axial and peripheral symptoms, as well as extramuskuloskeletal disease manifestations and comorbidities, over time. The classification of axSpA is based on the 2009 Assessment of Spondyloarthritis International Society (ASAS) criteria (5), which are used to differentiate between

[1]David Kiefer, MD, Jürgen Braun, MD, Uta Kiltz, MD, Varvara Chatzistefanidi, MD, Xenofon Baraliakos, MD: Ruhr-Universität Bochum, Bochum, Germany; [2]Daniela Adolf, PhD, Maria Kabelitz: Gesellschaft für klinische und Versorgungsforschung mbH Magdeburg, Magdeburg, Germany; [3]Ilka Schwarze, MD: Praxis für Rheumatologie, Leipzig, Germany; [4]Uwe Lange, MD: Universität Gießen, Bad Nauheim, Germany; [5]Jan Brandt-Jürgens, MD: Rheumatologische Schwerpunktpraxis, Berlin, Germany; [6]Edgar Stemmler, PhD, Sabine Sartingen, PhD: AbbVie Deutschland GmbH & Co. KG, Wiesbaden, Germany.

These clinical trial data can be requested by any qualified researchers who engage in rigorous, independent, scientific research, and will be provided following review and approval of a research proposal, Statistical Analysis Plan (SAP), and execution of a Data Sharing Agreement (DSA). Data requests can be submitted at any time after approval in the US and Europe and after acceptance of this manuscript for publication. The data will be accessible for 12 months, with possible extensions considered. For more information on the process or to submit a request, visit the following link: https://www.abbvie.com/our-science/clinical-trials/clinical-trials-data-and-information-sharing/data-and-information-sharing-with-qualified-researchers.html

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25204.

Address correspondence via email to David Kiefer, MD, at davidkiefer@gmx.de.

Submitted for publication February 26, 2023; accepted in revised form July 20, 2023.

SIGNIFICANCE & INNOVATIONS

- Axial spondyloarthritis (AxSpA) disease-specific patient-reported outcomes impact global functioning measured by the Assessment of Spondyloarthritis International Society Health Index, whereas spinal mobility scores, even if objectively assessed by electronic devices, as well as the degree of structural damage in the axial skeleton, have only limited impact.
- Patients with AxSpA may find factors related to spinal mobility less relevant than others or may use coping strategies to overcome these limitations.

nonradiographic axSpA (nr-axSpA) and radiographic axSpA (r-axSpA) depending on the presence or absence of definite structural damage in the sacroiliac joints (SIJ) (5–7). R-axSpA is largely equivalent to the classical ankylosing spondylitis (AS) (8).

Disease activity and physical function assessed by patient-reported outcomes (PROs), such as the Bath AS disease activity index (BASDAI) (9), the AS Disease Activity Score (ASDAS) (10), the Bath AS functional index (BASFI) (11), and spinal mobility measured by physical examination mostly using the Bath AS metrology index (BASMI) (12), are influenced by both inflammation assessed by magnetic resonance imaging and structural changes assessed by conventional radiography (13,14). These issues may result in compromised health status and an impaired quality of life. Restrictions in physical and global functioning are major drivers of costs in this disease (14–17).

To assess the burden of disease, which is certainly based on many factors (18), various instruments, such as the axSpA-specific ASAS Health Index (HI), have been developed (19–21). Although these complex constructs largely rely on the subjective judgement of patients (1,22), other assessments, such as imaging of SIJ and spine, as well as assessment of spinal mobility, have a more objective basis—even though findings reported by radiologists and rheumatologists may not match, be wrong, or be misinterpreted (6,23–26).

Nevertheless, the ASAS HI based on the International Classification of Functioning, Disability and Health (1) was developed in close cooperation with patients to cover the entire spectrum of possible improvements and limitations of global functioning and health in patients with axSpA (21). Additionally, the ASAS HI has been proven to be sensitive to change because of treatment and thus can serve as a reliable instrument to assess change (27). The Short Form-36 (SF-36) is a composite self-report measure designed as a short, generic assessment of health, including physical functioning, physical and emotional roles, bodily pain, general health, vitality, social functioning, and mental health that has been used in axSpA studies (28). The main components of SF-36 are subscores for physical (physical component summary [PCS]) and mental health (mental component summary [MCS]) (1). The former, but not the latter, is frequently impaired in patients with axSpA.

The Epionics SPINE (ES) device is a noninvasive electronic class IIa-certified movement analysis system to objectively assess spinal mobility that has been evaluated in healthy individuals and in several diseases and conditions, including axSpA (26,29–33). In addition to the exact assessment of spinal mobility in angular degrees (range of motion [RoM]), the speed of the executed movement can also be digitally recorded as the range of kinematics (RoK).

In this post hoc analysis of the SPINEtronic study (26), we investigated the relationship between global functioning and clinical measures of disease activity, physical function, spinal mobility, and radiographic damage.

The objective is to study the association between objective measurements of spinal mobility, including speed of spinal movements and global functioning in patients axSpA.

MATERIALS AND METHODS

The SPINEtronic study was designed as a national, noninterventional cross-sectional observational multicenter trial in which consecutive patients 18 years of age or older diagnosed with axSpA by a rheumatologist were prospectively included. For the present analysis, patients were only included if conventional radiographs of the SIJ and spine were available to enable classification according to the modified New York criteria and assessment of radiographic damage. Ethics approval was obtained from the Independent Ethical Committee of the Medical Association of Westphalia-Lippe and the University of Münster (reference number 2014-277-f-S). Written informed consent was obtained from all patients prior to study inclusion.

Assessment of patient demographics, disease characteristics, and disease status. Demographic data and disease characteristics were assessed in all patients. Self-reported questionnaires (PROs), including ASAS HI (sum score: range 0–17) and the physical (PCS) and mental component summary (MCS) of the SF-36, were completed (21,28). Back pain was quantified by using numerical rating scales (NRS, range 0–10). Furthermore, BASFI, BASDAI, and the linear version of BASMI (34) were assessed (27). In addition, all patients underwent electronic measurements of their spinal mobility by the ES device. This included the performance of a choreography of predefined exercises to record spinal mobility, including flexion, extension, rotation, and lateral flexion of the spine. These exercises had to be performed three times in a row and as fast as possible. The ES uses strain gauge sensors attached in predefined positions at the back to provide a sensitive measure of electrical resistance, and thus of the aperture angles, according to the curvature in each of six 50-mm sensor segments (Figure 1). The standardized paravertebral position of the sensors allows it to record movements and rotations outside

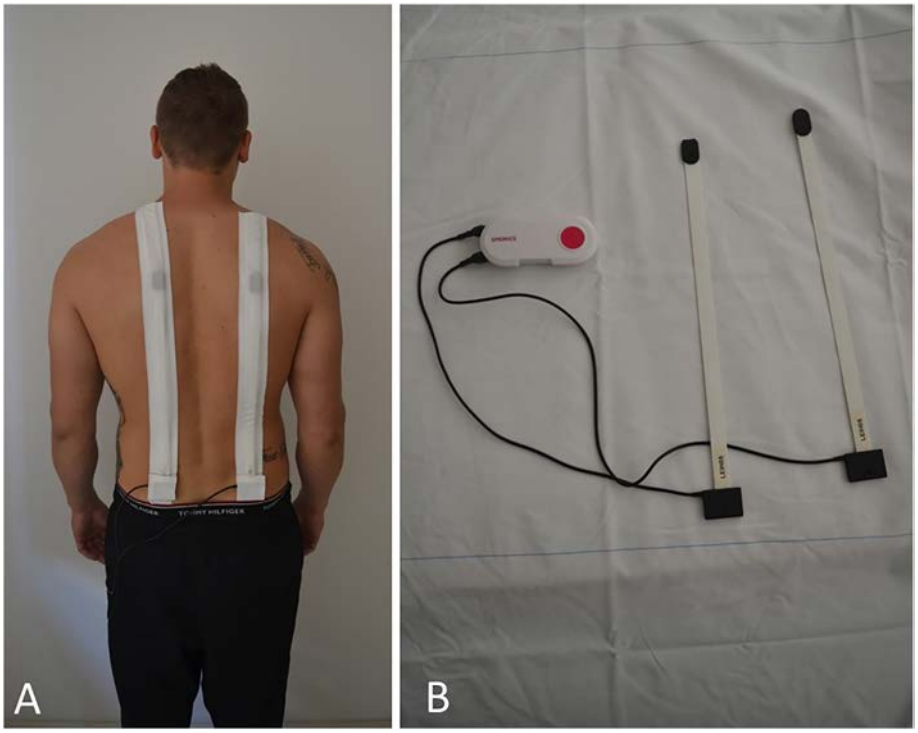


Figure 1. The ES device. **A,** The strain gauge sensors are located along flexible circuit board strips placed at predefined areas of the back. **B,** Sensor strips and memory unit.

of the sagittal plane. The ES is therefore capable to assess the RoM, measured and calculated in angular degrees and the maximum speed with which the exercises have been performed (RoK); this is measured in angular degrees and/or seconds. The accuracy of the system has been shown with an intraclass correlation coefficient (ICC) average greater than 0.98 and a very good test–retest repeatability with ICCs greater than 0.98 (29). The standardized examination with the ES in this study does not exceed 12 minutes and is simple to use because all measurements are automatic and computer guided. This guarantees accurate measures with obviously almost no intra- or interobserver variability. The performance and validity of the ES has been previously demonstrated, which included healthy individuals and patients with low back pain for different reasons. For example, it was shown that the lower lumbar spine retains its lordosis and mobility and becomes less mobile related to increasing age (35,36).

Imaging. Conventional radiographs of the SIJs (performed in routine care) were scored according to the grading system used in the 1984 modified New York criteria (grade 0–4). The sacroiliitis sum score for both SIJs was calculated for analysis as the sum of grades for the left and right SIJ for each patient (sum score ranging between 0 and 8) (37). Radiographs of cervical, thoracic, and lumbar spine were evaluated to count the total number of syndesmophytes if the respective images had been performed in routine care.

Statistical analysis. For the analyses of patient demographics and ES variables and to differentiate between groups, the scores of spinal measurements (ES and BASMI) and the presence and quantification of structural changes in the SIJ and the spine were directly compared by *t*-tests. The correlation between PROs, ES scores, and structural damage was calculated by Kendall's tau correlation coefficients. For categorical data, Fisher's exact test was used. All explorative statistical analyses were performed using the software SAS version 9.4 (SAS Institute) and intentionally calculated to a full significance level of 5%; that is, they were not corrected with respect to multiple testing, and each *P* value less than or equal to 0.05 represents a significant result.

RESULTS

A total of 103 patients were included in the analyses: 72 with r-axSpA (69.9%) and 31 with nr-axSpA (30.1%). Participants (n = 74) were mostly male (71.8%), had a mean (± SD) age 45.8 (11.8), and had a mean (± SD) body mass index (BMI) 27.2 kg/m^2 (± 5.8 kg/m^2), whereas 37 patients (35.9%) had an elevated (>0.5 mg/dL) C-reactive protein (mean ± SD 1.4 ± 4.7 mg/dL, median: 0.4 mg/dL) (Table 1). Overall, 72 patients (77.4%) were human leukocyte antingen-B27 positive (10 values missing). The mean disease duration was 9.0 (± 10.4) years, and the mean (± SD) onset of symptoms was 17.1 years (± 11.8 years) (Table 1). The mean (± SD) ASAS HI score was

Table 1. Patient demographics and disease characteristics comparing r-axSpA and nr-axSpA*

Characteristic	Total (N = 103)	nr-axSpA (n = 31)	r-axSpA (n = 72)	*P* value
Sex, male, n (%)	74 (71.8)	17 (54.8)	57 (79.1)	0.017
Age, y, mean (± SD)	45.8 (± 11.8)	40.9 (± 11.6)	47.9 (± 11.3)	0.006
HLA-B27 (positive), n (%)	72 of 93 (77.4)	20 (64.5)	52 (83.8)†	0.022
BMI, kg/m^2, mean (± SD)	27.2 (± 5.8)	26.7 (± 5.5)	27.4 (± 5.9)	0.558
CRP, mg/dL, mean (± SD)	1.4 (± 4.7)	0.7 (± 1.0)	1.7 (± 5.6)	0.166
Time since symptoms, y, mean (± SD)	17.1 (± 11.8)	12.3 (± 10.8)	19.2 (± 11.7)	0.005
Time since diagnosis, y, mean (± SD)	9.0 (± 10.4)	4.4 (± 7.7)	10.9 (± 10.9)	<0.001
BASDAI, mean (± SD)	4.3 (± 2.2)	4.2 (± 2.1)	4.3 (± 2.2)	0.868
BASFI, mean (± SD)	4.3 (± 2.6)	3.2 (± 2.3)	4.7 (± 2.6)	0.007
BASMI, mean (± SD)	3.2 (± 1.8)	2.0 (± 1.2)	3.7 (± 1.8)	<0.001
ASAS HI, mean (± SD)	7.5 (± 3.6)	7.1 (± 3.5)	7.6 (± 3.7)	0.472
SF-36, PCS, mean (± SD)	36.0 (± 10.1)	38.6 (± 10.7)	34.8 (± 9.7)	0.112
SF-36, MCS, mean (± SD)	43.1 (± 11.9)	42.2 (± 11.1)	43.5 (± 12.3)	0.626
Back pain (NRS 0–10), mean (± SD)	5.2 (± 2.6)	5.3 (± 2.7)	5.1 (± 2.6)	0.786

* AS = ankylosing spondylitis; ASAS HI = Assessment of Spondyloarthritis International Society Health Index; BASDAI = Bath Ankylosing Spondylitis Disease Activity Index; BASFI = Bath Ankylosing Spondylitis Functional Index; BASMI = Bath Ankylosing Spondylitis Metrology Index; BMI = body mass index; CRP = C-reactive protein; HLA-27 = human leukocyte antingen-B27; MCS = mental component summary; nr-axSpA = non-radiographic axial spondyloarthritis; NRS = numerical rating scales; PCS = physical component summary; r-axSpA = radiographic axial spondyloarthritis; SF-36 = Short Form-36.
† HLA-B27 was not tested in 10 patients with r-axSpA.

7.5 (± 3.6). In the population studied, 29.1% patients reported good (≤5), and 53.4% reported moderate (>5 to <12) global functioning, whereas only 12.6% rated themselves as poor (≥12). The mean (± SD) patients' disease activity assessed by BASDAI was 4.3 (± 2.2), whereas the (± SD) level of back pain (NRS 0–10) was 5.2 (± 2.6). Limitations in physical function (BASFI mean [± SD] 4.3 [± 2.6]) and impairments in spinal mobility (BASMI mean [± 3.2], [±1.8]) were documented. SF-36 scores revealed a mean (± SD) PCS of 36.0 (± 10.1) and a mean (± SD) MCS of 43.1 (± 11.9); see Table 1.

Radiographs of the SIJs were available in all patients, and the cervical, thoracic, and lumbar spine radiographs were available in 69, 61, and 97 patients, respectively. Overall, syndesmophytes were present in 43 patients (41.7%) with a mean of 5.9, with 6.2 syndesmophytes per patient. Syndesmophytes in all three spinal segments were seen in 11 patients. As previously reported for axSpA and also for this cohort, ES measurements for RoK and RoM were significantly worse in patients with axSpA than in healthy controls (26,32). The mean (± SD) results for flexion (RoM 26.3 [± 14.8]) and extension (8.9 [± 7.1]) in patients with r-axSpA and nr-axSpA (RoM flexion 40.0 [± 14.0]; extension, 18.0 [± 14.2]) were lower than normal values in the age group of 40 to 49 years (RoM flexion, 51.7 [± 10.2]; extension, 22.2, [± 11.5]) as defined for the ES in a historical cohort (36). ES measurements of RoK and RoM did not significantly correlate with ASAS HI scores (all $P > 0.2$; all r between –0.23 and 0.14) (Table 2).

For PCS and MCS, few of the ES measurements showed weak correlations. In detail, PCS correlated weakly with rotation and lateral flexion (all $P < 0.048$; $r = 0.14$ and 0.19) and MCS with

Table 2. Correlations of patient-reported outcomes, clinical and ES measurements in patients with axSpA*

Correlations	ASAS HI	PCS	MCS
BASMI, r (*P*)	0.14 (0.05)†	−0.24 (<0.001)‡	0.09 (0.21)
Chest expansion, r (*P*)	−0.08 (0.29)	0.14 (0.06)	−0.18 (0.091)
Flexion (RoK), r (*P*)	0.005 (0.95)	0.14 (0.06)	−0.08 (0.24)
Flexion (RoM), r (*P*)	−0.01 (0.85)	0.13 (0.06)	−0.1 (0.17)
Extension (RoK), r (*P*)	0.09 (0.2)	0.01 (0.89)	−0.19 (0.007)†
Extension (RoM), r (*P*)	0.06 (0.4)	0.04 (0.59)	−0.18 (0.01)†
Rotation (RoK), r (*P*)	−0.03 (0.64)	0.14 (0.048)†	−0.08 (0.246)
Rotation (RoM), r (*P*)	−0.09 (0.2)	0.19 (0.008)†	−0.1 (0.16)
Lateral flexion (RoK), r (*P*)	−0.03 (0.71)	0.16 (0.025)†	−0.1 (0.14)
Lateral flexion (RoM), r (*P*)	−0.02 (0.74)	0.16 (0.024)†	−0.15 (0.046)†

* AS = ankylosing spondylitis; ASAS HI = Assessment of Spondyloarthritis International Society Health Index; axSpA = axial spondyloarthritis; BASMI = Bath Ankylosing Spondylitis Metrology Index; ES = Epionics SPINE; MCS = mental component summary; PCS = physical component summary; RoK = range of kinematics; RoM = range of motion.
† Statistically significant correlation ($P < 0.05$).
‡ Statistically significant correlation ($P < 0.001$).

Table 3. Correlations of patient-reported outcomes, clinical and radiographic measurements*

	ASAS HI	PCS	MCS
BASDAI, r (*P*)	0.36 (<0.001)‡	−0.37 (<0.001)‡	−0.23 (<0.002)‡
BASFI, r (*P*)	0.48 (<0.001)‡	−0.54 (<0.001)‡	−0.21 (<0.003)‡
ASAS HI, r (*P*)	—	−0.51 (<0.001)‡	−0.36 (<0.001)‡
PCS, r (*P*)	−0.51 (<0.001)‡	—	0.14 (<0.045)‡
MCS, r (*P*)	−0.39 (<0.001)‡	0.14 (<0.045)‡	—
Back pain, r (*P*)	0.41 (<0.001)‡	−0.32 (<0.001)‡	−0.22 (0.003)†
Nocturnal back pain, r (*P*)	0.4 (<0.001)‡	−0.3 (<0.001)‡	−0.26 (<0.001)‡
Global pain, r (*P*)	0.41 (<0.001)‡	−0.36 (<0.001)‡	−0.24 (0.001)†
Number of syndesmophytes, r (p)	0.03 (0.79)	−0.12 (0.33)	−0.04 (0.77)
Sacroiliitis sum score, r (p)	0.004 (0.96)	−0.11 (0.16)	0.02 (0.78)

* AS = ankylosing spondylitis; ASAS HI = Assessment of Spondyloarthritis International Society Health Index; BASDAI = Bath Ankylosing Spondylitis Disease Activity Score; BASFI = Bath Ankylosing Spondylitis Functional Index; MCS = mental component summary; PCS = physical component summary; RoK = range of kinematics; RoM = range of motion.
† Statistically significant correlation ($P < 0.05$).
‡ Statistically significant correlation ($P < 0.001$).

extension (all $P < 0.01$; extension: RoK, $r = –0.19$; RoM, $r = –0.18$) and RoM of lateral flexion ($P = 0.046$; $r = 0.15$), all other correlations were not significant; see Table 2.

The correlation between BASMI and ASAS HI was weaker ($P = 0.05$; $r = 0.14$) than with PCS ($P = <0.0001$; $r = –0.24$), and BASMI did not correlate with MCS ($P = 0.21$; $r = 0.09$); see Table 2.

Similarly, ASAS HI, as well as MCS, PCS, back pain, and BASDAI, did not correlate significantly with the extent of radiographic damage in the axial skeleton (Table 3).

In contrast, BASFI scores correlated significantly with ASAS HI and also with PCS and MCS (Table 3). BASDAI scores of patients with nr-axSpA and r-axSpA were comparable and correlated with ASAS HI, MCS, and PCS (Tables 1 and 3).

The comparison (mean ± SD) of patients with r-axSpA with nr-axSpA showed that they were older (47.9 ± 11.3 years vs. 40.9 ± 11.6 years, $P = 0.017$), whereas the BMI was comparable (Table 1). The mean (± SD) disease duration was 10.9 (± 10.9) years in patients with r-axSpA and 4.4 (± 7.7) years in patients with nr-axSpA ($P < 0.001$), whereas the mean (± SD) symptom duration was 19.2 (± 11.7) years and 12.3 (± 10.8) years in these groups ($P = 0.005$), respectively (Table 1). BASDAI, MCS, PCS, and levels of back pain were comparable in these groups (Table 1). The mean ± SD BASFI was 4.7 (± 2.6) in patients with r-axSpA and 3.2 (± 2.3) in patients with nr-axSpA ($P = 0.007$). Spinal mobility (BASMI, mean [± SD]) was also worse in patients with r-axSpA than in patients with nr-axSpA: 3.7 (± 1.8) versus 2.0 (± 1.2), respectively ($P < 0.001$). ASAS HI scores (mean ± SD) were not different in patients with r-axSpA (7.5 ± 3.6) and nr-axSpA (7.4 ± 3.6) (Table 1).

DISCUSSION

This study shows that in patients with axSpA, disease-specific PROs have a much higher impact on ASAS HI scores compared with measures of spinal mobility even when objectively assessed by the ES. Even though there were a few weak associations between mobility measures and PCS and MCS subscores of the SF36, these correlations are not as strong as the correlations among the individual PROs.

In addition, this study confirms that assessment of spinal mobility with the ES, an electronic system capable of measuring spinal mobility with high accuracy, provides more information on the range and the speed of spinal motion than what is obtained by physical examination (26,32). Our results also indicate that, in this study population, global functioning does not depend much on the degree of structural changes in the axial skeleton. However, as a limitation, the number of patients with syndesmophytes was relatively low.

On the one hand, patients' subjective perceptions can be influenced by multiple different parameters. On the other hand, PROs, which are capable of covering multiple domains, have the advantage of reflecting the patients' point of view, and they are potentially brief, inexpensive, and not usually prone to be influenced by observer bias. In line with our results, a striking discrepancy between the perceptions and the performance of patients has been previously reported; this has also been seen in those with axSpA (38–44). In general, it is not surprising that PROs correlate better among each other compared with more objective assessments, such as spinal mobility measurements. However, impairment of spinal mobility, regardless of its cause and at least not to the degree patients in this study were affected, may not be that burdensome for patients with axSpA. This is consistent with more recent data showing a limited correlation with structural changes in the axial skeleton of patients with axSpA (45–48). Furthermore, patients with axSpA may develop coping strategies to avoid pain by adaptation most often by decreasing their physical activity, possibly by reducing physical stress, and by changing patterns of movement in everyday life in order to overcome impaired mobility. Such coping strategies may explain the lack of correlation between spinal mobility and structural changes in

contrast to patients' perception of function. There is limited evidence that patients' illness perceptions and coping strategies may have short- and long-term effects on patients' perceptions, and they may also influence PROs (49,50). However, in other studies, back pain, disease activity, and health outcomes clearly improved over two years, whereas illness perceptions and coping strategies did not change (49–51).

Moreover, the results of our study demonstrate that neither the range (RoM) nor the speed of motion (RoK) correlate with ASAS HI scores. This result may indicate that patients are able to develop coping strategies (52) or find alternative movement strategies to bypass their functional limitations. However, patients may also just accept the disadvantage and/or simply get used to it. Whether speed of motion can be positively influenced by medication or physical activity has not been studied to date. Future studies in axSpA should study the speed of motion and what it means to patients in more detail. The ES is able visualize the patient's deficits in RoK and RoM in the movements performed using a spider diagram that is automatically generated immediately after the examination. It can be helpful in developing individualized training programs to directly address the patient's impairments and to assess and visualize changes in disease course. The assessment of spinal mobility and the relationship between global functioning and clinical measures of disease activity, physical function, spinal mobility, and radiographic damage is also of increasing importance because spinal mobility has recently been included in the ASAS-OMERACT core outcome set in axSpA (53). Thus, this is relevant for clinical trials and also as a socioeconomic factor because it is known that decreases in physical function are a major driver of costs in health care systems (17). However, as already mentioned, it is also important for individual patients to monitor the course of their disease. The ES is an objective measurement tool of spinal mobility that is likely to be increasingly important (54).

A limitation of our study is the relatively small sample size, which made it difficult to perform subgroup analyses—for example, according to age and sex. In addition, because of the nature of the study, which only allowed for syndesmophyte counts, it was not possible to quantify structural changes in the spine by an established scoring system, such as the modified Stoke Ankylosing Spondylitis Spinal Score with data collection for radiographic images from daily practice. Nevertheless, our data do still provide useful information for the field of axSpA including nr- and r-axSpA subgroup analyses.

Taken together, this study shows for the first time that global functioning assessed by the ASAS HI is affected by different PROs but only fairly by spinal mobility and speed of motion or radiographic damage.

Thus, our data suggest that patients with axSpA may find factors related to spinal mobility less relevant than others or may use coping strategies to overcome these limitations. More research is needed to explain this in more detail.

ACKNOWLEDGMENT

Open Access funding enabled and organized by Projekt DEAL.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Kiefer, Braun, Chatzistefanidi, Adolf, Schwarze, Lange, Brandt-Jürgens, Stemmler, Sartingen, Baraliakos.

Acquisition of data. Kiefer, Braun, Chatzistefanidi, Lange, Brandt-Jürgens, Baraliakos.

Analysis and interpretation of data. Kiefer, Braun, Kiltz, Chatzistefanidi, Adolf, Schwarze, Kabelitz, Lange, Brandt-Jürgens, Stemmler, Sartingen, Baraliakos.

ROLE OF THE STUDY SPONSOR

AbbVie sponsored this study and contributed to the design, study conduct, and analysis, and interpretation of data; in writing, reviewing and approval of the final version. No honoraria or payments were made for authorship. Statistical analysis support was provided by StatConsult, which was funded by AbbVie.

REFERENCES

1. Kiltz U, Kiefer D, Boonen A. (Health-related) Quality of life as an outcome in studies of axial spondyloarthritis [review]. Rheum Dis Clin North Am 2020;46:379–93.
2. Singh JA, Strand V. Spondyloarthritis is associated with poor function and physical health-related quality of life. J Rheumatol 2009;36:1012–20.
3. Sieper J, Poddubnyy D. Axial spondyloarthritis [review]. Lancet 2017;390:73–84.
4. Feldtkeller E. Age at disease onset and delayed diagnosis of spondyloarthropathies. Z Rheumatol 1999;58:21–30. In German.
5. Rudwaleit M, van der Heijde D, Landewe R, et al. The development of Assessment of SpondyloArthritis International Society classification criteria for axial spondyloarthritis (part II): validation and final selection. Ann Rheum Dis 2009;68:777–83.
6. Van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. Arthritis Rheum 1984;27:361–8.
7. Kiltz U, Baraliakos X, Karakostas P, et al. Do patients with non-radiographic axial spondylarthritis differ from patients with ankylosing spondylitis? Arthritis Care Res (Hoboken) 2012;64:1415–22.
8. Boel A, Molto A, van der Heijde D, et al. Do patients with axial spondyloarthritis with radiographic sacroiliitis fulfil both the modified New York criteria and the ASAS axial spondyloarthritis criteria? Results from eight cohorts. Ann Rheum Dis 2019;78:1545–9.
9. 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;21:2286–91.
10. Machado P, Landewé R, Lie E, et al. Ankylosing Spondylitis Disease Activity Score (ASDAS): defining cut-off values for disease activity states and improvement scores. Ann Rheum Dis 2011;70:47–53.
11. 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;21:2281–5.

12. Jones SD, Porter J, Garrett SL, et al. A new scoring system for the Bath Ankylosing Spondylitis Metrology Index (BASMI). J Rheumatol 1995;22:1609.
13. Maksymowych WP. The role of imaging in the diagnosis and management of axial spondyloarthritis [review]. Nat Rev Rheumatol 2019;15: 657–72.
14. Landewe R, Dougados M, Mielants H, et al. Physical function in ankylosing spondylitis is independently determined by both disease activity and radiographic damage of the spine. Ann Rheum Dis 2009;68: 863–7.
15. Zink A, Braun J, Listing J, et al. Disability and handicap in rheumatoid arthritis and ankylosing spondylitis–results from the German rheumatological database. German Collaborative Arthritis Centers. J Rheumatol 2000;27:613–22.
16. Machado P, Landewe R, Braun J, et al. Both structural damage and inflammation of the spine contribute to impairment of spinal mobility in patients with ankylosing spondylitis. Ann Rheum Dis 2010;69: 1465–70.
17. Ward MM. Functional disability predicts total costs in patients with ankylosing spondylitis. Arthritis Rheum 2002;46:223–31.
18. Wanders A, Landewe R, Dougados M, et al. Association between radiographic damage of the spine and spinal mobility for individual patients with ankylosing spondylitis: can assessment of spinal mobility be a proxy for radiographic evaluation? Ann Rheum Dis 2005;64: 988–94.
19. Ware J Jr, Snow KK, Kosinski M, et al. SF-36 health survey. Manual and interpretation guide. Lincoln (RI): Quality Metric Inc.; 2000.
20. EuroQol Group. EuroQol–a new facility for the measurement of health-related quality of life. Health Policy 1990;16:199–208.
21. Kiltz U, van der Heijde D, Boonen A, et al. Development of a health index in patients with ankylosing spondylitis (ASAS HI): final result of a global initiative based on the ICF guided by ASAS. Ann Rheum Dis 2015;74:830–5.
22. Kiltz U, Tsiami S, Kiefer D, et al. Assessments und outcome-parameter bei axialer Spondyloarthritis. Aktuelle Rheumatologie 2019;44:332–8.
23. Van der Heijde D, Braun J, Deodhar A, et al. Modified stoke ankylosing spondylitis spinal score as an outcome measure to assess the impact of treatment on structural progression in ankylosing spondylitis [review]. Rheumatology (Oxford) 2019;58:388–400.
24. Van Bentum RE, Ibáñez Vodnizza SE, Poblete de la Fuente MP, et al. The Ankylosing Spondylitis Performance Index: reliability and feasibility of an objective test for physical functioning. J Rheumatol 2020; 47:1475–82.
25. Van Weely SF, van Denderen JC, Steultjens MP, et al. Moving instead of asking? Performance-based tests and BASFI-questionnaire measure different aspects of physical function in ankylosing spondylitis. Arthritis Res Ther 2012;14:R52.
26. Kiefer D, Baraliakos X, Adolf D, et al. Successful evaluation of spinal mobility measurements with the Epionics SPINE device in patients with axial spondyloarthritis compared to controls. J Rheumatol 2022;49:44–52.
27. Zochling J. Measures of symptoms and disease status in ankylosing spondylitis: Ankylosing Spondylitis Disease Activity Score (ASDAS), Ankylosing Spondylitis Quality of Life Scale (ASQoL), Bath Ankylosing Spondylitis Disease Activity Index (BASDAI), Bath Ankylosing Spondylitis Functional Index (BASFI), Bath Ankylosing Spondylitis Global Score (BAS-G), Bath Ankylosing Spondylitis Metrology Index (BASMI), Dougados Functional Index (DFI), and Health Assessment Questionnaire for the Spondylarthropathies (HAQ-S) [review]. Arthritis Care Res (Hoboken) 2011;63 Suppl 11:S47–58.
28. Revicki DA, Rentz AM, Luo MP, et al. Psychometric characteristics of the short form 36 health survey and functional assessment of chronic illness Therapy-Fatigue subscale for patients with ankylosing spondylitis. Health Qual Life Outcomes 2011;9:36.
29. Taylor WR, Consmuller T, Rohlmann A. A novel system for the dynamic assessment of back shape. Med Eng Phys 2010;32:1080–3.
30. Consmuller T, Rohlmann A, Weinland D, et al. Comparative evaluation of a novel measurement tool to assess lumbar spine posture and range of motion. Eur Spine J 2012;21:2170–80.
31. Dideriksen JL, Gizzi L, Petzke F, et al. Deterministic accessory spinal movement in functional tasks characterizes individuals with low back pain. Clin Neurophysiol 2014;125:1663–8.
32. Kiefer D, Baraliakos X, Bühring B, et al. Epionics SPINE-use of an objective method to examine spinal mobility in patients with axial spondyloarthritis. Z Rheumatol 2020;79:143–52. In German.
33. Kiefer D, Braun J, Chatzistefanidi V, et al. Clinical relevance of axial radiographic damage in axial spondyloarthritis: evaluation of functional consequences by an objective electronic device. J Rheumatol 2023;jrheum.2022–1240.
34. Van der Heijde D, Deodhar A, Inman RD, et al. Comparison of three methods for calculating the Bath Ankylosing Spondylitis Metrology Index in a randomized placebo-controlled study. Arthritis Care Res (Hoboken) 2012;64:1919–22.
35. Pries E, Dreischarf M, Bashkuev M, et al. The effects of age and gender on the lumbopelvic rhythm in the sagittal plane in 309 subjects. J Biomech 2015;48:3080–7.
36. Dreischarf M, Albiol L, Rohlmann A, et al. Age-related loss of lumbar spinal lordosis and mobility–a study of 323 asymptomatic volunteers. PLoS One 2014;9:e116186.
37. Rodriguez VR, Llop M, Protopopov M, et al. Assessment of radiographic sacroiliitis in anteroposterior lumbar vs conventional pelvic radiographs in axial spondyloarthritis. Rheumatology (Oxford) 2021; 60:269–76.
38. Van Weely SF, van Denderen CJ, van der Horst-Bruinsma IE, et al. Reproducibility of performance measures of physical function based on the BASFI, in ankylosing spondylitis. Rheumatology (Oxford) 2009;48:1254–60.
39. Hoeymans N, Feskens EJ, van den Bos GA, et al. Measuring functional status: cross-sectional and longitudinal associations between performance and self-report (Zutphen Elderly Study 1990-1993). J Clin Epidemiol 1996;49:1103–10.
40. Kempen GI, van Heuvelen MJ, van den Brink RH, et al. Factors affecting contrasting results between self-reported and performance-based levels of physical limitation. Age Ageing 1996;25:458–64.
41. Guralnik JM, Branch LG, Cummings SR, et al. Physical performance measures in aging research. J Gerontol 1989;44:M141–6.
42. Wittink H, Rogers W, Sukiennik A, et al. Physical functioning: self-report and performance measures are related but distinct. Spine (Phila Pa 1976) 2003;28:2407–13.
43. Sager MA, Dunham NC, Schwantes A, et al. Measurement of activities of daily living in hospitalized elderly: a comparison of self-report and performance-based methods. J Am Geriatr Soc 1992;40: 457–62.
44. Terwee CB, van der Slikke RM, van Lummel RC, et al. Self-reported physical functioning was more influenced by pain than performance-based physical functioning in knee-osteoarthritis patients. J Clinical Epidemiol 2006;59:724–31.
45. Protopopov M, Sieper J, Haibel H, et al. Relevance of structural damage in the sacroiliac joints for the functional status and spinal mobility in patients with axial spondyloarthritis: results from the German Spondyloarthritis Inception cohort. Arthritis Res Ther 2017;19:240.
46. Poddubnyy D, Listing J, Haibel H, et al. Functional relevance of radiographic spinal progression in axial spondyloarthritis: results from the GErman SPondyloarthritis Inception Cohort. Rheumatology (Oxford) 2018;57:703–11.

47. Poddubnyy D, Fedorova A, Listing J, et al. Physical function and spinal mobility remain stable despite radiographic spinal progression in patients with ankylosing spondylitis treated with TNF-alpha inhibitors for up to 10 years. J Rheumatol 2016;43:2142–8.
48. Kiltz U, Wiatr T, Redeker I, et al. Effects of patient and disease characteristics on global functioning in patients with axial spondyloarthritis in routine care. Semin Arthritis Rheum 2022;55:152006.
49. Van Lunteren M, Scharloo M, Ez-Zaitouni Z, et al. The impact of illness perceptions and coping on the association between back pain and health outcomes in patients suspected of having axial spondyloarthritis: data from the SPondyloArthritis Caught Early cohort. Arthritis Care Res (Hoboken) 2018;70:1829–39.
50. Peláez-Ballestas I, Boonen A, Vázquez-Mellado J, et al. Coping strategies for health and daily-life stressors in patients with rheumatoid arthritis, ankylosing spondylitis, and gout: STROBE-compliant article. Medicine (Baltimore) 2015;94:e600.
51. Van Lunteren M, Landewé R, Fongen C, et al. Do illness perceptions and coping strategies change over time in patients recently diagnosed with axial spondyloarthritis? J Rheumatol 2020;47:1752–9.
52. Günther V, Mur E, Traweger C, et al. Stress coping of patients with ankylosing spondylitis. J Psychosom Res 1994;38:419–27.
53. Navarro-Compán V, Boel A, Boonen A, et al. The ASAS-OMERACT core domain set for axial spondyloarthritis. Semin Arthritis Rheum 2021;51:1342–9.
54. Garrido-Castro JL, Collantes-Estévez E, Alburquerque-Sendín F, et al. Future challenges and critical approach to metrology in patients with axial spondyloarthritis. Diagnostics (Basel) 2021;11:1533.

Arthritis Care & Research
Vol. 77, No. 3, March 2025, pp 393–401
DOI 10.1002/acr.25199

American College of Rheumatology
Empowering Rheumatology Professionals

Differential Item Functioning on the Cochin Hand Function Scale Among People With Systemic Sclerosis by Language, Sex, and Disease Subtype: A Scleroderma Patient-Centered Intervention Network (SPIN) Cohort Study

Mingyao Xu,[1] Daphna Harel,[1] Marie-Eve Carrier,[2] Linda Kwakkenbos,[3] Susan J. Bartlett,[4] Karen Gottesman,[5] Geneviève Guillot,[6] Laura Hummers,[7] Vanessa L. Malcarne,[8] Michelle Richard,[9] and Brett D. Thombs,[10] on behalf of the SPIN Investigators

Objective. To evaluate the degree that the Cochin Hand Function Scale (CHFS) generates scores that are comparable across language, sex, and disease subtype.

Methods. We included participants enrolled in the Scleroderma Patient-centered Intervention Network (SPIN) Cohort who completed the CHFS at their baseline assessment between April 2014 and September 2020. Confirmatory factor analysis (CFA) was used to test unidimensionality, and multiple indicator multiple cause (MIMIC) models were used for differential item functioning (DIF) analysis based on language, sex, and disease subtype. Both intraclass correlation coefficient (ICC) and Pearson's correlation were calculated using factor scores obtained from unadjusted and DIF-adjusted MIMIC models to evaluate agreement and correlation between scores.

Results. A total of 2,155 participants were included. CFA with covarying error terms supported a good fit of the model ($\chi^2[127] = 1{,}754.671$; $P < 0.001$; Tucker-Lewis index = 0.985; comparative fit index = 0.987; root mean square error of approximation = 0.077). Nine items displayed statistically significant DIF for language of administration, 10 items for sex, and 10 items for disease subtype. However, the overall impact of DIF was negligible when comparing factor scores that did and did not account for DIF (ICC = 0.999; r = 0.999).

Conclusion. The CHFS has score comparability in systemic sclerosis regardless of participants' language, sex, and disease subtype.

INTRODUCTION

Systemic sclerosis (SSc) is a rare chronic autoimmune disease characterized by fibrosis of the skin and internal organs.[1] Digital ulcers, contractures, and deformities of the hand can lead to decreased flexion and limited extension.[2] These symptoms impact hand function and can result in substantial impairment.[3] The Cochin Hand Function Scale (CHFS) was developed to measure the functional ability of the hand among people with rheumatic diseases[4] and has been validated[5,6] and used extensively in patients with SSc.[2,6–9] The self-report CHFS consists of 18 items used to assess a person's ability to perform daily hand-related activities.[4]

The cross-language validity of the CHFS is important in SSc because SSc is a rare disease,[10] and people who complete a scale in different languages are commonly included in the same

No sponsor had any role in the study design; in the collection, analysis, and interpretation of the data; in the writing of the report; or in the decision to submit the article for publication.

Supported by the Canadian Institutes of Health Research, the Lady Davis Institute for Medical Research of the Jewish General Hospital, the Jewish General Hospital Foundation, McGill University, the Scleroderma Society of Ontario, Scleroderma Canada, Sclérodermie Québec, Scleroderma Manitoba, Scleroderma Atlantic, the Scleroderma Association of BC, Scleroderma SASK, Scleroderma Australia, Scleroderma New South Wales, Scleroderma Victoria, and Scleroderma Queensland.

[1]Mingyao Xu, MSc, Daphna Harel, PhD: New York University, New York, New York; [2]Marie-Eve Carrier, MSc: Jewish General Hospital, Montreal, Quebec, Canada; [3]Linda Kwakkenbos, PhD: Radboud University, Nijmegen, The Netherlands; [4]Susan J. Bartlett, PhD: McGill University, Montreal, Quebec, Canada; [5]Karen Gottesman, BA: National Scleroderma Foundation, Los Angeles, California; [6]Geneviève Guillot, PDt: Sclérodermie Québec, Longueuil, Quebec, Canada; [7]Laura Hummers, MD: Johns Hopkins University, Baltimore, Maryland; [8]Vanessa L. Malcarne, PhD: San Diego State University and the University of California, San Diego; [9]Michelle Richard, DSW: Scleroderma Atlantic, Halifax, Nova Scotia, Canada; [10]Brett D. Thombs, PhD: Jewish General Hospital and McGill University, Montreal, Quebec, Canada.

Additional supplementary information cited in this article can be found online in the Supporting Information section (http://onlinelibrary.wiley.com/doi/10.1002/acr.25199).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25199.

Address correspondence via email to Daphna Harel, PhD, at daphna.harel@nyu.edu.

Submitted for publication April 20, 2023; accepted in revised form July 18, 2023.

SIGNIFICANCE & INNOVATIONS

- Hand function is an important contributor to disability in systemic sclerosis (SSc), and the Cochin Hand Function Scale (CHFS) is commonly used in SSc clinical trials and multinational observational studies.
- This is the first study to evaluate if CHFS items display differential item functioning (DIF) by language (English and French), sex, and disease subtype.
- Some CHFS items display DIF for participants taking the CHFS in different languages, are of different sexes, and have different disease subtypes, but the impact on total scale scores is negligible.
- The CHFS can be used and compared among participants with SSc across different languages, sexes, and disease subtypes.

study,[11] especially when the study is carried out in countries or regions with more than one commonly spoken language, such as Canada (eg, French and English). Additionally, for rare diseases such as SSc, international collaboration and recruitment of participants from different countries who use different languages is often necessary to include sufficient numbers of participants in a given study.[7,9,12,13]

In addition, because approximately 85% of people with SSc are female,[14,15] it is important to ensure the measurements obtained from the CHFS are comparable regardless of sex. Previous validations have been done with very small numbers of male participants, and thus it is hard to evaluate the equivalence of measurement. For example, out of 40 participants in the first study that validated the CHFS, then called the Duruöz Hand Index, in SSc, only 6 participants were male.[5]

SSc has two main subtypes—limited and diffuse,[16] and disease severity, which is reflected in subtypes, is an important indicator of hand function.[12] Therefore, it is important to assess the degree to which scores from the CHFS may systematically differ by disease subtype.

Differential item functioning (DIF) occurs when members of one group (eg, English-language responders) have a different expected score on an item compared with members of another group (eg, French-language responders), after controlling for any differences in the construct being measured (eg, hand function).[17,18] Therefore, the responses to an item are influenced, not only by the level of the hand function the person has, but also by the grouping factor (eg, whether they completed the scale in French or English).

The purpose of the study was to evaluate whether (1) the CHFS displays DIF with respect to language (English or French), sex (male or female), and disease subtype (limited or diffuse); and (2) if any identified statistically significant DIF influences CHFS scores to a nonnegligible extent.

METHODS

This was a cross-sectional study evaluating baseline data from the Scleroderma Patient-centered Intervention Network (SPIN) Cohort.[7] For a list of the SPIN Cohort Investigators, please see Appendix A. A protocol was published online prior to study initiation (https://osf.io/qb8m3/). Because of overlap with previous studies, we adopted part of the methods from previous work,[12] including the description of the SPIN Cohort in the Participants and Procedure section, and procedures and study variables in the Measures section. This is in line with guidance from the Text Recycling Research Project.[19]

Participants and procedure. The SPIN Cohort is a convenience sample. Eligible patients at SPIN recruiting sites are invited by the attending physician or a nurse coordinator to participate. Eligible participants must be classified as having SSc according to 2013 American College of Rheumatology (ACR)/European Alliance of Associations for Rheumatology (EULAR) classification criteria; at least 18 years old; and fluent in English, French, or Spanish.[20] After written informed consent is obtained, the recruiting site physician or nurse coordinator completes and submits an online medical data form. An automated email is then sent to participants with instructions on activating their SPIN online account and completing measures. SPIN Cohort participants complete outcome measures via an online portal upon enrollment and subsequently every three months. The SPIN Cohort consists of data from 51 centers in Canada, the US, the United Kingdom, France, Spain, Mexico, and Australia. The SPIN Cohort study was approved by the Research Ethics Committee of the Centre intégré universitaire de santé et de services sociaux du Centre-Ouest-de-l'Île-de-Montréal (#MP-05-2013-150) and by the ethics committees of all recruiting sites.

The present study used baseline assessment data from participants enrolled between April 2014 and September 2020 who completed the CHFS in English or French only, and with complete item-level data for the CHFS and complete data on language of instrument completion, sex, and disease subtype.

Measures. *Sociodemographic and medical data.* Participants provided marital status, years of education, number of cigarettes smoked per week, and number of alcoholic drinks per week. SPIN physicians completed a medical data form that included all items of the 2013 ACR/EULAR SSc classification criteria[20] and provided age, sex, time since the first non-Raynaud phenomenon symptoms and diagnosis, SSc subtype (limited or diffuse cutaneous SSc),[16] presence of overlap syndromes (systemic lupus erythematosus, rheumatoid arthritis, Sjögren syndrome, idiopathic inflammatory myositis), and presence of joint contractures (no/mild [0%–25%] vs moderate/severe [>25%] limit in range of motion). Standard numeric rating scales were completed by patients for Raynaud severity in the past week and

severity of finger ulcers, ranging from 0 (not severe at all) to 10 (unbearable).

In the SPIN Cohort, participants self-report race or ethnicity data using the standard categories that are used in each country. Because categories differ across countries, and categories used in one country may not be recognized by participants from other countries, we characterized study participants by aggregating them as White, Black, and Other. The categories used in each country are presented in Supplemental Material A.

Hand function (CHFS). The 18-item CHFS[4] was developed to measure the ability to perform daily hand-related activities. Items reflecting five content areas (ie, kitchen, dressing oneself, hygiene, writing/typing, other) are scored on a 0 to 5 Likert scale (0 = without difficulty; 5 = impossible). The total score is obtained by adding the scores of all items (range 0–90), and higher scores indicate more difficulty in hand function. Validity and reliability of the CHFS have been confirmed in SSc.[5,6]

Statistical analysis. Descriptive statistics were calculated for all variables and all participants in the sample. We fit a unidimensional confirmatory factor analysis (CFA) model to the CHFS data using a robust weighted least squares variance estimator[21] to test the unidimensionality of the underlying latent trait (hand function). We chose to assess a unidimensional model to evaluate whether the standard practice of scoring the CHFS with a simple summed score is justified. To evaluate the unidimensional model, we determined fit via a mean- and variance-adjusted chi-square test statistic, the comparative fit index (CFI), the Tucker-Lewis index (TLI), and the root mean square error of approximation (RMSEA). The CFI, TLI, and RMSEA indices were prioritized, because the chi-square test is highly dependent on the sample size of the study and may reject the model despite its good fit.[22] Values of CFI and TLI greater than or equal to 0.95 and RMSEA less than or equal to 0.08 were considered to indicate a good fit for the model.[23–25] We also calculated modification indices to recognize item pairs for which measurement errors correlate highly.[26] If there was also theoretical justification for shared effects within these pairs of items, we then allowed their errors to covary if this improved model fit.

We then used multiple indicator multiple cause (MIMIC) models to determine if items of the CHFS exhibited DIF when different groups were compared by language, sex, and disease subtype. We first fit a baseline MIMIC model that only included the path between the grouping variables and the latent variable, hand function.[27,28] This model is a unidimensional CFA model with the additional paths between all groups and the latent variable to capture any mean differences in scores for participants in different groups. Next, we used the constrained baseline approach. Specifically, for each grouping variable, we fit 18 augmented models, each with paths from the grouping variable to the individual CHFS item. For each of the grouping variables, we iterated this process 18 times for each item separately. Meanwhile, we noted the statistical significance of the coefficient of the path between the grouping variables and each item. Once we identified all the items that displayed DIF and which grouping variable(s) were the sources of DIF, we constructed the final MIMIC model by adding paths between all the DIF items and the corresponding grouping variables to the baseline MIMIC model, even if these paths were no longer statistically significant in the final DIF-adjusted MIMIC model to be conservative in our model choice. We did not employ a Type I error correction for the *P* values from the original sets of models to capture any possible DIF across items.

Lastly, we assessed the effect of DIF on latent factor scores. This is important because we included 2,155 participants and, because of the large sample size, we expected to detect statistically significant DIF for potentially many items. Use of an effect size measure indicates whether any statistically significant DIF has an actual, meaningful impact on the reason the CHFS is administered—to obtain scores for hand function for participants. Ideally, clinical decisions are based on highly precise estimates and effect sizes and not on analyses of statistical significance. Therefore, we calculated the agreement between the scores obtained from the MIMIC baseline model and the final DIF-adjusted MIMIC model through the intraclass correlation coefficient (ICC) and its 95% confidence interval (CI).[29] As a secondary measure, we also calculated Pearson's correlation coefficient and its 95% CI.[30] Upon analyzing all the MIMIC models, we identified if any CHFS item exhibited DIF and which grouping variables contributed to DIF, as well as whether any observed DIF impacted the factor scores that were estimated from the participants' responses. A high ICC or correlation would indicate that any statistically significant DIF had meaningful impact, whereas a low ICC or correlation would indicate that, although there was statistically significant DIF, it may not have clinical impact. All analyses were conducted in R,[31] with the CFA and MIMIC models fit using the MplusAutomation package.[32]

Sample size calculation. Recommendations for CFA sample size vary. In the present study, we performed a single-factor CFA and multiple MIMIC models with 18 indicators, using a sample of 2,155 participants. This number significantly surpassed the minimum sample size recommended by many established recommendations and standards[33,34] to ensure a substantial agreement between true sample characteristics and model estimates.

RESULTS

Sample characteristics. Within the SPIN Cohort, 2,240 participants had complete data for all CHFS items with 2,178 of those in English or French. However, only 2,155 participants had complete data for all variables in the CFA and MIMIC model analyses (ie, sex, disease subtype) and were included in this study. There were 1,882 female participants (87.3%) and 273 male

participants (12.7%; see Table 1); 1,459 people responded to the CHFS in English (67.7%) and 696 in French (32.3%); 842 respondents presented with diffuse SSc (39.1%), and 1,313 respondents presented with limited or sine SSc (60.9%). A total of 1,788 (83.0%) self-identified as White. Most participants were married or living as married (61.8%). The mean (±SD) time since first non-Raynaud symptoms was 11.1 (±8.8) years, and the mean (±SD) time since diagnosis was 9.4 (±8.1) years. The mean (±SD) CHFS score was 13.5 (±16.1). There were 63 (2.9%) participants with sine disease subtype who were grouped together with the 1,250 (58.0%) participants with limited disease subtype for all following analyses.

Table 1. Participant demographic and disease characteristics (n = 2,155)*

Variable	Participants
Demographic	
English language, n (%)	1,459 (67.7)
Female sex, n (%)	1,882 (87.3)
Race or ethnicity, n (%)[a,b]	
White	1,788 (83.0)
Black	149 (6.9)
Other	216 (10.0)
Age, years, mean (SD)	55.0 (12.6)
Marital status, n (%)	
Married	1,332 (61.8)
Living as married	196 (9.1)
Separated or divorced	257 (11.9)
Widowed	97 (4.5)
Single	273 (12.7)
Education, years, mean (SD)	14.9 (3.7)
Alcohol consumption (drinks/week), n (%)	
0	1,224 (56.8)
1–7	773 (35.9)
8+	158 (7.3)
Cigarette consumption (cigarettes/d), n (%)	
0	2001 (92.9)
1–9	75 (3.5)
10–19	57 (2.6)
20+	22 (1.0)
Disease characteristics	
Time since onset first non-Raynaud symptom in years, mean (SD)[c]	11.1 (8.8)
Time since diagnosis, n (%)[d]	9.4 (8.1)
Disease subtype, n (%)	
Limited	1,250 (58.0)
Diffuse	842 (39.1)
Sine	63 (3.0)
Patient-reported severity of Raynaud, mean (SD)[e]	3.8 (2.8)
Patient-reported severity of finger ulcers, mean (SD)[f]	1.7 (2.7)
Small joint contractures, n (%)	
None or mild	1,505 (69.8)
Moderate	383 (17.8)
Severe	151 (7.0)
Not available	116 (5.4)
Large joint contractures, n (%)	
None or mild	1,743 (80.9)
Moderate	185 (8.6)
Severe	70 (3.2)
Not available	157 (7.3)
Presence of systemic lupus erythematosus, n (%)	63 (2.9)
Presence of Sjögren syndrome, n (%)	164 (7.6)
Presence of rheumatoid arthritis, n (%)	119 (5.5)
Presence of idiopathic inflammatory myositis, n (%)	107 (5.0)
CHFS total score, mean (SD)	13.5 (16.1)

* CHFS, Cochin Hand Function Scale; SSc, systemic sclerosis.
[a] Because ethnicity/race information is collected differently across countries, it is aggregated here into the categories "White," "Black," and "Other." See Supplementary Material A for further details about race or ethnicity grouping.
[b] n = 2,153.
[c] n = 1975.
[d] n = 2,072.
[e] n = 2,132.
[f] n = 2,131.

Confirmatory factor analysis. A unidimensional CFA model of the CHFS items, in which covariance of item residuals was restricted to zero, resulted in a less than ideal fit (χ^2[135] = 5,232.629; $P < 0.001$; TLI = 0.955; CFI = 0.960; RMSEA = 0.132).

The modification indices suggested allowing error measurements of the following items to covary: items 1 and 2, items 2 and 3, items 2 and 4, items 3 and 4, items 9 and 10, items 9 and 12, items 9 and 17, and items 13 and 14. For example, item 13 measures how well participants can write a short sentence with a pencil or an ordinary pen, and item 14 measures how well participants can write a letter with a pencil or an ordinary pen, which are extremely similar. Because of the high degree of similarity across the content or wording of these CHFS items, we allowed all pairs of items with large modification indices to have correlated covariance terms until the CFA model had adequate fit. Therefore, the CFA model was refitted with allowing the error terms of these items to covary, and the refitted model indicated a good fit (χ^2[127] = 1,754.671; $P < 0.001$; TLI = 0.985; CFI = 0.987; RMSEA = 0.077).

DIF analysis. The baseline MIMIC model with paths between each grouping variable and the latent variable demonstrated good fit (χ^2[178] = 2,173.740; $P < 0.001$; TLI = 0.982; CFI = 0.984; RMSEA = 0.072). The baseline MIMIC model's parameters can be found in Table 2.

Using iterations to identify DIF for each grouping variable, we found that 9 items displayed DIF for the grouping variable of language of CHFS administration, 10 items displayed DIF for the grouping variable of the respondent's sex, and 10 items displayed DIF for the grouping variable of the respondent's disease subtype. See Table 3 for the *P* values of each of statistically significant paths in the MIMIC models.

Table 4 shows the final MIMIC model parameters after correcting for DIF. Estimated group differences on the latent factor did not differ meaningfully depending on whether we controlled for DIF. The difference between the two language groups (French – English) on the latent factor was not statistically significant for either the model with DIF adjustment (standardized mean differences [SMD] = –0.048; 95% CI –0.150 to 0.053; $P = 0.352$) or without adjustment (SMD= –0.049; 95% CI –0.149 to 0.052; $P = 0.343$). The difference between the two sex groups (male –

Table 2. Factors loadings of the baseline MIMIC model*

Variable	Estimate	95% CI
Item 1 Hold bowl	0.874	(0.858–0.890)
Item 2 Raise bottle	0.817	(0.797–0.836)
Item 3 Hold plate	0.857	(0.841–0.874)
Item 4 Pour liquid	0.865	(0.849–0.881)
Item 5 Unscrew lid	0.810	(0.793–0.828)
Item 6 Cut meat	0.887	(0.874–0.899)
Item 7 Prick fork	0.876	(0.857–0.896)
Item 8 Peel fruit	0.895	(0.883–0.907)
Item 9 Button shirt	0.870	(0.858–0.883)
Item 10 Zipper	0.879	(0.865–0.893)
Item 11 Toothpaste tube	0.875	(0.857–0.893)
Item 12 Hold toothbrush	0.865	(0.848–0.882)
Item 13 Write short	0.842	(0.824–0.860)
Item 14 Write letter	0.786	(0.766–0.807)
Item 15 Doorknob	0.885	(0.873–0.897)
Item 16 Cut paper	0.897	(0.884–0.911)
Item 17 Pick up coins	0.854	(0.841–0.867)
Item 18 Turn key	0.906	(0.895–0.917)
Item 2 with Item 1	0.110	(0.095–0.125)
Item 2 with Item 3	0.164	(0.145–0.182)
Item 2 with Item 4	0.144	(0.126–0.161)
Item 3 with Item 4	0.124	(0.107–0.141)
Item 9 with Item 10	0.103	(0.086–0.120)
Item 9 with Item 17	0.069	(0.056–0.082)
Item 11 with Item 12	0.087	(0.071–0.103)
Item 13 with Item 14	0.243	(0.221–0.266)
Hand function on language	−0.049	(−0.149 to 0.052)
Hand function on sex	−0.292	(−0.439 to −0.146)
Hand function on disease subtype	0.638	(0.541–0.735)

* CI, confidence interval; MIMIC, multiple indicator multiple cause.

female) on the latent factor was statistically significant for both the final MIMIC model with DIF adjustment (SMD = –0.282; 95% CI –0.432 to –0.131; $P < 0.001$) and the baseline MIMIC model (SMD = –0.292; 95% CI –0.439 to –0.146; $P < 0.001$). The difference between the two disease subtype groups (diffuse – limited) on the latent factor was statistically significant for both the model with DIF adjustment (SMD = 0.624; 95% CI 0.526–0.722; $P < 0.001$) and without adjustment (SMD = 0.638; 95% CI 0.541–0.735; $P < 0.001$).

The ICC between the factor scores obtained from the baseline MIMIC model and the ones from the final MIMIC model was 0.999 (95% CI 0.999–0.999). Pearson's correlation coefficient between the factor scores obtained from the baseline MIMIC model and the ones from the final MIMIC model was 0.998 (95% CI 0.998–0.998).

DISCUSSION

We tested the unidimensional structure of the CHFS and examined whether there were meaningful differences in measurement properties on the latent variable with three grouping variables—language, sex, and disease subtype—in a sample of participants with SSc. We confirmed the one-dimensionality of the latent trait and found that, although there was statistically significant DIF in items of the CHFS, the overall impact of DIF on scores was negligible.

Although there was statistically significant DIF for 9 items between English- and French-language participants, 10 items between male and female participants, and 10 items between participants with limited and diffuse disease subtype, the cumulative effect of DIF was minimal and did not meaningfully influence estimates of hand function differences of participants, regardless of their language, sex, or disease subtype. The high Pearson's correlation (0.998) and ICC (0.999) between factor scores from models that did and did not account for DIF allowed us to conclude that CHFS scores of French- and English-language, male and female, and diffuse and limited subtype participants can be aggregated and compared without concerns of bias because of

Table 3. *P* values for items displaying DIF*

	Initial Models			MIMIC models		
Variable	Item on language	Item on sex	Item on disease subtype	Item on language	Item on sex	Item on disease subtype
Item 2	<0.001	0.001	–	<0.001	<0.001	–
Item 3	<0.001	–	<0.001	<0.001	–	0.029
Item 4	0.028	0.006	0.035	0.065	0.006	0.186
Item 5	<0.001	<0.001	–	<0.001	<0.001	–
Item 6	0.026	0.001	<0.001	0.095	<0.001	<0.001
Item 8	–	–	0.003	–	–	0.002
Item 9	–	<0.001	–	–	<0.001	–
Item 10	–	–	0.003	–	–	0.020
Item 11	–	–	0.012	–	–	0.035
Item 12	0.006	0.004	0.031	0.013	0.006	0.058
Item 13	<0.001	0.001	0.002	<0.001	<0.001	0.001
Item 14	<0.001	–	–	<0.001	–	–
Item 15	–	0.006	–	–	0.002	–
Item 16	–	–	0.020	–	–	0.011
Item 17	–	<0.001	<0.001	–	<0.001	–
Item 18	0.006	–	–	0.008	–	–

* DIF, differential item functioning; MIMIC, multiple indicator multiple cause.

Table 4. Factors loadings of the final MIMIC model*

Variable	Estimate (95% CI)	Item on language estimate (95% CI)	Item on sex estimate (95% CI)	Item on disease subtype estimate (95% CI)
Item 1	0.874 (0.858–0.890)	–	–	–
Item 2	0.816 (0.797–0.836)	0.195 (0.121–0.270)	−0.198 (−0.304 to −0.092)	–
Item 3	0.858 (0.841–0.874)	0.426 (0.356–0.496)	–	−0.070 (−0.133 to −0.007)
Item 4	0.865 (0.849–0.881)	0.069 (−0.004 to 0.141)	−0.131 (−0.224 to −0.037)	−0.043 (−0.106 to 0.020)
Item 5	0.81 (0.793–0.828)	−0.18 (−0.256 to −0.104)	−0.342 (−0.448 to −0.236)	–
Item 6	0.886 (0.874–0.899)	−0.057 (−0.124 to 0.010)	−0.187 (−0.278 to −0.096)	0.132 (0.072–0.192)
Item 7	0.876 (0.857–0.896)	–	–	–
Item 8	0.895 (0.883–0.906)	–	–	0.093 (0.034–0.151)
Item 9	0.871 (0.858–0.883)	–	0.185 (0.097– 0.273)	–
Item 10	0.879 (0.865–0.893)	–	–	−0.072 (−0.132 to −0.011)
Item 11	0.875 (0.857–0.893)	–	–	−0.078 (−0.15 to −0.006)
Item 12	0.865 (0.848–0.882)	−0.100 (−0.179 to −0.021)	0.140 (0.039–0.241)	0.069 (−0.002 to 0.14)
Item 13	0.842 (0.824–0.86)	−0.170(−0.248 to −0.093)	0.171 (0.079–0.264)	−0.115 (−0.182 to −0.047)
Item 14	0.787 (0.766–0.807)	−0.152 (−0.23 to −0.074)	–	–
Item 15	0.885 (0.873–0.897)	–	−0.142 (−0.233 to −0.05)	–
Item 16	0.897 (0.884–0.911)	–	–	0.079 (0.018–0.14)
Item 17	0.854 (0.840–0.867)	–	0.211 (0.109–0.314)	0.174 (0.112–0.235)
Item 18	0.906 (0.895–0.917)	−0.087 (−0.150 to −0.023)	–	–
Item 1 with Item 2	0.110 (0.095–0.125)	–	–	–
Item 2 with Item 3	0.164 (0.145–0.182)	–	–	–
Item 2 with Item 4	0.144 (0.126–0.161)	–	–	–
Item 3 with Item 4	0.124 (0.106–0.141)	–	–	–
Item 9 with Item 10	0.103 (0.086–0.120)	–	–	–
Item 9 with Item 12	0.087 (0.071–0.103)	–	–	–
Item 9 with Item 17	0.069 (0.056–0.082)	–	–	–
Item 13 with Item 14	0.243 (0.220–0.266)	–	–	–
Hand function on language	−0.048 (−0.150 to 0.053)	–	–	–
Hand function on sex	−0.282 (−0.432 to −0.131)	–	–	–
Hand function on disease subtype	0.624 (0.526–0.722)	–	–	–

* CI, confidence interval; MIMIC, multiple indicator multiple cause.

the grouping factors we studied. The lack of impact of DIF on the CHFS may be due, in part, to the wording of the items. Specifically, all items assess concrete abilities to perform a certain task that requires the use of hands and not abstract concepts. This, in turn, may reduce the likelihood of DIF based on participant characteristics.

The present study is the first to assess DIF of CHFS using MIMIC models and the first to compare measurement properties based on participants' language, sex, and disease subtype. Our findings have important implications for research. This study's result demonstrated the comparability of CHFS scores across English and French languages in SSc. Furthermore, regardless of participants' sex and disease subtype, their CHFS scores can be compared without scaling or DIF correction. Considering SSc is a rare disease, with its overall pooled prevalence of approximately 17.6 per 100,000 people,[35] local or regional samples can be limited. Our study supports the use of the CHFS in larger-scale collaborations and promotes broader use in international participants cohorts, such as the SPIN Cohort. Additionally, interventions and treatments aimed at improving hand functionality have been shown to reduce symptom burden among individuals with SSc to some degree.[8] The CHFS is a valid outcome measure that can be used to measure hand function in patients with SSc across language, sex, and disease severity. Future work may investigate sensitivity to change for the CHFS, therefore allowing it to be used to test interventions and treatments.

There are several noteworthy strengths of our study, including its international cohort recruited from 51 clinical sites, its large sample, and the assessment of measurement properties among people with SSc in multiple languages. Although this study focused on determining the impact of DIF for the CHFS for people with SSc based on their language, sex, and disease subtype, the MIMIC models we used could be applied to other participant populations and other measures for DIF identification and correction.

The present study, however, represents only a first step in using the DIF approach to attempt to standardize processes for validating CHFS among people with SSc with different backgrounds and medical histories. There are also limitations to our study. First, the SPIN Cohort is a convenience sample and thus may not represent the SSc population. For example, the cohort participants completed all the required measures online. Second, the examination of DIF was limited to English- and French-speaking participants, and therefore the generalizability of the findings based on our sample population is unknown. Third, we only examined uniform DIF in this study with the assumption of a constant relationship between measures and grouping variables; we did not examine nonuniform DIF.[36] We only examined the differences in mean across groups and did not examine the patterns. However, because in practice the CHFS is scored with a summed score that does not allow for varying factor loadings, any nonuniform DIF would not influence how the CHFS is scored. Lastly, future research may investigate whether our results are replicable under other well-known methods for DIF detection, such as those that use item response theory methods.

Overall, the results of this study indicated that, although the CHFS displayed statistically significant DIF across language of administration, participant sex, and disease subtype, the impact of this DIF was negligible on scores obtained from the scale. This means that participants' CHFS scores can be compared without DIF adjustment, which supports the use of the CHFS in studies that administer the scale in different languages or recruit participants with SSc of different sexes or with different levels of disease severity.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Xu, Harel, Kwakkenbos, Thombs.

Acquisition of data. Carrier, Kwakkenbos, Bartlett, Gottesman, Guillot, Hummers, Malcarne, Richard, Thombs.

Analysis and interpretation of data. Xu, Harel, Thombs.

REFERENCES

1. Allanore Y, Simms R, Distler O, et al. Systemic sclerosis. Nat Rev Dis Primer 2015;1:15,002.
2. Mouthon L, Carpentier PH, Lok C, et al. Ischemic digital ulcers affect hand disability and pain in systemic sclerosis. J Rheumatol 2014;41:1317–1323.
3. Young A, Namas R, Dodge C, et al. Hand impairment in systemic sclerosis: various manifestations and currently available treatment. Curr Treat Options Rheumatol 2016;2:252–269.
4. Duruöz MT, Poiraudeau S, Fermanian J, et al. Development and validation of a rheumatoid hand functional disability scale that assesses functional handicap. J Rheumatol 1996;23:1167–1172.
5. Brower LM, Poole JL. Reliability and validity of the Duruoz Hand Index in persons with systemic sclerosis (scleroderma). Arthritis Rheum 2004;51:805–809.
6. Rannou F, Poiraudeau S, Berezné A, et al. Assessing disability and quality of life in systemic sclerosis: construct validities of the Cochin Hand Function Scale, Health Assessment Questionnaire (HAQ), Systemic Sclerosis HAQ, and Medical Outcomes Study 36-Item Short Form Health Survey. Arthritis Rheum 2007;57:94–102.
7. Kwakkenbos L, Jewett LR, Baron M, et al. The Scleroderma Patient-centered Intervention Network (SPIN) Cohort: protocol for a cohort multiple randomised controlled trial (cmRCT) design to support trials of psychosocial and rehabilitation interventions in a rare disease context. BMJ Open 2013;3:e003563.
8. Rannou F, Boutron I, Mouthon L, et al. Personalized physical therapy versus usual care for patients with systemic sclerosis: a randomized controlled trial. Arthritis Care Res (Hoboken) 2017;69:1050–1059.
9. Kwakkenbos L, Carrier ME, Welling J, et al. Randomized controlled trial of an internet-based self-guided hand exercise program to improve hand function in people with systemic sclerosis: the Scleroderma Patient-centered Intervention Network Hand Exercise Program (SPIN-HAND) trial. Trials 2022;23:994.

10. Bergamasco A, Hartmann N, Wallace L, et al. Epidemiology of systemic sclerosis and systemic sclerosis-associated interstitial lung disease. Clin Epidemiol 2019;11:257–273.
11. Bairwa D, Kavadichanda CG, Adarsh MB, et al. Cultural adaptation, translation and validation of Cochin Hand Function Scale and evaluation of hand dysfunction in systemic sclerosis. Clin Rheumatol 2021; 40:1913–1922.
12. Kwakkenbos L, Sanchez TA, Turner KA, et al. The association of sociodemographic and disease variables with hand function: a Scleroderma Patient-centered Intervention Network Cohort study. Clin Exp Rheumatol 2018;36 Suppl 113:88–94.
13. Levis AW, Harel D, Kwakkenbos L, et al. Using optimal test assembly methods for shortening patient-reported outcome measures: development and validation of the Cochin Hand Function Scale-6: a Scleroderma Patient-Centered Intervention Network Cohort study: using optimal test assembly to shorten the CHFS. Arthritis Care Res (Hoboken) 2016;68:1704–1713.
14. Peoples C, Medsger TA Jr, Lucas M, et al. Gender differences in systemic sclerosis: relationship to clinical features, serologic status and outcomes. J Scleroderma Relat Disord 2016;1:177–240.
15. Hughes M, Pauling JD, Armstrong-James L, et al. Gender-related differences in systemic sclerosis. Autoimmun Rev 2020;19:102494.
16. LeRoy EC, Black C, Fleischmajer R, et al. Scleroderma (systemic sclerosis): classification, subsets and pathogenesis. J Rheumatol 1988; 15:202–205.
17. Holland PW, Wainer H, eds. Differential item functioning. New York (NY): Routledge; 2012.
18. Chen WH, Revicki D. Differential item functioning (DIF). In: Michalos AC, ed. Encyclopedia of quality of life and well-being research. New York (NY): Springer; 2014:1611–1614.
19. Hall S, Moskovitz C, Pemberton M. Text Recycling Research Project. Understanding text recycling: a guide for researchers. 2021. URL: https://textrecycling.org.
20. Van den Hoogen F, Khanna D, Fransen J, et al. Classification criteria for systemic sclerosis: an American College of Rheumatology/-European League Against Rheumatism collaborative initiative. Ann Rheum Dis 2013;72:1747–1755.
21. Muthén LK, Muthén BO. Mplus user's guide. 8th ed. Los Angeles (CA): Muthén & Muthén; 2017.
22. Reise SP, Widaman KF, Pugh RH. Confirmatory factor analysis and item response theory: 2 approaches for exploring measurement invariance. Psychol Bull 1993;114:552–566.
23. Chen F, Curran PJ, Bollen KA, et al. An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models. Sociol Methods Res 2008;36:462–494.
24. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Model Multidiscip J 1999;6:1–55.
25. Browne MW, Cudeck R. Alternative ways of assessing model fit. Sociol Methods Res 1992;21:230–258.
26. McDonald RP, Ho MH. Principles and practice in reporting structural equation analyses. Psychol Methods 2002;7:64–82.
27. Chun S, Stark S, Kim ES, et al. MIMIC methods for detecting DIF among multiple groups: exploring a new sequential-free baseline procedure. Appl Psychol Meas 2016;40:486–499.
28. Woods CM, Oltmanns TF, Turkheimer E. Illustration of MIMIC-model DIF testing with the schedule for nonadaptive and adaptive personality. J Psychopathol Behav Assess 2009;31:320–330.
29. Ranganathan P, Pramesh CS, Aggarwal R. Common pitfalls in statistical analysis: measures of agreement. Perspect Clin Res 2017;8: 187–191.
30. Liu J, Tang W, Chen G, et al. Correlation and agreement: overview and clarification of competing concepts and measures. Shanghai Arch Psychiatry 2016;28:115–120.
31. R Core Team. R: A language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing; 2021. URL: https://www.R-project.org.
32. Hallquist MN, Wiley JF. MplusAutomation : an R package for facilitating large-scale latent variable analyses in Mplus. Struct Equ Modeling 2018;25:621–638.
33. Mundfrom DJ, Shaw DG, Ke TL. Minimum sample size recommendations for conducting factor analyses. Int J Test 2005;5:159–168.
34. Wolf EJ, Harrington KM, Clark SL, et al. Sample size requirements for structural equation models: an evaluation of power, bias, and solution propriety. Educ Psychol Meas 2013;73:913–934.
35. Bairkdar M, Rossides M, Westerlind H, et al. Incidence and prevalence of systemic sclerosis globally: a comprehensive systematic review and meta-analysis. Rheumatology (Oxford) 2021;60:3121–3133.
36. Woods CM, Grimm KJ. Testing for nonuniform differential item functioning with multiple indicator multiple cause models. Appl Psychol Meas 2011;35:339–361.

APPENDIX A: SPIN STUDY INVESTIGATORS

Investigators in the SPIN Cohort study are as follows: Catherine Fortuné, Ottawa Scleroderma Support Group, Ottawa, Ontario, Canada; Amy Gietzen, National Scleroderma Foundation, Tri-State Chapter, Buffalo, New York, USA; Marie Hudson, McGill University, Montreal, Quebec, Canada; Amanda Lawrie-Jones, Scleroderma Australia, and Scleroderma Victoria, Melbourne, Australia; Maureen D. Mayes, Department of Internal Medicine, University of Texas McGovern School of Medicine, Houston, Texas, USA; Warren R. Nielson, Department of Psychology, Western University, and Lawson Research Institute, London, Ontario, Canada; Maureen Sauvé, Scleroderma Society of Ontario, and Scleroderma Canada, Hamilton, Ontario, Canada; Robyn K. Wojeck, University of Rhode Island, Kingston, Rhode Island, USA; Claire Elizabeth Adams, Jewish General Hospital, Montreal, Quebec, Canada; Richard S. Henry, Jewish General Hospital, Montreal, Quebec, Canada; Shervin Assassi, University of Texas McGovern School of Medicine, Houston, Texas, USA; Andrea Benedetti, McGill University, Montreal, Quebec, Canada; Ghassan El-Baalbaki, Université du Québec à Montréal, Montreal, Quebec, Canada; Kim Fligelstone, Scleroderma & Raynaud's UK, London, UK; Tracy Frech, University of Utah, Salt Lake City, Utah, USA; Monique Hinchcliff, Yale School of Medicine, New Haven, Connecticut, USA; Sindhu R. Johnson, Toronto Scleroderma Program, Mount Sinai Hospital, Toronto Western Hospital, and University of Toronto, Toronto, Ontario, Canada; Maggie Larche, McMaster University, Hamilton, Ontario, Canada; Catarina Leite, University of Minho, Braga, Portugal; Christelle Nguyen, Université Paris Descartes, Université de Paris, Paris, France, and Assistance Publique - Hôpitaux de Paris, Paris, France; Karen Nielsen, Scleroderma Society of Ontario, Hamilton, Ontario, Canada; Janet Pope, University of Western Ontario, London, Ontario, Canada; François Rannou, Université Paris Descartes, Université de Paris, Paris, France, and Assistance Publique - Hôpitaux de Paris, Paris, France; Tatiana Sofia Rodriguez-Reyna, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico City, Mexico; Anne A. Schouffoer, Leiden University Medical Center, Leiden, The Netherlands; Maria E. Suarez-Almazor, University of Texas MD Anderson Cancer Center, Houston, Texas, USA; Christian Agard, Centre Hospitalier Universitaire - Hôtel-Dieu de Nantes, Nantes, France; Marc André, Centre Hospitalier Universitaire Gabriel-Montpied, Clermont-Ferrand, France; Elana J. Bernstein, Columbia University, New York, New York, USA; Sabine Berthier, Centre Hospitalier Universitaire Dijon Bourgogne, Dijon, France; Lyne Bissonnette, Université de Sherbrooke, Sherbrooke, Quebec, Canada; Alessandra Bruns, Université de

Sherbrooke, Sherbrooke, Quebec, Canada; Carlotta Cacciatore, Assistance Publique – Hôpitaux de Paris, Hôpital St-Louis, Paris, France; Patricia Carreira, Servicio de Reumatologia del Hospital 12 de Octubre, Madrid, Spain; Marion Casadevall, Assistance Publique - Hôpitaux de Paris, Hôpital Cochin, Paris, France; Benjamin Chaigne, Assistance Publique - Hôpitaux de Paris, Hôpital Cochin, Paris, France; Lorinda Chung, Stanford University, Stanford, California, USA; Benjamin Crichi, Assistance Publique - Hôpitaux de Paris, Hôpital St-Louis, Paris, France; Robyn Domsic, University of Pittsburgh, Pittsburgh, Pennsylvania, USA; James V. Dunne, St. Paul's Hospital and University of British Columbia, Vancouver, British Columbia, Canada; Bertrand Dunogue, Assistance Publique - Hôpitaux de Paris, Hôpital Cochin, Paris, France; Regina Fare, Servicio de Reumatologia del Hospital 12 de Octubre, Madrid, Spain; Dominique Farge-Bancel, Assistance Publique - Hôpitaux de Paris, Hôpital St-Louis, Paris, France; Paul R. Fortin, CHU de Québec - Université Laval, Quebec, Quebec, Canada; Jessica Gordon, Hospital for Special Surgery, New York City, New York, USA; Brigitte Granel-Rey, Université, and Assistance Publique - Hôpitaux de Marseille, Hôpital Nord, Marseille, France; Aurélien Guffroy, Les Hôpitaux Universitaires de Strasbourg, Nouvel Hôpital Civil, Strasbourg, France; Genevieve Gyger, Jewish General Hospital and McGill University, Montreal, Quebec, Canada; Eric Hachulla, Centre Hospitalier Régional Universitaire de Lille, Hôpital Claude Huriez, Lille, Sabrina Hoa, Centre hospitalier de l'Université de Montréal – CHUM, Montreal, Quebec, Canada; Alena Ikic, CHU de Québec - Université Laval, Quebec, Quebec; Niall Jones, University of Alberta, Edmonton, Alberta, Canada; Nader Khalidi, McMaster University, Hamilton, Ontario, Canada; Kimberly Lakin, Hospital for Special Surgery, New York City, New York, USA; Marc Lambert, Centre Hospitalier Régional Universitaire de Lille, Hôpital Claude Huriez, Lille, France; David Launay, Centre Hospitalier Régional Universitaire de Lille, Hôpital Claude Huriez, Lille, France; Yvonne C. Lee, Northwestern University, Chicago, Illinois, USA; Hélène Maillard, Centre Hospitalier Régional Universitaire de Lille, Hôpital Claude Huriez, Lille, France; Nancy Maltez, University of Ottawa, Ottawa, Ontario, Canada; Joanne Manning, Salford Royal NHS Foundation Trust, Salford, UK; Isabelle Marie, CHU Rouen, Hôpital de Bois-Guillaume, Rouen, France; Maria Martin Lopez, Servicio de Reumatologia del Hospital 12 de Octubre, Madrid, Spain; Thierry Martin, Les Hôpitaux Universitaires de Strasbourg, Nouvel Hôpital Civil, Strasbourg, France; Ariel Masetto, Université de Sherbrooke, Sherbrooke, Quebec, Canada; François Maurier, Uneos - Groupe hospitalier associatif, Metz, France; Arsene Mekinian, Assistance Publique - Hôpitaux de Paris, Hôpital St-Antoine, Paris, France; Sheila Melchor Díaz, Servicio de Reumatologia del Hospital 12 de Octubre, Madrid, Spain; Mandana Nikpour, St Vincent's Hospital and University of Melbourne, Melbourne, Victoria, Australia; Louis Olagne, Centre Hospitalier Universitaire Gabriel-Montpied, Clermont-Ferrand, France; Vincent Poindron, Les Hôpitaux Universitaires de Strasbourg, Nouvel Hôpital Civil, Strasbourg, France; Susanna Proudman, Royal Adelaide Hospital and University of Adelaide, Adelaide, South Australia, Australia; Alexis Régent, Assistance Publique - Hôpitaux de Paris, Hôpital Cochin, Paris, France; Sébastien Rivière, Assistance Publique - Hôpitaux de Paris, Hôpital St-Antoine, Paris, France; David Robinson, University of Manitoba, Winnipeg, Manitoba, Canada; Esther Rodríguez Almazar, Servicio de Reumatologia del Hospital 12 de Octubre, Madrid, Spain; Sophie Roux, Université de Sherbrooke, Sherbrooke, Quebec, Canada; Perrine Smets, Centre Hospitalier Universitaire Gabriel-Montpied, Clermont-Ferrand, France; Vincent Sobanski, Centre Hospitalier Régional Universitaire de Lille, Hôpital Claude Huriez, Lille, France; Robert Spiera, Hospital for Special Surgery, New York City, New York, USA; Virginia Steen, Georgetown University, Washington, DC, USA; Evelyn Sutton, Dalhousie University, Halifax, Nova Scotia, Canada; Carter Thorne, Southlake Regional Health Centre, Newmarket; Pearce Wilcox, St. Paul's Hospital and University of British Columbia, Vancouver, British Columbia, Canada; Mara Cañedo Ayala, Jewish General Hospital, Montreal, Quebec, Canada; Vanessa Cook, Jewish General Hospital, Montreal, Quebec, Canada; Sophie Hu, Jewish General Hospital, Montreal, Quebec, Canada; Bianca Matthews, Jewish General Hospital, Montreal, Quebec, Canada; Elsa-Lynn Nassar, Jewish General Hospital, Montreal, Quebec, Canada; Marieke A. Neyer, Jewish General Hospital, Montreal, Quebec, Canada; Julia Nordlund, Jewish General Hospital, Montreal, Quebec, Canada; Sabrina Provencher, Jewish General Hospital, Montreal, Quebec, Canada.

Arthritis Care & Research
Vol. 77, No. 3, March 2025, pp 402–411
DOI 10.1002/acr.25189

American College of Rheumatology
Empowering Rheumatology Professionals

Patterns of Imaging Requests By General Practitioners for People With Musculoskeletal Complaints: An Analysis From a Primary Care Database

Romi Haas,[1] Alexandra Gorelik,[1] Denise A. O'Connor,[1] Christopher Pearce,[2] Danielle Mazza,[1] and Rachelle Buchbinder[1]

Objective. The aim of this study was to examine imaging requested by general practitioners (GPs) for patients with low back, neck, shoulder, and knee complaints over 5 years (2014–2018).

Methods. This analysis from the Australian Population Level Analysis and Reporting database included patients presenting with a diagnosis of low back, neck, shoulder, and/or knee complaints. Eligible imaging requests included low back and neck x-ray, computed tomography (CT), and magnetic resonance imaging (MRI); knee x-ray, CT, MRI, and ultrasound; and shoulder x-ray, MRI, and ultrasound. We determined number of imaging requests and examined their timing, associated factors, and trends over time. Primary analysis included imaging requests from 2 weeks before diagnosis to 1 year after diagnosis.

Results. There were 133,279 patients (57% low back, 25% knee, 20% shoulder, and 11% neck complaints). Imaging was most common among those with a shoulder (49%) complaint, followed by knee (43%), neck (34%), and low back complaints (26%). Most requests occurred simultaneously with the diagnosis. Imaging modality varied by body region and, to a lesser extent, by gender, socioeconomic status, and primary health network. For low back, there was a 1.3% (95% confidence interval [95% CI] 1.0–1.6) annual increase in proportion of MRI and a concomitant 1.3% (95% CI 0.8–1.8) decrease in CT requests. For neck, there was a 3.0% (95% CI 2.1–3.9) annual increase in proportion of MRI and a concomitant 3.1% (95% CI 2.2–4.0) decrease in x-ray requests.

Conclusion. GPs commonly request early diagnostic imaging for musculoskeletal complaints at odds with recommended practice. We observed a trend towards more complex imaging for neck and back complaints.

INTRODUCTION

Diagnostic imaging for the majority of people with nonspecific regional musculoskeletal complaints has limited value. First, degenerative findings are common in asymptomatic people and increase with age. For example, a systematic review including 33 studies (n = 3,110) found vertebral disc degeneration is present in approximately 37% and 96% of asymptomatic 20- and 80-year-olds, respectively (1). Similarly, other reviews have found knee osteophytes (2) and rotator cuff abnormalities (3) are seen in imaging of asymptomatic people, more commonly in those with increasing age. Second, because imaging abnormalities are so common in asymptomatic people, their clinical relevance in symptomatic individuals is questionable. For example, two recent longitudinal studies, one based on lumbar spine x-ray (4) and the other on MRI (5), reported no association between degenerative findings and current or future back pain. Finally, evidence suggests patient-reported outcomes such as pain and function do not improve in patients who receive imaging compared with those who do not (6,7).

For these reasons, clinical care standards and clinical practice guidelines discourage imaging for regional

Supported by Arthritis Queensland, Arthritis South Australia, Arthritis Australia (Allan and Beryl Stephens grant), and the National Health and Medical Research Council of Australia (grant APP-1113532). Arthritis Australia did not contribute to the conduct of this study. Dr. Buchbinder's work was supported by the National Health and Medical Research Council of Australia (Investigator fellowship).

[1]Romi Haas, BPhysio(Hons), MPH, PhD, Alexandra Gorelik, BA, MSc, Denise A. O'Connor, PhD, Danielle Mazza, MD, MBBS, FRACGP, DRANZCOG, Grad Dip Women's Health, GAICD, CF, Rachelle Buchbinder, MBBS(Hons), MSc, PhD, FRACP, FAHMS: Monash University, Melbourne, Victoria, Australia; [2]Christopher Pearce, PhD, MFM, MBBS: Outcome Health, Blackburn, Victoria, Australia.

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.25189).

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25189.

Address correspondence via email to Romi Haas, BPhysio(Hons), MPH, PhD, at romi.haas@monash.edu.

Submitted for publication March 15, 2023; accepted in revised form June 29, 2023.

SIGNIFICANCE & INNOVATIONS

- This study presents data from a large sample of people with regional musculoskeletal complaints that are broadly representative of the wider population using routinely collected data.
- We observed general practitioners commonly request early diagnostic imaging for musculoskeletal complaints and a trend toward more complex imaging for neck and low back complaints.
- These findings are at odds with recommended practice.
- Multifaceted strategies to improve appropriate imaging requests for people with musculoskeletal complaints are urgently needed.

musculoskeletal complaints unless serious pathology is suspected, there is an unsatisfactory response to conservative care, and/or imaging is likely to change management (8–10). In addition, Choosing Wisely recommendations such as "do not undertake imaging for low back pain for patients without indications of an underlying serious condition" and "do not request shoulder ultrasound to diagnose nonspecific shoulder pain which on clinical evaluation is suggestive of rotator cuff pathology and in which surgery is not planned" (11) have been developed as part of a global campaign encouraging conversations about reducing unnecessary tests, treatments, and procedures.

Despite these recommendations, there is evidence that imaging rates for musculoskeletal complaints have increased over time. For example, a systematic review including 27 studies found a 53.5% increase in complex imaging requested for low back pain over 21 years from 1995–2015 (12). In Australia, the likelihood of general practitioner (GP)–requested imaging tests increased by 9% from one period between 2002 and 2005 to a second period between 2009 and 2012 for patients with shoulder and knee complaints (13). Similarly, the likelihood of GP-requested CT scans for patients with neck complaints more than doubled from 2.6%–5.9% during the same time periods (13).

Monitoring imaging trends over time is important to ensure the provision of high-quality care. However, in Australia, previous studies seeking to understand imaging requests in primary care have relied on data from the Medicare Benefits Schedule (14) or cross-sectional databases such as the Bettering the Evaluation and Care of Health dataset (13), which cannot evaluate changes in GP imaging requests over time for people with specific conditions.

General practice databases using routinely collected data from electronic medical records (EMRs) provide an efficient way of examining imaging requests over time. To our knowledge, these databases have not yet been used to examine patterns of care over time for people with musculoskeletal complaints. The objective of this study was to investigate the imaging requested by GPs for people with low back, neck, shoulder, and knee complaints using prospectively collected longitudinal data among general practices participating in the Population Level Analysis and Reporting (POLAR) database. It specifically examined the proportion of patients with imaging requests, imaging modalities, timing of diagnostic imaging requests, associated factors, and changes in diagnostic imaging requests over a 5-year study period. This study forms part of a larger project that has examined patterns of care for people with musculoskeletal complaints provided also including GP consultations, referrals to other health care providers, and prescriptions for pain relief (15).

MATERIALS AND METHODS

Data source and setting. This is a retrospective longitudinal analysis of deidentified data from the POLAR database. This database extracts patient-related information from every GP/patient encounter directly from the EMRs of consenting general practices (n = 301) within the Primary Health Networks (PHNs) of Eastern Melbourne, South-Eastern Melbourne, and Gippsland within Victoria, Australia. After excluding practices with inconsistent activity recording during the study period, 269 practices were included in the data analysis. The study protocol, including a detailed description of the rationale, aims, and methods (including data cleaning and sample size consideration), has been previously published (15). We conducted the study following the Reporting of studies Conducted using Observational Routinely-collected Data guidelines (16).

Participants. We included patients with at least one GP face-to-face consultation and a diagnosis of an eligible atraumatic low back (≥18 years), and/or neck, shoulder, or knee complaint (≥45 years) between January 1, 2014, and December 31, 2018. These age restrictions were chosen because the prevalence of most musculoskeletal conditions increases markedly after the age of 45 years, except for low back pain, which increases after the age of 18 years (17). We also excluded traumatic injuries because these are not likely to be primarily managed by a GP, imaging may be warranted, and these injuries are more common in people aged 18–44 years (18).

Data extracted. For this analysis, we extracted patient characteristics, dates, and diagnoses of eligible musculoskeletal complaints and dates and modalities of eligible imaging requests, which included low back and neck x-ray, CT, and MRI; knee x-ray, CT, MRI, and ultrasound; and shoulder x-ray, MRI, and ultrasound. Shoulder CT was excluded because it is rarely requested by GPs for atraumatic shoulder pain in our setting. We separated diagnostic imaging from image-guided procedural requests (ie, intraarticular or soft-tissue injections and hydrodilatation). Compared with publicly available Medicare statistics, our dataset captured all GP imaging requests regardless of their funding source. This included those eligible for a Medicare rebate paid

by the Australian government and those that were fully patient funded.

The POLAR database extracts structured data from various fields of the EMR, deidentifies it, and uses a combination of automated and manual processing to code the data so that they can be used for research purposes. Eligible musculoskeletal complaints were derived from diagnoses that are mapped and coded to Systemized Nomenclature of Medicine Clinical Terms within POLAR. A list of eligible diagnostic codes was reported in our protocol (15) and is available from https://clinicalcodes.rss.mhs.man.ac.uk/medcodes/article/174/. Our research team used an inductive coding process to select and categorize eligible imaging records. Our coding accounted for over 95% of the 845,400 diagnostic and procedural imaging records identified (15).

Data analysis. All relevant data were extracted from the POLAR Structured Query Language database and exported into Stata version 15 (StataCorp LP) for data management and analysis. For this analysis, we included imaging requested for an individual patient from 2 weeks prior to the date of the first eligible musculoskeletal diagnosis until 1 year following diagnosis or for patients with an eligible musculoskeletal complaint diagnosed in 2018 until the end of 2018. We included the 2 weeks before the date of diagnosis because imaging often precedes diagnosis (19). In addition, a preliminary analysis showed the majority of images requested in the 6 months before diagnosis occurred in the 2 weeks before (Supplementary Figure 1). We also conducted a sensitivity analysis including imaging requested during the entire follow-up period (until December 31, 2018).

We determined the number (percentage) of patients with at least one eligible diagnostic imaging request (overall and for each modality), number (percentage) and modality of diagnostic imaging requests, and number (percentage) of image-guided procedure requests by body region. Requested imaging for multiple modalities or procedures were counted separately. For the timing, regression, and trend analyses, we only included diagnostic imaging requests. The median (interquartile range [IQR]) time (days) from index diagnosis until the first diagnostic imaging request for each body region was also determined.

Multivariable logistic regression was used to examine the association between diagnostic imaging modality requested (x-ray, CT, MRI, and ultrasound) and patient- and practice-related characteristics including gender, socioeconomic status (lowest quintile or other), residential remoteness (metropolitan or other), practice PHN (Eastern Melbourne, South-Eastern Melbourne, or Gippsland), and body region affected (low back, neck, shoulder, or knee).

We reported odds ratios (ORs) with a 95% confidence interval (95% CI) adjusted for age and time since diagnosis. *P* values less than 0.01 were interpreted as statistically significant to account for multiplicity, and a change in OR of ≥10% was interpreted as clinically relevant. In the absence of published data about what would be clinically important, we determined this a priori based upon our clinical judgement. Gippsland was chosen as the reference PHN because this is a predominantly regional and remote area compared with Eastern Melbourne and South-Eastern Melbourne PHNs, which are predominantly metropolitan (20). Knee was chosen as the body region reference because this was the only site that included all imaging modalities, enabling reporting of the odds of receiving an imaging request of a specific modality for a low back, neck, or shoulder complaint compared with a knee complaint.

Trend analysis was used to examine the longitudinal changes in the proportion of patients with imaging requested and the proportion of imaging requests for each modality and body region between 2014 and 2018. *P* values less than 0.05 were interpreted as statistically significant, and a change in either direction of 1% or more per year was considered clinically relevant also determined a priori based upon clinical judgement. Based on the results of a recent trial that evaluated the effect of audit and feedback for reducing musculoskeletal imaging, a 1% reduction in imaging rate would result in approximately 4,700 fewer imaging requests per year (21). Based upon an estimated average of 1.5 imaging requests per person and assuming an imaged proportion of 25% within our cohort (12), this would translate into a change of 10% or more in the OR.

Ethics approval and consent to participate. This study was approved by the Cabrini Human Research Ethics Committee and Monash University Human Research Ethics Committee (reference numbers 02-21-01-19 and 16975, respectively) and was conducted in accordance with the Declaration of Helsinki. We did not obtain participant consent because all data were anonymized. Outcome Health holds a standing ethics approval for its collection and custodianship of the deidentified data from the Royal Australian College of General Practice. Outcome Health and the individual PHNs granted permission to access the data used in this study.

RESULTS

Study cohort. Our eligible study cohort (133,279 patients, 4,538 GPs, and 269 general practices) has been described previously (22). More than half the cohort were female (n = 73,428, 55%), and approximately two thirds had at least one comorbidity (n = 83,816, 63%). Mean (SD) age of the study cohort at diagnosis was 49.2 (18.5) years for those with low back complaints and 61.9 (12.0), 62.8 (11.8), and 64.2 (11.5) years for those with neck, shoulder, and knee complaints, respectively. Based on diagnostic codes, more than half (n = 76,504, 57%) had a low back complaint, a quarter (n = 33,438) had a knee complaint, a fifth (n = 26,335) had a shoulder complaint, and 11% (n = 14,492) had a neck complaint. This includes one tenth (n = 15,176, 11%)

of the cohort who had multiple body regions affected by an eligible musculoskeletal complaint.

Imaging requests. Over one-third (n = 49,174, 37%) of the cohort had at least one eligible imaging request (diagnostic or procedural) within the eligible study period. There were 76,249 imaging requests overall, with a median (IQR) of 1 (1–2) requests per patient.

Imaging requests varied by musculoskeletal complaint. From the 2 weeks before the index diagnosis to 1 year after diagnosis, patients with a shoulder complaint had diagnostic imaging requested most commonly (n = 12,959, 49% patients), followed by knee (n = 14,405, 43%), neck (n = 4,871, 34%), and low back complaints (n = 19,545, 26%) (Table 1). Ultrasound (n = 12,329, 57% requests) and x-ray (n = 8,718, 40%) were the most frequently requested modality for shoulder complaints, x-rays for knee complaints (n = 13,879, 62%), CT for low back complaints (n = 11,160, 50%), and MRI (n = 2,064, 37%) and x-ray (n = 1,960, 36%) were most frequently requested for neck complaints. Over 1 in 10 patients with a shoulder complaint had at least one image-guided procedure requested compared with less than 1% of patients with low back, neck, and knee complaints. There were 500 requests for shoulder hydrodilatation among 445 (2%) patients with a shoulder complaint.

Many patients had requests for more than one type of diagnostic imaging (n = 6254, 48% patients with at least one shoulder imaging request; n = 3419, 24% for knee; n = 1856, 9% for low back; and n = 504, 10% for neck complaints). The most frequent combinations were shoulder x-ray and ultrasound (n = 5,856, 45% patients with at least one imaging request), knee x-ray and either MRI (n = 1,521, 11%) or ultrasound (n = 1,333, 9%), and low back x-ray and CT (n = 1,070, 5%) (Figure 1).

Timing of diagnostic imaging requests. Almost a third (n = 22,571, 30%) of eligible imaging requests were made within the 2-week period before diagnosis (Supplementary Figure 1). The median timing was on the same day as the diagnosis except for shoulder MRI, knee CT, and low back MRI, which were all requested at a later time (median [IQR] days: 42 [2–147], 11 [0–73], and 3 [–1 to 60]) (Figure 2).

Association between diagnostic imaging and patient- and practice-related characteristics. Body region was the strongest predictor of diagnostic imaging requests (Table 2). Compared with patients with knee complaints, the odds of receiving any imaging request were 49% (OR 0.51, 95% CI 0.49–0.52) and 31% (OR 0.69, 95% CI 0.66–0.73) lower for patients with low back and neck complaints, respectively, and 56% higher (OR 1.56, 95% CI 1.50–1.62) for those with shoulder

Table 1. Number (%) of diagnostic and procedural imaging requests and number of patients (%) with imaging requests by modality and body region within 2 weeks before to 1 year after index diagnosis*

	Total study cohort (133,279 patients) n (%)	Low back (76,504 patients) n (%)	Neck (14,492 patients) n (%)	Shoulder (26,335 patients) n (%)	Knee (33,438 patients) n (%)
Patients with diagnostic imaging requests					
At least one request	48,253 (36.2)	19,545 (25.5)	4,871 (33.6)	12,959 (49.2)	14,405 (43.1)
X-ray	26,232 (19.7)	7,485 (9.8)	1,923 (13.3)	7,787 (29.6)	10,258 (30.7)
CT scan	12,379 (9.3)	10,854 (14.2)	1,464 (10.1)	N/A	160 (0.5)
MRI scan	11,031 (8.3)	3,140 (4.1)	2,012 (13.9)	719 (2.7)	5,332 (15.9)
Ultrasound	13,156 (9.9)	N/A	N/A	10,898 (41.4)	2,440 (7.3)
Requests for diagnostic imaging					
Total	71,865 (100)	22,153 (30.8)†	5,513 (7.7)†	21,812 (30.3)†	22,387 (31.2)†
X-ray	32,313 (45.0)	7,763 (35.0)	1,960 (35.6)	8,718 (40.0)	13,872 (62.0)
CT scan	12,821 (17.8)	11,160 (50.4)	1,489 (27.0)	N/A	172 (0.8)
MRI scan	11,737 (16.3)	3,230 (14.6)	2,064 (37.4)	765 (3.5)	5,678 (25.4)
Ultrasound	14,994 (20.9)	N/A	N/A	12,329 (56.5)	2,665 (11.9)
Patients with procedural imaging request‡					
At least one request	3,731 (2.8)	284 (0.4)	52 (0.4)	3,227 (12.3)	181 (0.5)
Image-guided injection	3,370 (2.5)	284 (0.4)	52 (0.4)	2,866 (10.9)	181 (0.5)
Hydrodilatation	445 (0.3)	N/A	N/A	445 (1.7)	N/A
Requests for procedural imaging‡					
Total	4,384 (100)	327 (7.5)†	61 (1.4)†	3,793 (86.5)†	203 (4.6)†
Image-guided injection	3,884 (88.6)	327 (100)	61 (100)	3,293 (86.8)	203 (100)
Hydrodilatation	500 (11.4)	N/A	N/A	500 (13.2)	N/A

* The number of participants with a musculoskeletal condition affecting each body region sums to more than 133,279, because n = 15,176 participants were diagnosed with musculoskeletal symptoms affecting multiple body regions. For proportion of total radiology requests, each patient may have had requests for multiple images and/or modalities for the same body region. CT = computed tomography; MRI = magnetic resonance imaging; N/A = not applicable.
† Percentage of imaging requests was calculated from the total study cohort (n = 71,865 diagnostic and n = 4,384 procedural).
‡ Image-guided injections were intra-articular or bursal injection of glucocorticoid, or it was not specified or hydrodilatation (arthrographic distension with glucocorticoid and saline, or it was not specified).

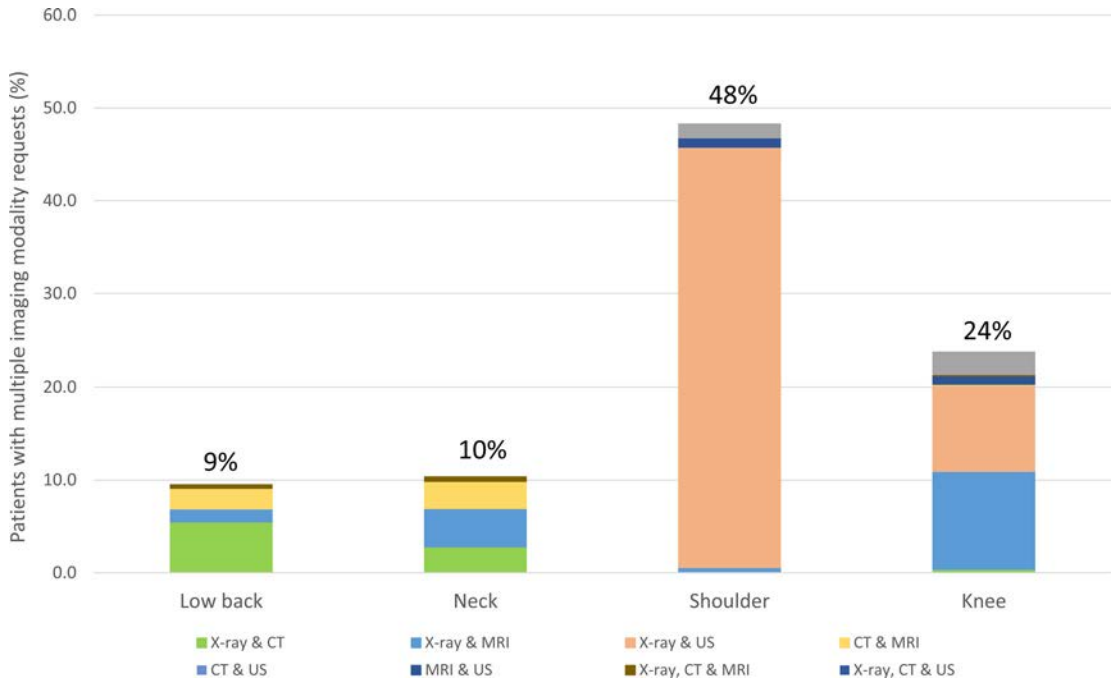


Figure 1. Proportion of patients with multiple diagnostic imaging requests by modality and region. CT = computed tomography; MRI = magnetic resonance imaging; US = ultrasound.

complaints. Compared with patients with knee complaints, those with shoulder complaints were 86% less likely (OR 0.14, 95% CI 0.13–0.15) to receive an x-ray request than those with knee complaints but were 11 times more likely (OR 11.33, 95% CI 10.23–12.56) to receive an ultrasound request.

Irrespective of the musculoskeletal complaint, men were 17% less likely (OR 0.83, 95% CI 0.80–0.87) to receive an x-ray request and 21% more likely to receive requests for CT (OR 1.21, 95% CI 1.15–1.28) and MRI (OR 1.16, 95% CI 1.10–1.23). Patients living in an area of low socioeconomic advantage were more likely to receive a request for an x-ray (OR 1.18, 95% CI 1.10–1.27), CT (OR 1.17, 95% CI 1.08–1.28), and ultrasound (OR 1.23, 95% CI 1.07–1.43) but were less likely to receive an MRI request (OR 0.54, 95% CI 0.48–0.90).

Compared with those attending a Gippsland PHN practice, patients attending predominantly metropolitan practices were less likely (Eastern Melbourne [OR 0.63, 95% CI 0.56–0.71] and South-Eastern Melbourne PHNs [OR 0.63, 95% CI 0.56–0.72]) to receive an x-ray request but were more likely to receive requests for ultrasounds (Eastern Melbourne [OR 2.54, 95% CI 1.98–3.26] and South-Eastern Melbourne [OR 2.44, 95% CI 1.90–3.14]) and MRI scans (Eastern Melbourne [OR 1.69, 95% CI 1.42–2.01] and South-Eastern Melbourne [OR 1.66, 95% CI 1.39–1.97]).

Trends in diagnostic imaging over time. There was no appreciable change in the proportion of participants with imaging requested over the study period (Supplementary Figure 2). However, there was a change in the modalities requested for people with low back and neck conditions. There was a 1.3% (95% CI 1.0–1.6) annual increase in the proportion of requests for low back MRI and a corresponding 1.3% (95% CI 0.8–1.8) decrease in low back CT requests (Figure 3). There was a 3.0% (95% CI 2.1–3.9) annual increase in the proportion of neck MRI requests and a corresponding 3.1% (95% CI 2.2–4.0) reduction in neck x-ray requests. There were no changes over time in the imaging modalities requested for people with shoulder or knee complaints.

Sensitivity analyses. Sensitivity analyses including all eligible diagnostic imaging requests made during the study period (n = 94,451 images requested, 41% patients) did not appreciably change the results (Supplementary Table 1).

DISCUSSION

Our study of routine Australian general practice based upon longitudinal consultation data indicates that GPs frequently request diagnostic imaging for people with regional musculoskeletal complaints, and this most commonly occurs at the time the diagnosis is made. Although we observed no change in overall imaging rates over the 5-year study period, there was a trend toward more complex imaging requests in patients with neck and low back complaints. Diagnostic imaging modality varied by musculoskeletal complaint and, to a lesser extent, by patient socioeconomic status and gender and practice location. We also observed that over 1 in 10 people with a shoulder complaint had at least one image-guided procedure.

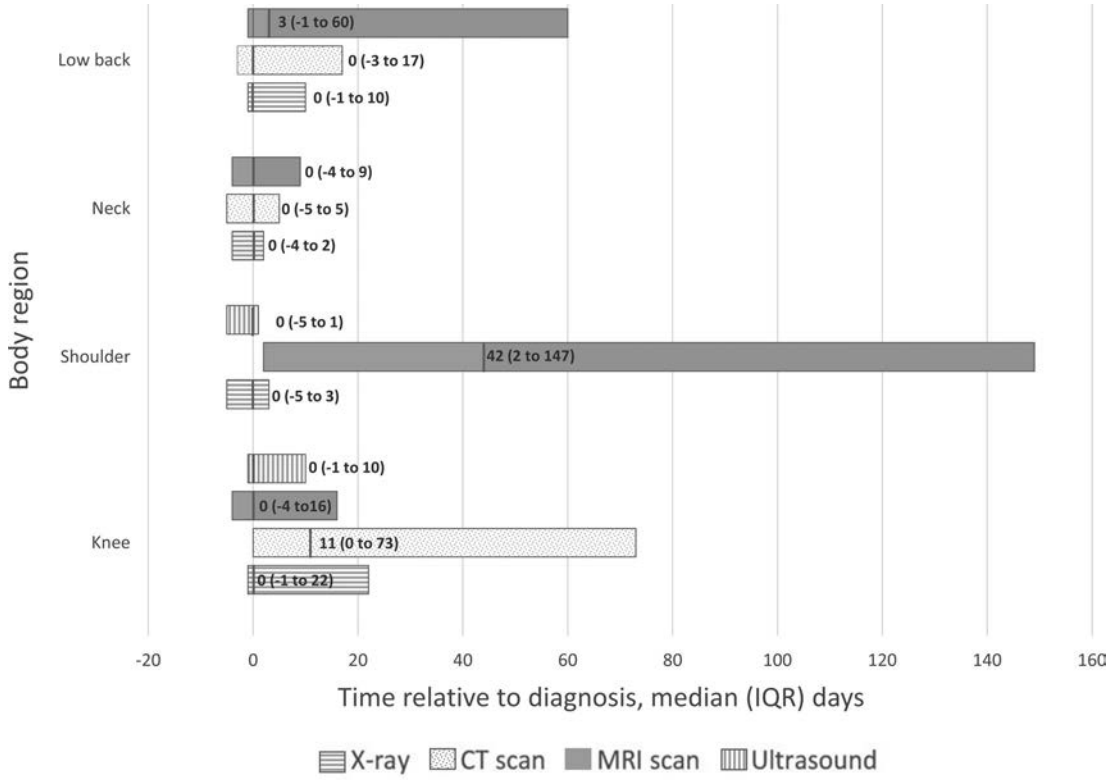


Figure 2. Timing of first diagnostic imaging request by body region and imaging modality, median (interquartile range [IQR]) days since diagnosis. CT = computed tomography; MRI = magnetic resonance imaging.

Our findings are broadly consistent with previous studies. Previous estimates of imaging request rates by Australian GPs have varied between 33% and 43% for shoulder complaints (13,23), 25% and 36% for knee complaints (13,24), and 15% and 24% for patients with low back complaints (13,25). Another study found 23% (95% CI 21–24%) of patients with a first visit for neck pain had imaging requested (25). The slightly lower rates compared with our study is likely because these studies reported point prevalence rates from a single consultation, whereas our study measured cumulative imaging request rates over the course of 1 year.

Clinical practice guidelines across all four musculoskeletal complaints and clinical care standards for low back and knee complaints (8–10) consistently recommend against routine

Table 2. Associations between imaging types and body region affected, patient variables, and GP practice*

	Imaging vs. no imaging, adjusted OR (95% CI)	X-ray only, adjusted OR (95% CI)	CT scan only, adjusted OR (95% CI)	MRI scan only, adjusted OR (95% CI)	Ultrasound only, adjusted OR (95% CI)
Body region affected					
Knee	1	1	N/A[a]	1	1
Low back	0.51 (0.49–0.52)	0.50 (0.48–0.53)	1	0.37 (0.35–0.40)	N/A
Neck	0.69 (0.66–0.73)	0.49 (0.45–0.53)	0.33 (0.30–0.36)	1.82 (1.66–1.98)	N/A
Shoulder	1.56 (1.50–1.62)	0.14 (0.13–0.15)	N/A	0.08 (0.07–0.10)	11.33 (10.23–12.56)
Patient-related variables					
Male	1.01 (0.98–1.03)	0.83 (0.80–0.87)	1.21 (1.15–1.28)	1.16 (1.10–1.23)	1.0 (0.93–1.09)
Lowest socioeconomic quintile	0.83 (0.79–0.86)	1.18 (1.10–1.27)	1.17 (1.08–1.28)	0.54 (0.48–0.90)	1.23 (1.07–1.43)
Metropolitan residential location	0.99 (0.94–1.04)	1.01 (0.92–1.12)	0.94 (0.83–1.07)	0.94 (0.83–1.07)	0.98 (0.83–1.17)
GP practice					
Gippsland	1	1	1	1	1
Eastern Melbourne	0.92 (0.86–0.98)	0.63 (0.56–0.71)	1.16 (0.99–1.36)	1.69 (1.42–2.01)	2.54 (1.98–3.26)
South-Eastern Melbourne	0.97 (0.91–1.04)	0.63 (0.56–0.72)	1.26 (1.08–1.47)	1.66 (1.39–1.97)	2.44 (1.90–3.14)

* Bold indicates statistically signficant ($P < 0.01$). All regression models are adjusted for age and time since index diagnosis. They include participants with a single body region affected by a musculoskeletal complaint. $P < 0.01$ was statistically significant. 95% CI = 95% confidence interval; CT = computed tomography; GP = general practitioner; MRI = magnetic resonance imaging; N/A = not applicable; OR = odds ratio.
[a] There were too few knee CTs to compare (report neck relative to low back).

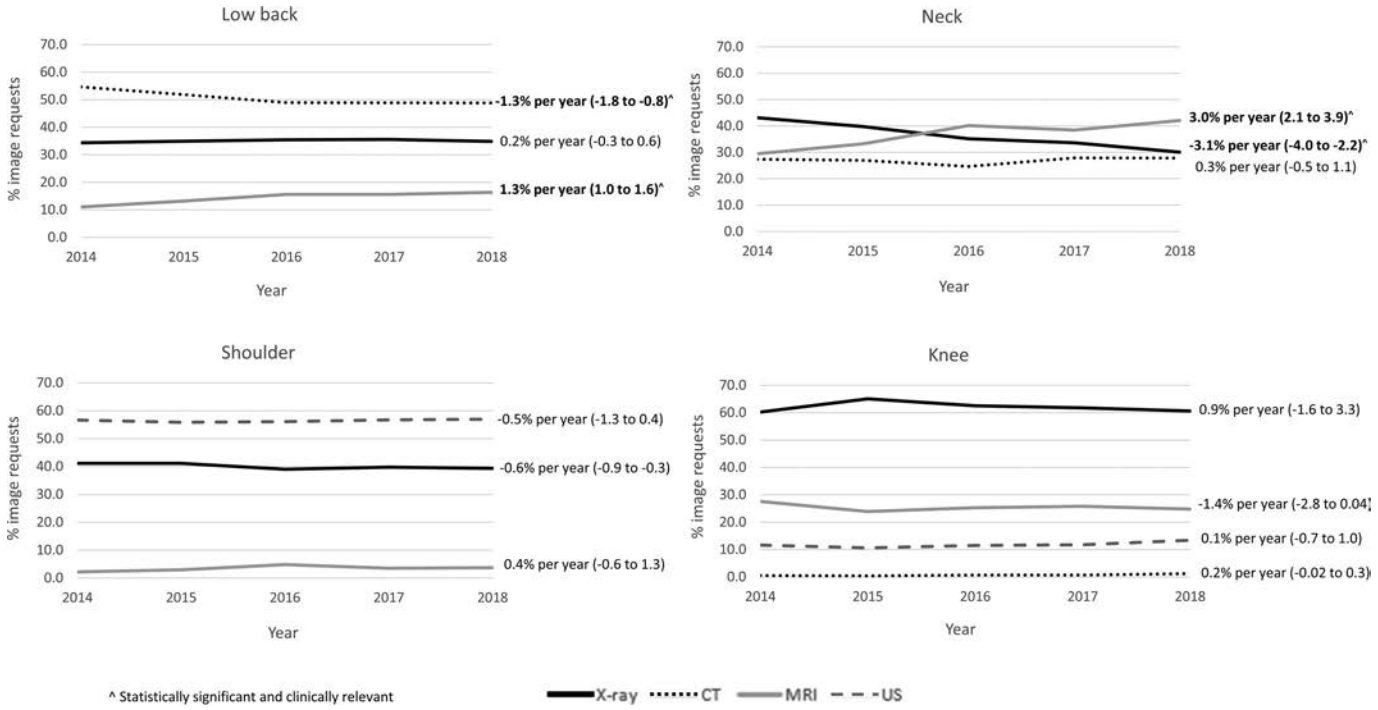


Figure 3. Trends in diagnostic imaging requests over time by modality and body region. CT = computed tomography; MRI = magnetic resonance imaging; US = ultrasound.

imaging unless there are clinical features suggestive of serious pathology. Based upon the small number of people attending general practice with serious pathology, it is likely our findings are at odds with recommended practice. For example, the estimated proportion of people attending primary care with low back pain who have serious pathology ranges from 1% to 6% (26–28). Overuse of imaging for musculoskeletal complaints has also been recognized as a low-value practice that should be questioned across several Australian and international Choosing Wisely recommendations (11,29). Yet, our data indicate that, to date, these have had little if any impact in changing practice.

To our knowledge, this is the first study to examine the timing of GP imaging requests relative to diagnosis. We found the majority of diagnostic imaging requests occurred at or around the same time as the diagnosis with 31% occurring in the 2 weeks before diagnosis, although we cannot discount the possibility that our cohort received previous care for the same complaint or that imaging was performed only after a period of unsatisfactory improvement. Qualitative research, GP surveys, and Australian Medicare Statistics suggest both clinicians and patients continue to have misconceptions about the value of diagnostic imaging for musculoskeletal complaints (24,30–33). Shoulder MRI, knee CT, and low back MRI were the only modalities to be requested after diagnosis, suggesting perhaps that these are requested if initial management does not help or if symptoms do not subside.

Our finding that about one quarter of people who present with low back pain receive diagnostic imaging is consistent with a systematic review that found one in four patients attending primary care receive imaging (12). Our finding of a trend towards more MRI and less CT requests for low back pain over our 5-year study period is also largely in keeping with this review, which found an increase in complex imaging from 7.4% in 1995 to 11.4% in 2015. Other studies have also reported a trend toward more complex imaging over time for low back (12,13,34) and neck complaints (13), although our study has now demonstrated that neck MRI requests have surpassed neck x-ray requests. Our finding of a relative increase in lumbar spine MRI over CT requests may be partially because of a concerted effort to reduce unnecessary radiation exposure from imaging (14) even though GP-requested lumbar spine MRI is nonrebatable in Australia. A move towards more complex imaging is not only concerning because of the heightened risk of overdiagnosis and overtreatment and greater financial cost (35), but also the carbon footprint of MRI and CT is 23 and 12 times greater than plain x-ray, respectively (36), indicating these tests also have a significant environmental impact.

Almost 30% of our cohort with a shoulder complaint received an x-ray request. Although plain radiographs may be worthwhile to diagnose glenohumeral joint arthritis and assess its severity, the radiographic prevalence of this condition in primary care is only 17% and mainly affects older adults (37), indicating likely overuse. Over 40% of our cohort had a shoulder ultrasound request, although the utility of diagnostic ultrasound for shoulder complaints in primary care is of questionable, if any, utility (7). Age-related abnormalities of the rotator cuff in

asymptomatic people are common and may not be of clinical relevance to a patient's complaint (3). This overuse may also partially explain the high number of patients who received an ultrasound-guided shoulder injection. Although there is high-quality evidence that glucocorticoid injection provides worthwhile short-term benefit for people with rotator cuff disease (38), there is also moderate-certainty evidence that ultrasound guidance does not improve pain or function over landmark-guided injection (39). It also has significantly greater cost. Medicare statistics data indicate an increasing number of ultrasound-guided injections are being performed over time, which led to it being identified as a low-value practice, which should be questioned by the Australian Rheumatology Association (11). Other than a lack of awareness of the evidence and promotion of the procedure by vested interests, the removal of subsidized landmark-guided intra-articular steroid injection from the Medicare Benefits Schedule in 2009 (40) may also explain the increased use of ultrasound-guided shoulder injection. Although meta-analyses support the use of hydrodilatation with glucocorticoid over glucocorticoid injection alone for frozen shoulder (41), only 1.7% (n = 445) of our cohort with a shoulder complaint had a request for hydrodilatation compared with nearly 11% for injection. This may be explained by the relatively low prevalence of frozen shoulder compared with rotator cuff disease for which there is also evidence supporting the effectiveness of shoulder injection (42).

Our study has demonstrated disparities in imaging by gender, socioeconomic status, and practice location, which are consistent with known gender, socioeconomic, and geographic disparities in access to health care (43). In particular, we found patients attending practices within predominantly metropolitan PHNs were more likely to receive requests for MRI scans and ultrasound but less likely to receive x-ray requests than those attending practices within predominantly regional and remote PHNs. This is related to limited availability of both ultrasound and MRI services and availability of experts to operate the equipment and interpret results in regional and remote areas (44). For example, less than 4% of Victorian sonographers are known to be located in our predominantly regional and remote PHN (Gippsland) (45), yet this services an area containing 8% of the Victorian population (46). It is also possible these geographic disparities are partially related to supplier-induced demand in metropolitan areas (47).

There are many reasons why practice differs from guideline recommended care. A metasynthesis of 11 studies (n = 270) identified social influence from patients, beliefs that a scan will reassure patients, and a lack of time to discuss why a scan is not needed as the major barriers to reducing imaging for low back pain (48). Policies to address inequitable access to imaging may also inadvertently facilitate inappropriate imaging (49) as well as fee-for-service models that do not remunerate for the time taken to explain why imaging is not necessary and advise on alternate management approaches (50). Successful implementation of tailored interventions to improve the appropriate use of imaging will therefore likely require a multifaceted approach targeting patients, clinicians, and health care policy.

Few interventions have been proven to reduce unnecessary imaging for musculoskeletal complaints. A Victorian mass media campaign that aimed to alter societal and clinician beliefs about low back pain performed in the late 1990s successfully improved beliefs about imaging (51), and this was sustained over time (52). Further study of ways of changing societal views about diagnostic imaging is also necessary. A metasynthesis of qualitative studies found that the general public value the information that imaging provides and also have differences in comprehension and acceptance of overuse concepts (53).

A more recent successful approach was an Australia-wide factorial cluster trial of individualized audit and feedback targeting top requestors of 11 commonly overused musculoskeletal diagnostic imaging tests. This significantly decreased the rate of imaging requested over 6, 12, and 18 months compared with no audit and feedback (54). Further study of this relatively simple, low cost, and easily scalable intervention targeting clinicians to reduce overused diagnostic imaging tests is warranted.

We examined imaging requested by GPs in a large sample of people with regional musculoskeletal complaints that are broadly representative of the wider population (22), which is a study strength. Another strength was that we were able to capture all imaging requests irrespective of whether or not they would attract a government subsidy, whereas Medicare statistics are only able to capture tests that receive a government subsidy. Limitations of our study include that we do not know how many patients received imaging because the POLAR dataset did not include these data. We also could not determine the clinical appropriateness of the imaging requested. Our estimates of the patients receiving imaging requests are likely to be an undercount because we did not include an entire year of follow-up for patients diagnosed in 2018. We excluded younger adults (<45 years) with knee, shoulder, and neck complaints as well. Nonaccredited, corporate-owned general practices and those without EMRs are also likely underrepresented, and our findings may not be generalizable to them. Further, there may be some variability in our estimates of the timing of imaging requests because GPs may record a diagnosis at the first presentation or at a later visit when the diagnosis is confirmed. We also assumed that the imaging request was related to the diagnosis. Although it is possible that an imaging test could be requested for another reason, it is unlikely their frequency would be sufficient to alter our results. Similarly, it is unlikely exclusion of a small proportion of general practices because of inconsistent activity recording (11%) or uncoded tests (<5%) would have substantially altered our findings.

GPs frequently request diagnostic imaging for people with regional musculoskeletal complaints, and this most commonly

occurs simultaneously with the diagnosis. It is likely a substantial proportion of requests are discordant with evidence-based practice. Identification and testing of strategies that target patients, clinicians, and policy to improve appropriate use of imaging in people with musculoskeletal complaints is urgently needed.

ACKNOWLEDGMENT

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

AUTHOR CONTRIBUTIONS

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

Study conception and design. Haas, O'Connor, Buchbinder.
Acquisition of data. Haas, O'Connor, Buchbinder.
Analysis and interpretation of data. Haas, Gorelik, Pearce, Mazza, Buchbinder.

REFERENCES

1. Brinjikji W, Luetmer PH, Comstock B, et al. Systematic literature review of imaging features of spinal degeneration in asymptomatic populations [review]. AJNR Am J Neuroradiol 2015;36:811–6.
2. Culvenor AG, Øiestad BE, Hart HF, et al. Prevalence of knee osteoarthritis features on magnetic resonance imaging in asymptomatic uninjured adults: a systematic review and meta-analysis. Br J Sports Med 2019;53:1268–78.
3. Teunis T, Lubberts B, Reilly BT, et al. A systematic review and pooled analysis of the prevalence of rotator cuff disease with increasing age [review]. J Shoulder Elbow Surg 2014;23:1913–21.
4. Chen L, Perera RS, Radojčić MR, et al. Association of lumbar spine radiographic changes with severity of back pain–related disability among middle-aged, community-dwelling women. JAMA Netw Open 2021;4:e2110715.
5. Kasch R, Truthmann J, Hancock MJ, et al. Association of lumbar MRI findings with current and future back pain in a population-based cohort study. Spine (Phila Pa 1976) 2022;47:201–11.
6. Karel YH, Verkerk K, Endenburg S, et al. Effect of routine diagnostic imaging for patients with musculoskeletal disorders: a meta-analysis [review]. Eur J Int Med 2015;26:585–95.
7. Ottenheijm RP, Cals JW, Winkens B, et al. Ultrasound imaging to tailor the treatment of acute shoulder pain: a randomised controlled trial in general practice. BMJ Open 2016;6:e011048.
8. Australian Commission on Safety and Quality in Health Care. Osteoarthritis of the knee clinical care standard. 2017. URL: https://www.safetyandquality.gov.au/standards/clinical-care-standards/osteoarthritis-knee-clinical-care-standard.
9. Australian Commission on Safety and Quality in Health Care. Low back pain clinical care standard. 2022. URL: https://www.safetyandquality.gov.au/standards/clinical-care-standards/low-back-pain-clinical-care-standard.
10. Lin I, Wiles L, Waller R, et al. What does best practice care for musculoskeletal pain look like? Eleven consistent recommendations from high-quality clinical practice guidelines: systematic review. Br J Sports Med 2020;54:79–86.
11. Morrisroe K, Nakayama A, Soon J, et al. EVOLVE: the Australian Rheumatology Association's 'top five' list of investigations and interventions doctors and patients should question. Intern Med J 2018;48:135–43.
12. Downie A, Hancock M, Jenkins H, et al. How common is imaging for low back pain in primary and emergency care? Systematic review and meta-analysis of over 4 million imaging requests across 21 years. Br J Sports Med 2020;54:642–51.
13. Britt H, Miller GC, Valenti L, et al. Evaluation of imaging ordering by general practitioners in Australia: 2002–03 to 2011–12 Family Medicine Research Centre. Parramatta, New South Wales (Australia): Sydney University Press; 2014.
14. Morgan T, Wu J, Ovchinikova L, et al. A national intervention to reduce imaging for low back pain by general practitioners: a retrospective economic program evaluation using Medicare Benefits Schedule data. BMC Health Serv Res 2019;19:983.
15. Haas R, Busija L, Gorelik A, et al. Patterns of care for people presenting to Australian general practice with musculoskeletal complaints based on routinely collected data: protocol for an observational cohort study using the Population Level Analysis and Reporting (POLAR) database. BMJ Open 2021;11:e055528.
16. Benchimol EI, Smeeth L, Guttmann A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. PLoS Med 2015;12:e1001885.
17. Australian Institute of Health and Welfare. Musculoskeletal conditions and comorbidity in Australia. 2019. URL: https://www.aihw.gov.au/reports/chronic-musculoskeletal-conditions/musculoskeletal-conditions-comorbidity-australia/summary.
18. Pollak AN, Watkins-Castillo SI. The burden of musculoskeletal diseases in the United States: musculoskeletal injuries by type for age and sex. 2014. URL: https://www.boneandjointburden.org/2014-report/via22/musculoskeletal-injuries-type-age-and-sex.
19. Blankenbaker DG. Diagnostic imaging: musculoskeletal non-traumatic. Amsterdam (The Netherlands): Elsevier Health Sciences; 2016.
20. Australian Government Department of Health and Aged Care. Victoria Primary Health Networks (PHN) resource collection. 2018. URL: https://www.health.gov.au/resources/collections/victoria-primary-health-networks-phn-resource-collection#gippsland-phn.
21. O'Connor DA, Glasziou P, Maher CG, et al. Effect of audit and feedback on rates of musculoskeletal diagnostic imaging by high requesting Australian general practitioners: a nation-wide factorial cluster randomised trial. JAMA 2022;328(9):850–60. https://jamanetwork.com/journals/jama/fullarticle/2795890
22. Haas R, Gorelik A, Busija L, et al. Prevalence and characteristics of musculoskeletal complaints: an analysis from the POpulation Level and Analysis Reporting (POLAR) database. BMC Primary Care 2023;24:40.
23. Naunton J, Harrison C, Britt H, et al. General practice management of rotator cuff related shoulder pain: a reliance on ultrasound and injection guided care. PloS One 2020;15:e0227688.
24. Bennell KL, Bayram C, Harrison C, et al. Trends in management of hip and knee osteoarthritis in general practice in Australia over an 11-year window: a nationwide cross-sectional survey. Lancet Reg Health West Pac 2021;12:100187.
25. Michaleff ZA, Harrison C, Britt H, et al. Ten-year survey reveals differences in GP management of neck and back pain. Eur Spine J 2012;21:1283–9.
26. Budtz CR, Hansen RP, Thomsen JN, et al. The prevalence of serious pathology in musculoskeletal physiotherapy patients–a nationwide register-based cohort study. Physiotherapy 2021;112:96–102.
27. Enthoven WT, Geuze J, Scheele J, et al. Prevalence and "red flags" regarding specified causes of back pain in older adults presenting in general practice. Phys Ther 2016;96:305–12.

28. Henschke N, Maher CG, Refshauge KM, et al. Prevalence of and screening for serious spinal pathology in patients presenting to primary care settings with acute low back pain. Arthritis Rheum 2009; 60:3072–80.
29. Canadian Academy of Sport and Exercise Medicine. Sport and exercise medicine: five things physicians and patients should question. 2018. URL: https://choosingwiselycanada.org/wp-content/uploads/2017/05/Sport-and-exercise-medicine.pdf.
30. Buchbinder R, Staples MP, Shanahan EM, et al. General practitioner management of shoulder pain in comparison with rheumatologist expectation of care and best evidence: an Australian national survey. PLoS One 2013;8:e61243.
31. Sharma S, Traeger AC, Reed B, et al. Clinician and patient beliefs about diagnostic imaging for low back pain: a systematic qualitative evidence synthesis. BMJ Open 2020;10:e037820.
32. Artus M, van der Windt DA, Afolabi EK, et al. Management of shoulder pain by UK general practitioners (GPs): a national survey. BMJ Open 2017;7:e015711.
33. Smith L, Barratt A, Buchbinder R, et al. Trends in knee magnetic resonance imaging, arthroscopies and joint replacements in older Australians: still too much low-value care? ANZ J Surg 2020;90: 833–9.
34. Wang Y, Hussain SM, Wluka AE, et al. Rates, costs and determinants of lumbar spine imaging in population-based women born in 1973–1978: data from the Australian Longitudinal Study on Women's Health. PloS One 2020;15:e0243282.
35. Australian Government: Department of Health. Medicare benefits schedule book: Operating from 1 March 2022. 2022. URL: http://www6.health.gov.au/internet/mbsonline/publishing.nsf/Content/E528C73047125137CA258797000C9597/$File/MBSBook%201%20March%202022-v2.pdf.
36. McAlister S, McGain F, Petersen M, et al. The carbon footprint of hospital diagnostic imaging in Australia. Lancet Reg Health West Pac 2022;24:100459.
37. Oh JH, Chung SW, Oh CH, et al. The prevalence of shoulder osteoarthritis in the elderly Korean population: association with risk factors and function. J Shoulder Elbow Surg 2011;20:756–63.
38. Hopewell S, Keene DJ, Marian IR, et al. Progressive exercise compared with best practice advice, with or without corticosteroid injection, for the treatment of patients with rotator cuff disorders (GRASP): a multicentre, pragmatic, 2 × 2 factorial, randomised controlled trial. Lancet 2021;398:416–28.
39. Zadro J, Rischin A, Johnston RV, et al. Image-guided glucocorticoid injection versus injection without image guidance for shoulder pain [review]. Cochrane Database Syst Rev 2021;8:CD009147.
40. Herbert B, Australian Broadcasting Corporation. Arthritis injection rebates axed. 2009. URL: https://www.abc.net.au/news/2009-11-02/arthritis-injection-rebates-axed/1126588.
41. Lädermann A, Piotton S, Abrassart S, et al. Hydrodilatation with corticosteroids is the most effective conservative management for frozen shoulder [review]. Knee Surg Sports Traumatol Arthrosc 2021;29: 2553–63.
42. Ostor A, Richards C, Prevost A, et al. Diagnosis and relation to general health of shoulder disorders presenting to primary care. Rheumatology (Oxford) 2005;44:800–5.
43. Meyer SB, Luong TC, Mamerow L, et al. Inequities in access to healthcare: analysis of national survey data across six Asia-Pacific countries. BMC Health Serv Res 2013;13:238.
44. Parliament of Australia, Senate Community Affairs References Committee. Availability and accessibility of diagnostic imaging equipment around Australia. 2018. URL: https://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Community_Affairs/Diagnosticimaging/Report.
45. Victoria State Government. Victorian Allied Health workforce research project: sonography workforce report. Melbourne (Australia): State of Victoria, Department of Health and Human Services; 2016. URL: https://www.health.vic.gov.au/sites/default/files/migrated/files/collections/research-and-reports/v/victorian-allied-health-workforce-project-2016/sonography-workforce-report.pdf.
46. Australian Bureau of Statistics. 3218.0 - regional population growth, Australia 2017–2018: population estimates by local government area (ASGS 2018), 2017 to 2018. 2019. URL: https://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/3218.02017-18.
47. Baras JD, Baker LC. Magnetic resonance imaging and low back pain care for Medicare patients. Health Aff (Millwood) 2009;28:w1133–40.
48. Hall AM, Scurrey SR, Pike AE, et al. Physician-reported barriers to using evidence-based recommendations for low back pain in clinical practice: a systematic review and synthesis of qualitative studies using the Theoretical Domains Framework. Implement Sci 2019; 14:49.
49. Docking S, Haddock R, Buchbinder R. Australian health policies related to diagnostic imaging: too much of a good thing? Aust Health Rev 2022;46:635–8.
50. Pike A, Patey A, Lawrence R, et al. Barriers to following imaging guidelines for the treatment and management of patients with low-back pain in primary care: a qualitative assessment guided by the Theoretical Domains Framework. BMC Prim Care 2022;23:143.
51. Buchbinder R, Jolley D, Wyatt M. Population based intervention to change back pain beliefs and disability: three part evaluation. BMJ 2001;322:1516–20.
52. Buchbinder R, Jolley D. Effects of a media campaign on back beliefs is sustained 3 years after its cessation. Spine (Phila Pa 1976) 2005;30: 1323–30.
53. Rozbroj T, Haas R, O'Connor D, et al. How do people understand overtesting and overdiagnosis? Systematic review and meta-synthesis of qualitative research [review]. Soc Sci Med 2021;285: 114255.
54. O'Connor DA, Glasziou P, Maher CG, et al. Effect of an individualized audit and feedback intervention on rates of musculoskeletal diagnostic imaging requests by Australian general practitioners: a randomized clinical trial. JAMA 2022;328:850–60.

Arthritis Care & Research
Vol. 77, No. 3, March 2025, pp 412–418
DOI 10.1002/acr.25192

American College of Rheumatology
Empowering Rheumatology Professionals

BRIEF REPORT

Urban and Rural Patterns of Health Care Utilization Among People With Rheumatoid Arthritis and Osteoarthritis in a Large US Patient Registry

Luke W. Desilet,[1] Sofia Pedro,[2] Patricia Katz,[3] and Kaleb Michaud[4]

Objective. Rural residence has been associated with health disparities in rheumatic diseases and other chronic conditions in the United States. This study aimed to determine if a relationship exists between geographic residence and health care utilization outcomes for people with rheumatoid arthritis (RA) and osteoarthritis (OA) in a US-wide rheumatic disease registry.

Methods. Participants were in FORWARD, The National Databank for Rheumatic Diseases, a US-wide rheumatic disease longitudinal cohort completing questionnaires between 1999 and 2019. Health care utilization variables (ie, medical visits and diagnostic tests) from six-month questionnaires were analyzed by geographic categories (small rural/isolated, large rural, and urban). Double selection LASSO with Poisson regression was used to assess the best model when examining the association between health care utilization variables and geographic residence.

Results. Among 37,802 participants with RA, urban residents were more likely than small rural residents to use in-person health care by most measures including physician visits and diagnostic tests. Urban residents reported more rheumatologist visits (incidence rate ratio [IRR], 1.22; 95% confidence interval [95% CI], 1.18–1.27) but fewer primary care visits (IRR 0.90; 95% CI 0.85–0.94). Among 8,248 participants with OA, urban residents were also more likely than rural residents to report health care utilization by most measures.

Conclusion. Individuals residing in urban areas were more likely than those in rural areas to report in-person health care utilization. Specifically, urban residents with RA were more likely to report rheumatologist visits, but less likely to report primary care visits. Less disparity existed in OA health care utilization, although an urban–rural disparity still existed by most measures.

INTRODUCTION

Rural residence has been associated with disparities in health care-related outcomes in both rheumatic diseases and other chronic conditions in the United States (1). Studies have demonstrated that those with rheumatic disease in rural areas have more osteoarthritis (OA)–related disability, higher rates of surgical reconstruction for rheumatoid hand deformities, and lower rheumatoid arthritis (RA)–associated health-related quality of life (2–4).

In people with RA, multiple studies have shown that access to health care with a rheumatologist is associated with improved quality of care (5–7). Rural and micropolitan areas are generally more susceptible to ongoing shortages of rheumatologists in the United States and current workforce projections predict a shortage and maldistribution of rheumatologists that appears to be increasing (8). Although innovative care models including telehealth are promising for addressing rural disparities in access to care, they are not a panacea. The COVID-19 pandemic has highlighted the complexity of providing rheumatic disease care to rural communities. Recent work has demonstrated that in community-based rheumatology practices, rural residents were more likely to stop medications or experience interruptions in receiving medications prescribed for rheumatic disease early

[1]Luke W. Desilet, DO: University of Nebraska Medical Center, Omaha; [2]Sofia Pedro, MSc: FORWARD - The National Databank for Rheumatic Diseases, Wichita, Kansas; [3]Patricia Katz, PhD: University of California San Francisco, San Francisco; [4]Kaleb Michaud, PhD: University of Nebraska Medical Center, Omaha and FORWARD - The National Databank for Rheumatic Diseases, Wichita, Kansas.

Author disclosures are available at https://onlinelibrary.wiley.com/doi/10.1002/acr.25192.

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.25192).

Address correspondence via email to Kaleb Michaud, PhD, at kmichaud@unmc.edu.

Submitted for publication February 28, 2023; accepted in revised form July 6, 2023.

SIGNIFICANCE & INNOVATIONS

- Disparities exist in rheumatic disease outcomes for individuals residing in rural areas compared with their urban counterparts.
- This study is among the first to comprehensively assess patient-reported health care utilization disparities by urban/rural residence specifically in rheumatoid arthritis and osteoarthritis care.
- This analysis of >45,000 participants demonstrated significant health care utilization disparities among rural residents, especially relating to rheumatoid arthritis-specific care.

in the COVID-19 pandemic in the United States (9). Although the implementation and optimization of telehealth and other care delivery methods is ongoing in rheumatology, a current assessment of in-person health care utilization for rheumatic disease is warranted. We propose that a potential urban–rural health disparity exists in access to in-person care, which may contribute to other observed urban–rural rheumatic disease disparities. Work is needed to characterize the delivery of rheumatic disease care to geographically diverse communities to help address potential gaps and opportunities for improvement.

This study sought to examine the relationship between urban versus rural residence and in-person health care utilization for people with RA and OA in the FORWARD Databank, a US-wide rheumatic disease longitudinal registry. It intends to be a cross-sectional snapshot prior to the COVID-19 pandemic by selecting the most recent observation of each participant between 1999 and 2019. We propose that in-person health care utilization may serve as a proxy measure for access to rheumatologic care in the current health care landscape. We evaluated potential geographic differences in access to care as this may provide insight into observed geographic disparities and highlight targets for health care delivery interventions.

PATIENTS AND METHODS

Data source. The study participants were adults enrolled in FORWARD, the National Databank for Rheumatic Diseases, a US-wide rheumatic disease longitudinal registry. Participants in FORWARD are volunteers who have patient-reported, physician-confirmed diagnoses and are recruited primarily from rheumatology practices across the United States (10). A small proportion of participants are self-referred or enroll from other sources. Participants in this study had an RA or OA diagnosis and a completed six-month semiannual questionnaire between 1999 and 2019, prior to the COVID-19 pandemic. The OA population serves as an active control because this group is the closest comparison to a "healthy" population in FORWARD. For each participant, the most recent observation was selected and the years 2020 and onward were excluded. This study was approved by the Via Christi Hospitals Wichita, Inc., Institutional Review Board (IRB00001674), and all participants consented to participate. Current participant home addresses were converted into 10 codes using US Department of Agriculture Rural-Urban Commuting Area (RUCA) codes. RUCA codes are a commonly used methodology for classifying rurality in health policy and research and are based on primary commuting patterns to urban centers and smaller population centers (11). RUCA codes were further aggregated into three commonly accepted groupings for analysis: small rural towns/isolated rural areas (completely rural areas or commuting patterns to population centers of 2,500–9,999 residents), large rural areas (areas with micropolitan cores of 10,000–49,999 residents and significant commuting patterns to urban clusters), and urban areas (metropolitan core areas of ≥50,000 residents and adjacent areas).

Study variables. The following variables were assessed using reports from participant questionnaires and used as covariables for the analysis: demographics (age, sex, race/ethnicity [a fixed set of categories including "other"], education level, and marital status), primary rheumatic disease, duration of disease, history of smoking, body mass index, patient-reported outcome measures (including pain, global severity, fatigue, Health Assessment Questionnaire II [HAQ-II] score, Patient Activity Scale II [PAS-II], and health satisfaction scores), medication use (including disease-modifying antirheumatic drugs [DMARDs] and DMARD categories) and year of study entry.

Outcome variables. Health care utilization variables from questionnaires were analyzed by geographic categories (small rural/isolated, large rural, and urban). Small rural was used as the reference group for the analyses. The outcome variables were all health professional visits, rheumatologist visits, primary care/general practitioner (GP) visits, other physician visits, other non-physician health professional visits (ie, physical therapy/occupational therapy [PT/OT], nurse, chiropractor, other health worker), other nontraditional practitioners (ie, massage therapist, acupuncturist, herbalist, acupressurist, or homeopathic practitioner), diagnostic tests, and composite measures of these variables. Composite measures for both RA and OA included the following: all physician visits; a combination of rheumatologist, GP, and PT/OT visits, and a combination of rheumatologist and GP visits. Composite measures are not mutually exclusive.

Statistical analyses. Poisson regression models were used to assess the relationship between health care utilization outcomes. Analysis of variance and chi-squared tests were used to conduct a bivariate analysis between geographic location and the remaining variables characterizing the cohorts. Double selection least absolute shrinkage and selection operator (LASSO) with Poisson regression was used for variable selection

and inference (12–14). This technique developed by Belloni et al (15) uses LASSO regression as basis, that is, minimizing the mean squared error with a penalty term times the sum of coefficients' absolute values. The penalty term will shrink the estimated coefficients toward zero, performing variable selection. However, one of the problems of the LASSO selection method is that it selects covariates and estimates coefficients but does not provide the standard errors required for performing statistical inference. These authors developed a novel estimation and uniformly valid inference method for the treatment effect called "post–double-selection" method. Briefly, this consists of obtaining a post–model-selection estimator that applies ordinary least squares to the model selected by first-step penalized estimators, typically LASSO. The main attractive feature of this method is that it allows for imperfect selection of the control variables and provides confidence intervals that are valid uniformly across a large class of models. It resolves the problem of uniform inference after model selection.

Double-selection LASSO with Poisson regression estimates incidence rate ratios, standard errors, and confidence intervals for the treatment variables of interest, in which we wish to make inferences (in this study, the geographical location), while using LASSOs to select from potential control variables. However, double selection does not provide estimates of the coefficients on the control variables or their standard errors. A double selection LASSO algorithm is performed in two stages and the procedure is the following: first, a LASSO is fitted to choose from all the variables, the ones associated with the outcome (in our case, the utilization variables); second, a LASSO is performed for the exposure of interest (geographic location) to identify the variables associated with it. The union of the covariables found in steps 1 and 2 is fitted in a final regression. The intuition of this approach is to ensure few confounders are omitted. To select the optimal penalty parameter, there are several methods such as cross-validation, adaptive LASSO, or the plug-in iterative formula. We used the latter because the results were similar with the ones obtained by cross-validation.

RESULTS

Of the 37,802 participants with RA, 74.5% lived in urban areas, 12.2% in large rural areas and 13.2% in small rural areas. For the 8,248 participants with OA, the distribution was very similar: 75.0% urban, 12.5% large rural, and 12.5% small rural. Table 1 presents characteristics for individuals with RA and OA. In the RA population, geographic differences existed for most of the variables except in disease duration, smoking history, health satisfaction, nonsteroidal anti-inflammatory drug (NSAID) use, and Janus Kinase inhibitor use. In the OA population, geographic differences also existed in most variables except for sex, age, disease duration, smoking history, NSAID use, and health satisfaction. Standardized mean differences are also presented in Supplementary Table 1, which helped us identify important differences in variables such as educational level, age, race, and marital status for both populations of people with RA and OA.

The absolute number of health care utilization visits by the six-month survey are shown in Table 2, including relevant composite measures. Results from the unadjusted Poisson regression are also presented in Table 2. Figure 1 and Supplementary Table 2 show the incidence rate ratio (IRR) from the double selection LASSO with Poisson regression models for health care utilization outcome variables. Supplementary Table 3 details the selected variables.

Urban residents with RA were more likely than their small rural counterparts to use health care in the form of all health professional visits (IRR, 1.16; 95% CI, 1.11–1.20), rheumatologist visits (IRR, 1.22; 95% CI 1.18–1.27), "other" physician visits (IRR, 1.34; 95% CI 1.26–1.43), other non-physician health professional visits (IRR, 1.12; 95% CI, 0.95–1.32), and diagnostic tests (IRR, 1.12; 95% CI, 1.07–1.18) (Figure 1 and Supplementary Table 2). Urban residents were also more likely than those in small rural areas to use health care in composite measures of all physician visits (IRR, 1.15; 95% CI, 1.10–1.19), combined rheumatologist/PT/OT/GP (IRR, 1.10; 95% CI, 1.05–1.14), and combined rheumatologist/GP visits (IRR, 1.08; 95% CI, 1.04–1.12) (Supplementary Table 2). Urban residents with RA were less likely to use GP visits (IRR, 0.89; 95% CI, 0.90–0.94) (Figure 1 and Supplementary Table 2). Individuals from large rural areas were also more likely than small rural counterparts to use health care in the form of "other" physician visits (IRR, 1.17; 95% CI, 1.08–1.27) and diagnostic tests (IRR, 1.10; 95% CI, 1.03–1.18) (Figure 1 and Supplementary Table 2). These results remained very similar when DMARD medications were not added in the models (Supplementary Table 2, column 3).

Among individuals with OA, those from urban areas were more likely than their small rural counterparts to use all health professional visits (IRR, 1.12; 95% CI, 1.02–1.24), rheumatologist visits (IRR, 1.27; 95% CI, 1.11–1.46), "other" physician visits (IRR, 1.23; 95% CI, 1.09–1.39), diagnostic tests (IRR, 1.15; 95% CI, 1.03–1.28), all physician visits (IRR, 1.13; 95% CI, 1.04–1.24), and combined rheumatologist/PT/OT/GP visits (IRR, 1.12; 95% CI, 1.01–1.23) (Figure 1 and Supplementary Table 2). Individuals with OA from large rural areas were not more likely than their small rural counterparts to use health care in any specific outcome measure (Figure 1 and Supplementary Table 2).

DISCUSSION

We conducted a large retrospective analysis of patient-reported health care utilization in over 45,000 participants with RA and OA. This study appears to be among the first to examine whether an urban-rural health care utilization disparity exists specifically in arthritis care. In evaluating urban versus rural places of residence, several health care utilization outcome measures seem

Table 1. Participant characteristics by location and disease (RA and OA) measured at last observation*

Variable	RA (n = 37,802)				OA (n = 8,248)			
	Urban	Large rural	Small rural	*P*	Urban	Large rural	Small rural	*P*
Male sex, % (n)	19.70 (5,550)	21.61 (1,000)	23.48 (1,174)	0.000	16.14 (999)	15.47 (159)	16.80 (173)	0.715
White race, % (n)	94.34 (26,579)	96.65 (4,473)	97.14 (4,857)	0.000	95.83 (5,932)	97.37 (1,001)	98.83 (1,018)	0.000
Age, mean ± SD years	62.84 ± 14.37	63.71 ± 13.92	64.46 ± 13.50	0.000	67.79 ± 12.63	67.55 ± 12.42	67.04 ± 12.04	0.203
Education, mean ± SD years	13.59 ± 2.40	13.26 ± 2.34	13.02 ± 2.26	0.000	13.75 ± 2.41	13.45 ± 2.36	13.22 ± 2.25	0.000
Married (yes/no), % (n)	67.76 (19,091)	70.72 (3,273)	72.88 (3,644)	0.000	62.86 (3,891)	69.16 (711)	70.68 (728)	0.000
Disease duration, mean ± SD years	18.07 ± 13.03	18.28 ± 12.93	18.52 ± 12.87	0.061	20.46 ± 13.92	20.32 ± 13.67	21.20 ± 13.87	0.253
Comorbidity index (0–9), mean ± SD	1.98 ± 1.66	2.06 ± 1.70	2.04 ± 1.67	0.000	2.18 ± 1.63	2.39 ± 1.67	2.27 ± 1.63	0.000
Ever smoked?, % (n)	41.60 (11,721)	42.52 (1,968)	42.74 (2,137)	0.202	37.35 (2,312)	36.19 (372)	33.69 (347)	0.072
BMI, kg/m^2, mean ± SD	28.31 ± 7.01	28.62 ± 6.96	28.93 ± 32.70	0.007	29.60 ± 7.28	30.63 ± 7.79	30.63 ± 7.98	0.000
HAQ-DI (0–3 scale), mean ± SD	1.16 ± 0.75	1.21 ± 0.76	1.21 ± 0.75	0.000	1.16 ± 0.71	1.23 ± 0.70	1.19 ± 0.73	0.020
Patient activity score (0–10 scale), mean ± SD	4.02 ± 2.28	4.18 ± 2.29	4.19 ± 2.27	0.000	4.17 ± 2.20	4.41 ± 2.11	4.30 ± 2.22	0.002
Fatigue (VAS 0–10 scale), mean ± SD	4.75 ± 3.07	4.92 ± 3.07	4.85 ± 3.01	0.001	4.81 ± 3.03	5.18 ± 2.93	5.05 ± 2.98	0.000
Pain (VAS 0–10 scale), mean ± SD	4.24 ± 2.89)	4.43 ± 2.90	4.41 ± 2.88	0.000	4.55 ± 2.83	4.82 ± 2.72	4.67 ± 2.79)	0.015
Patient global (0–10 scale), mean ± SD	3.96 ± 2.57	4.09 ± 2.58	4.12 ± 2.56	0.000	4.09 ± 2.51	4.34 ± 2.48	4.24 ± 2.53	0.005
Health satisfaction (0–5 scale)								
Somewhat satisfied, % (n)	34.05 (9,592)	32.87 (1,521)	33.60 (1,680)	0.229	33.39 (2,067)	33.37 (343)	30.68 (316)	0.760
Neither dissatisfied or satisfied, % (n)	16.91 (4,765)	17.42 (806)	17.30 (865)		17.17 (1,063)	17.90 (184)	18.06 (186)	
Somewhat dissatisfied, % (n)	23.77 (6,696)	25.17 (1,165)	24.46 (1,223)		25.36 (1,570)	25.39 (261)	26.21 (270)	
Very dissatisfied, % (n)	12.06 (3,397)	11.99 (555)	12.38 (619)		12.91 (799)	13.52 (139)	13.69 (141)	
NSAID use, % (n)	44.43 (11,812)	43.82(1,905)	43.11(2,052)	0.210	33.39 (2,067)	33.37 (343)	30.68 (316)	0.760
DMARD lifetime count, mean ± SD	2.30 ± 1.68	2.22 ± 1.65	2.33 ± 1.68	0.005	N/A	N/A	N/A	N/A
Biologic lifetime count, mean ± SD	1.10 ± 1.31	1.00 ± 1.29	0.99 ± 1.24	0.000	N/A	N/A	N/A	N/A
DMARD use, % (n)	67.72 (18,003)	68.76 (2,989)	70.59 (3,360)	0.000	N/A	N/A	N/A	N/A
Biologics use, % (n)	42.40 (11,946)	39.87 (1845)	39.42 (1,971)	0.000	N/A	N/A	N/A	N/A
TNFi use, % (n)	32.91 (9,271)	30.88 (1429)	30.36 (1,518)	0.000	N/A	N/A	N/A	N/A
Non-TNFi bDMARD use, % (n)	8.65 (2,437)	7.58 (351)	7.86 (393)	0.017	N/A	N/A	N/A	N/A
JAKi use, % (n)	1.85 (521)	1.99 (92)	2.04 (102)	0.578	N/A	N/A	N/A	N/A
Year of questionnaire, median (IQR)	2009 (2003–2015)	2010 (2004–2010)	2009 (2003–2015)	0.0002	2009 (2003–2015)	2010 (2004–2010)	2009 (2003–2015)	0.3393
1998–2004, % (n)	30.79 (8,675)	26.17 (1211)	29.62 (1,481)	0.000	30.69 (1,900)	30.54 (314)	30.78 (317)	0.490
2005–2011, % (n)	30.55 (8,607)	33.88 (1,568)	31.02 (1,551)		30.02 (1,858)	32.00 (329)	31.94 (329)	
2012–2019, % (n)	38.66 (10,892)	39.95 (1,849)	39.36 (1,968)		39.29 (2,432)	37.45 (385)	37.28 (384)	

* bDMARD = biologic disease-modifying antirheumatic drug; BMI = body mass index; HAQ-DI = Health Assessment Questionnaire Disability Index; JAKi = Janus Kinase inhibitors; N/A = not applicable; NSAID = nonsteroidal anti-inflammatory drug; OA = osteoarthritis; RA = rheumatoid arthritis; TNFi, tumor necrosis factor inhibitor; VAS = visual analogue scale.

Table 2. Health care utilization by location and disease (RA and OA)*

Utilization	Urban, mean ± SD	Large rural, mean ± SD	Small rural, mean ± SD	Unadjusted		
				IRR urban vs rural	IRR large rural vs rural	P (χ^2)
RA	2.87 ± 4.78	2.85 ± 5.00	2.54 ± 4.33			
All health professional visits	5.27 ± 7.24	4.75 ± 6.65	4.47 ± 6.22	1.18 (1.16–1.20)	1.06 (1.04–1.08)	0.000
All physician visits	3.86 ± 4.72	3.56 ± 4.48	3.34 ± 4.32	1.16 (1.14–1.18)	1.07 (1.05–1.09)	0.000
Other health professional visits	1.41 ± 3.94	1.19 ± 3.46	1.13 ± 3.27	1.24 (1.21–1.28)	1.05 (1.01–1.09)	0.0137
GP/family physician visits	2.51 ± 2.00	2.80 ± 2.05	2.90 ± 2.16	0.86 (0.84–0.89)	0.97 (0.93–1.00)	0.0689
Rheumatologist visits	2.91 ± 2.02	2.53 ± 1.80	2.54 ± 1.82	1.15 (1.12–1.18)	1.00 (0.97–1.03)	0.9435
All physicians except rheumatology and GP	1.17 ± 2.29	1.04 ± 2.16	0.86 ± 1.88	1.37 (1.32–1.41)	1.21 (1.16–1.26)	0.000
Rheumatologist and GP and PT/OT	3.24 ± 4.15	2.99 ± 3.94	2.92 ± 3.94	1.11 (1.09–1.13)	1.02 (1.00–1.05)	0.0427
Rheumatologist and GP	2.75 ± 3.28	2.59 ± 3.18	2.53 ± 3.19	1.09 (1.07–1.11)	1.02 (1.00– 1.05)	0.0902
Gastroenterology visits	0.69 ± 1.26	0.71 ± 1.36	0.65 ± 1.25	1.06 (0.97–1.16)	1.08 (0.97–1.21)	0.1613
Dentist visits	1.62 ± 1.26	1.56 ± 1.21	1.56 ± 1.17	1.04 (0.98–1.11)	1.00 (0.92–1.09)	0.9633
Other doctor visits	2.20 ± 2.26	2.16 ± 2.26	2.03 ± 2.06	1.09 (1.04–1.13)	1.07 (1.01–1.12)	0.0169
PT/OT visits	1.99 ± 3.24	1.96 ± 3.20	2.18 ± 3.33	0.91 (0.87–0.95)	0.90 (0.84–0.95)	0.0007
Chiropractor visits	1.18 ± 2.55	1.56 ± 2.76	1.74 ± 2.80	0.68 (0.64–0.71)	0.90 (0.84–0.96)	0.0027
Other health worker visits	1.08 ± 2.38	0.96 ± 2.25	1.02 ± 2.32	1.06 (0.98–1.15)	0.94 (0.85–1.05)	0.2533
Nurse visits	0.87 ± 1.97	1.00 ± 2.21	1.03 ± 2.06	0.85 (0.75–0.96)	0.97 (0.83–1.14)	0.7295
Nontraditional therapies	0.94 ± 2.99	0.60 ± 2.21	0.60 ± 2.28	1.57 (1.48–1.67)	1.00 (0.92–1.09)	0.9721
Diagnostic tests	2.87 ± 4.78	2.85 ± 5.00	2.54 ± 4.33	1.13 (1.11–1.15)	1.12 (1.10–1.15)	0.000
OA	2.87 ± 4.78	2.85 ± 5.00	2.54 ± 4.33			
All health professional visits	5.24 ± 7.75	4.71 ± 7.16	4.62 ± 6.83	1.13 (1.10–1.17)	1.02 (0.98–1.06)	0.3540
All physician visits	3.18 ± 4.42	2.88 ± 4.21	2.84 ± 4.03	1.12 (1.07–1.16)	1.01 (0.96–1.07)	0.6003
Other health professional visits	2.07 ± 4.84	1.83 ± 4.34	1.78 ± 4.15	1.16 (1.11–1.22)	1.03 (0.96–1.10)	0.4061
GP/family physician visits	2.82 ± 2.15	3.14 ± 2.28	3.15 ± 2.25	0.90 (0.84–0.95)	1.00 (0.92–1.08)	0.9445
Rheumatologist visits	1.65 ± 1.69	1.61 ± 1.51	1.59 ± 1.50	1.04 (0.96–1.14)	1.01 (0.90–1.14)	0.8190
All physicians except rheumatology and GP	1.35 ± 2.43	1.10 ± 2.20	1.07 ± 2.06	1.26 (1.19–1.35)	1.03 (0.95–1.12)	0.4504
Rheumatologist and GP and PT/OT	2.65 ± 4.01	2.48 ± 3.90	2.41 ± 3.83	1.10 (1.05–1.15)	1.03 (0.97–1.09)	0.3551
Rheumatologist and GP	1.88 ± 2.73	1.84 ± 2.74	1.81 ± 2.72	1.04 (0.99–1.09)	1.02 (0.95–1.08)	0.6112
Gastroenterology visits	0.80 ± 1.30	0.89 ± 1.37	0.84 ± 1.53	0.95 (0.81–1.12)	1.07 (0.86–1.32)	0.5494
Dentist visits	1.75 ± 1.37	1.75 ± 1.16	1.59 ± 1.16	1.10 (0.96–1.26)	1.10 (0.92–1.31)	0.2846
Other doctor visits	2.47 ± 2.27	2.40 ± 2.22	2.42 ± 2.20	1.02 (0.94–1.11)	0.99 (0.89–1.10)	0.8676
PT/OT visits	3.08 ± 3.71	3.16 ± 3.68	2.89 ± 3.59	1.06 (0.98–1.16)	1.09 (0.98–1.22)	0.1100
Chiropractor visits	1.91 ± 3.13	2.40 ± 3.23	2.53 ± 3.08	0.75 (0.69–0.83)	0.95 (0.84–1.07)	0.3733
Other health worker visits	0.79 ± 1.73	1.10 ± 2.02	0.65 ± 1.45	1.31 (1.13–1.53)	1.16 (0.95–1.42)	0.1382
Nurse visits	1.58 ± 2.87	1.40 ± 2.54	1.21 ± 2.42	1.22 (0.86–1.73)	1.69 (1.12–2.57)	0.0128
Nontraditional therapies	1.29 ± 3.62	1.14 ± 3.32	0.98 ± 3.06	1.31 (1.18–1.46)	1.16 (1.02–1.33)	0.0292
Diagnostic tests	2.69 ± 4.61	2.60 ± 4.31	2.37 ± 4.13	1.13 (1.09–1.18)	1.10 (1.04–1.16)	0.0010

* GP = primary care/general practitioner; IRR = incidence rate ratio; OA = osteoarthritis; OT = occupational therapy; PT = physical therapy; RA = rheumatoid arthritis.

to be more prevalent in urban populations, especially those involving RA care.

People with RA residing in urban areas were more likely than their small rural area peers to use health care with all health professional visits, rheumatologist visits, other physician visits, other nonphysician health services, and diagnostic tests. Even people with RA residing in large rural areas were more likely than those in small rural areas to use other physician visits and diagnostic tests. Given the association between care from a rheumatologist and improved quality of care, the increased utilization of rheumatologist visits in the urban population likely indicates a disparity in both access and quality of care for people with RA residing in rural areas (5–7). Interestingly, people residing in small rural areas were more likely than urban counterparts to use primary care in this study. This may indicate that primary care providers are filling in care gaps in underserved rural areas with lack of access to rheumatology specialist care. More diagnostic tests and other non-physician health services in urban populations could potentially indicate better adherence to quality health care practices in RA such as frequent lab monitoring for DMARD medication safety and vigilant cancer screenings in patients with RA at an increased risk for malignancy. Alternatively, more diagnostic tests could also potentially indicate overuse of low value services and tests.

In people with OA in this study, there were fewer areas of health care utilization disparity than in RA, yet urban populations used more "all" professional visits, rheumatologist visits, "other" physician visits, diagnostic tests, "all" physician visits and combined rheumatologist/GP/PT/OT visits. This difference between RA and OA health care utilization outcome measures likely reflects the relative complexity of RA care in comparison to OA. There are not yet any disease-modifying therapies for OA, and the immunosuppressive medications used in RA care often require frequent

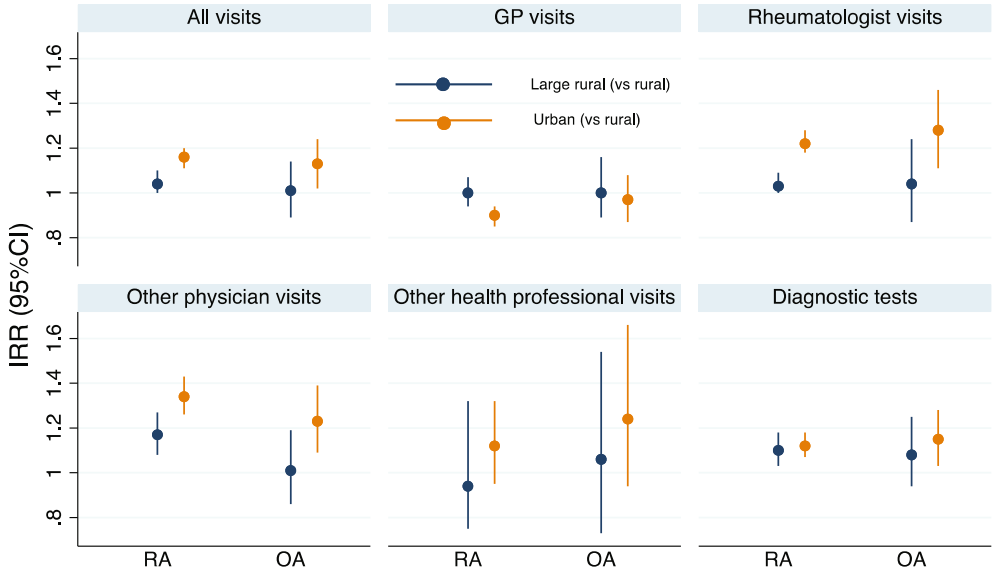


Figure 1. Incidence rate ratio (IRR) from the double selection least absolute shrinkage and selection operator (LASSO) with Poisson regression models for health care utilization outcome variables in six month questionnaires. Bars show the incidence rate ratio (IRR) ± 95% confidence interval (95% CI). GP = primary care/general practitioner; OA = osteoarthritis; RA = rheumatoid arthritis.

monitoring. More rheumatologist visits among patients with OA may reflect how the FORWARD cohort is primarily recruited from rheumatology practices.

This study has several limitations. First, this was a retrospective analysis, and potential for error exists in the accuracy of patient-reported measures. Some questionnaire items also do not capture more specific or granular data on measures of health care utilization such as the types of "other" health care providers that are used. This study relied upon disease-specific patient-reported data and could not characterize patients who are not yet diagnosed with arthritis. This presents another potential area of health care access–related disparity among undiagnosed patients that could not be captured in this study. The population in the study was also predominantly White and highly educated, which could limit the generalizability of the results. Because of the study population characteristics, our findings may underestimate actual disparities in utilization because individuals with higher educational attainment may be more likely to seek out specialty care. Future work within this and other registries should be done to enroll and retain diverse participants to examine health disparities in groups that better represent the population at-large. Lastly, there is a potential that participants could have moved from one geographic area to another during the course of the study, affecting their access to certain health care services. Theoretically, this is partially accounted for by updating participant information by surveys every six months.

Overall, disparities exist in a number of RA health care utilization measures. People living in urban areas had considerably more medical visits and diagnostic tests in comparison with those who lived in small rural areas. This may reflect a tendency toward less rheumatology-specific care in small rural area populations. Although the trend toward more primary care utilization in rural populations in encouraging, it is unclear whether people in rural areas are receiving care for their RA from primary care providers or whether some are less likely to receive specific care for RA altogether.

This study highlights the need to implement health care delivery techniques and incentives to broaden access to rheumatologic care in rural communities. Additional research is needed to examine the specific patient and system-related factors that impair access to rheumatologic care in rural communities. Future work is also needed to determine if differences between urban and rural populations in health care utilization correlate with clinically important disease-specific outcomes.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Desilet, Pedro, Michaud.
Acquisition of data. Pedro, Katz, Michaud.
Analysis and interpretation of data. Desilet, Pedro, Katz, Michaud.

REFERENCES

1. Garcia MC, Faul M, Massetti G, et al. Reducing potentially excess deaths from the five leading causes of death in the rural United States. MMWR Surveill Summ 2017;66:1–7.
2. Jordan JM, Linder GF, Renner JB, et al. The impact of arthritis in rural populations [review]. Arthritis Care Res 1995;8:242–50.

3. Zhong L, Chung KC, Baser O, et al. Variation in rheumatoid hand and wrist surgery among Medicare beneficiaries: a population-based cohort study. J Rheumatol 2015;42:429–36.
4. Groessl EJ, Ganiats TG, Sarkin AJ. Sociodemographic differences in quality of life in rheumatoid arthritis [review]. Pharmacoeconomics 2006;24:109–21.
5. MacLean CH, Louie R, Leake B, et al. Quality of care for patients with rheumatoid arthritis. JAMA 2000;284:984–92.
6. Lacaille D, Anis AH, Guh DP, et al. Gaps in care for rheumatoid arthritis: a population study. Arthritis Rheum 2005;53:241–8.
7. Ferucci ED, Donnithorne KJ, Koller KR, et al. Performance on rheumatoid arthritis quality indicators in an Alaska Native healthcare system. Qual Saf Health Care 2010;19:387–91.
8. Battafarano DF, Ditmyer M, Bolster MB, et al. 2015 American College of Rheumatology workforce study: supply and demand projections of adult rheumatology workforce, 2015–2030. Arthritis Care Res (Hoboken) 2018;70:617–26.
9. George MD, Baker JF, Banerjee S, et al. Social distancing, health care disruptions, telemedicine use, and treatment interruption during the COVID-19 pandemic in patients with or without autoimmune rheumatic disease. ACR Open Rheumatol 2021;3:381–9.
10. Wolfe F, Michaud K. The National Data Bank for rheumatic diseases: a multi-registry rheumatic disease data bank. Rheumatology (Oxford) 2011;50:16–24.
11. Hart LG, Larson EH, Lishner DM. Rural definitions for health policy and research. Am J Public Health 2005;95:1149–55.
12. Chen H, Xiao G, Xing J, et al. Empirical study on the impact of urban and rural resident basic medical insurance on children's educational outcomes. Front Public Health 2021;9:764364.
13. Torrats-Espinosa G. Using machine learning to estimate the effect of racial segregation on COVID-19 mortality in the United States. Proc Nat Acad Sci U S A 2021;118:e2015577118.
14. Everding J, Marcus J. The effect of unemployment on the smoking behavior of couples. Health Econ 2020;29:154–70.
15. Belloni A, Chernozhukov V, Wei Y. Post-selection inference for generalized linear models with many controls. J Bus Econ Stat 2016;34:606–19.

Arthritis Care & Research
Vol. 77, No. 3, March 2025, pp 419
DOI 10.1002/acr.25496

American College of Rheumatology
Empowering Rheumatology Professionals

Correction to Benefits of Early Versus Late Initiation of Intravenous Immunoglobulin in the Treatment of Patients With Anti–3-Hydroxy-3-Methylglutaryl-Coenzyme A Reductase Immune-Mediated Necrotizing Myopathy

Sharf K, Do T, Ghetie D, Choi D, Chahin N. Benefits of Early Versus Late Initiation of Intravenous Immunoglobulin in the Treatment of Patients With Anti-3-Hydroxy-3-Methylglutaryl-Coenzyme A Reductase Immune-Mediated Necrotizing Myopathy. Arthritis Care Res (Hoboken). 2024 Nov;76(11):1584-1592. doi: 10.1002/acr.25406.

In Table 2, all values for serum CK should be reported in units per liter (U/L), not L/L as mistakenly published. See below for the corrected Table 2.

Table 2. Outcome measures among the patient cohort with anti-HMGCR myopathy (n = 31) and comparison between patients who received delayed and nondelayed IVIG*

Characteristic	Cohort with nondelayed treatment (n = 19)	Cohort with delayed treatment (n = 12)	*P* value[a]	Bonferroni corrected *P* value[a]
Time difference between initiation of IVIG and three-month follow-up visit, average (±SD), d	107.42 (±89.19)	106.58 (±61.18)	0.97	0.97
Serum CK levels at zero months, average (±SD), U/L	5,800 (±2,994.16)	2,882.91 (±1,564.57)	**0.004**	**0.016**
Serum CK levels at three months, average (±SD), U/L	1915 (±2,744.24)	894 (±882.32)	0.22	0.88
Serum CK levels at six months, average (±SD), U/L	544 (±947.31)	611 (±749.88)	0.84	>0.999
Serum CK levels at 12 months, average (±SD), U/L	480 (±781.15)	384 (±454.33)	0.71	>0.999
MMT8 score at zero months, average (±SD)[b]	117 (±12)	127 (±15)	**0.03**	0.12
MMT8 score at three months, average (±SD)	141 (±10)	134 (±17)	0.13	0.52
MMT8 score at six months, average (±SD)	146 (±5)	138 (±13)	**0.02**	0.08
MMT8 score at 12 months, average (±SD)	149 (±3)	132 (±18)	**<0.001**	**<0.001**
Prednisone dosage at three months, average (±SD), mg/d	18 (±19.79)	23 (±18.31)	0.55	>0.999
Prednisone dosage at six months, average (±SD), mg/d	7 (±10.51)	13 (±13.10)	0.13	0.39
Prednisone dosage at 12 months, average (±SD), mg/d	3 (±5.08)	8 (±7.18)	**0.03**	0.09
Requiring a walker or wheelchair for ambulation at zero months, n (%)[b]	9 (47)	8 (66)	0.46	>0.999
Requiring a walker or wheelchair for ambulation at six-month evaluation, n (%)	1 (∼5)	7 (58)	**0.002**	**0.006**
Requiring a walker or wheelchair for ambulation at 12-month evaluation, n (%)	0 (0)	7 (58)	**0.0003**	**0.0009**

* The bolded values represent those of statistical significance, a *P*-value of less than 0.05. CK, creatine kinase; HMGCR, 3-hydroxy-3-methylglutaryl-coenzyme A reductase; IVIG, intravenous Ig; MMT, manual muscle testing. Patients receiving delayed IVIG are defined as those receiving it more than six months after symptom onset, and patients with nondelayed IVIG are defined as those receiving it six or fewer months after symptom onset.
[a] The *P* value of an unpaired two-tailed Student's *t*-test compares the cohorts with nondelayed and delayed treatment. Thin lines indicate *P* value subgroups used in Bonferroni correction.
[b] Zero months is defined as the clinical time point immediately before IVIG administration.

Thank you for correcting this,

Kyle Sharf, DO

Nizar Chahin, MD