Arthritis Care & Research

Aims and Scope

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

Volume 74 No. 3 March 2022

Arthritis Care & Research

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

Cover design: Sandra Pulmano

♾ This journal is printed on acid-free paper.

Arthritis Care & Research

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

VOLUME 74 • March 2022 • NO. 3

Cover image: The image on the cover (from Correll et al, page 340) shows the pediatric rheumatology distribution rate per 100,000 children, 2015 versus 2030, including Alaska and Hawaii.

Arthritis Care & Research
Vol. 74, No. 3, March 2022, pp 337–339
DOI 10.1002/acr.24810

American College of Rheumatology
Empowering Rheumatology Professionals

EDITORIAL

"Every New Beginning Comes From Some Other Beginning's End": Anterior Cruciate Ligament Injury and Posttraumatic Knee Osteoarthritis

Jackie L. Whittaker

Seneca's quote, "Every new beginning comes from some other beginning's end" (which was made popular by the Semisonic's 1998 song "Closing Time") reminds us that at a philosophical level there are neither beginnings nor ends, just an uninterrupted flow of events. Although we all know this to be true, at some level, as human beings we prefer to compartmentalize events and experiences by assigning them a start and an end. We struggle when we can't define something or put boundaries around it. This struggle is very real for Western health care practitioners and researchers in clinical fields who need health conditions to have a standardized start and end point in order to guide treatment and enable research to prevent or cure the conditions. Moreover, how we define when a condition starts and ends or makes way for a second condition depends on our point of view. A person living with a health condition might interpret the start and end of their condition much differently than a clinician faced with providing a diagnosis or a researcher deciding how to classify people for a clinical trial to assess the effectiveness of a new treatment. This challenge of knowing when one condition ends and a new condition begins is very relevant to the context of anterior cruciate ligament (ACL) tears and subsequent development of posttraumatic osteoarthritis (PTOA).

ACL tears represent a significant health burden in active adolescents and young adults and are so common (particularly in women's field and court sports) that they are basically treated as a rite of passage. Contrary to popular belief, ACL tears are not self-limiting. Recovery is not as simple as having an ACL reconstruction (ACLR) and/or committing to an extended period (i.e., 9–12 months) of intensive rehabilitation, and then simply returning to preinjury life. It is very well established that ACL tears (and other traumatic knee injuries) are associated with elevated symptoms and reduced function for many years (1) and increase susceptibility to additional injuries (2), inactivity (3), obesity (4) and PTOA (5). Recent estimates suggest that the risk of PTOA after an ACL tear is elevated 4- to 6-fold depending on whether it is an isolated or combined injury (6). Given that a significant proportion of people who tear their ACL develop PTOA, it raises the question: When do knee symptoms and the consequences of these symptoms stop being attributed to the ACL tear or ACLR and start being attributed to PTOA?

Our perspective of when PTOA starts after an ACL tear or ACLR depends upon our definition and our purpose. If the goal is to identify people in whom a specific intervention would be appropriate, we might use a different definition based on tools readily available in a clinical setting than if our goal is to classify people for a research study or to understand research findings. Traditionally, the diagnosis and classification of OA, both clinically and in research, has been based on indicators of disease (i.e., radiographic evidence of articular cartilage pathology). Recognition that cartilage pathology can, but does not always, manifest as illness (i.e., pain, disability, reduced quality of life, health-seeking behaviors) has given rise to combined definitions, such as symptomatic radiographic OA.

The study conducted by Harkey and colleagues, which is published in this issue of *Arthritis Care & Research* (7), considered 2 definitions of early OA illness after ACLR that are based on different Knee Injury and OA Outcome Score (KOOS) subscale (i.e., knee-related pain, symptoms, function in daily living, function in sport and recreation, and quality of life) thresholds. The first is a modification of an expert consensus definition of early OA (excluding joint line tenderness, crepitus, and Kellgren/Lawrence grade criteria) (8), and the second is a bespoke definition of a

Dr. Whittaker's work was supported by an Arthritis Society STARS Career Development Award and a Michael Smith Foundation for Health Research Scholar Award.

Jackie L. Whittaker PT, PhD: University of British Columbia and Arthritis Research Canada, Vancouver, British Columbia, Canada.

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.24810&file=acr24810-sup-0001-Disclosureform.pdf.

Address correspondence to Jackie Whittaker, PT, PhD, Department of Physical Therapy, University of British Columbia, Vancouver, British Columbia, Canada. Email: jackie.whittaker@ubc.ca.

Submitted for publication October 12, 2021; accepted in revised form October 14, 2021.

"symptomatic knee" or "individuals symptomatic enough to possibly seek medical care" proposed in a case–control study of persons who had undergone meniscectomy ~16 years previously (9). The authors also proposed modifications to these 2 definitions in which the KOOS subscale thresholds reflected 1 of the 6 published thresholds for the KOOS patient acceptable symptom state (PASS) for persons with an ACL tear (10–16). Using these 4 definitions, the authors identified that 5–8 months following an ACLR, up to 54% of young people meet 1 or more of author definitions for OA illness. Continued follow-up of this small cohort to determine the incidence of structural OA, symptomatic radiographic OA, or other indicators of OA illness (i.e., arthroplasty) is needed to determine the validity of the proposed definitions.

In addition to describing outcomes associated with OA illness after ACLR, this study highlights important considerations for ongoing efforts to define early-stage OA (17). First, it reaffirms that OA is a heterogeneous condition and that no single set of criteria may be applicable to all situations. Going forward, it is essential that criteria for early OA are sensitive to all phenotypes of OA, including PTOA (which typically has a relatively early age of onset in otherwise active persons), as well as consider definitions of both early OA disease (structural OA) and early OA illness (clinical or symptomatic OA). Ideally these definitions will be validated using samples consisting of persons with and without a previous knee injury (18) and with and without OA, applicable across race, ethnicity, gender, socioeconomic status, and capable of being applied in low-resource settings.

Second, the development of criteria for early OA should be data driven with a focus on statistical estimates that are of clinical and patient relevance, such as the PASS. Where data are not available, criteria should be driven by a rigorous consensus approach. When selecting the PASS (or similar) thresholds, study quality, contextual factors (i.e., study sample including age, previous knee trauma or surgery, follow-up time point and setting), and credibility (19) should be carefully considered. For example, the KOOS PASS thresholds are highly susceptible to methodologic heterogeneity and can differ by up to 26 points for the same subscale (10,15,20).

Third, much of our understanding of the trajectory of self-reported symptoms, function, and knee-related quality of life after ACL tear and ACLR are based on the KOOS subscale group mean values that are reported at set intervals of time (e.g., 3, 6, 9, 12, and 24 months) and do not take into consideration the considerable individual variability or daily fluctuation reported in clinical encounters. There is a need for high-quality studies that are specifically designed to characterize the KOOS trajectory and account for influence of baseline status, treatment approach (i.e., rehabilitation, ACLR), and other important factors to understand if there is a clear delineation of these constructs between ACLR and PTOA.

Perhaps the most important philosophical consideration that the study by Harkey and colleagues demonstrates is the utility of seeing PTOA as a condition that starts at some point after the ACL tear experience, as opposed to at the time of ACL tear. Similar to definitions of PTOA, health care providers and researchers have defined the end of the ACL tear experience as some combination of time from ACLR and/or symptomatic or functional-based criterion suggesting a person is ready to return to unrestricted physical activity, sport, or occupation to facilitate clinical management and research methodology. However, persons with lived experience of an ACL tear and subsequent PTOA view things very differently. A predominate viewpoint among this group is that the injury was the start of an ongoing journey with their knee (21,22) consisting of fluctuations in knee health that have varying and increasingly greater degrees of impact on their physical abilities and quality of life on an ongoing basis. This perspective suggests that, in our desire to differentiate the experience of ACL tear and PTOA to facilitate clinical management and research, we may be artificially creating "a new beginning from some other beginning's end." In doing so, we may be blinding ourselves to the lived experience of an individual with an ACL tear and PTOA, not to mention the importance and potential opportunities available to promote life-long knee health by considering ACL injury and PTOA as a continuum or uninterrupted flow of events.

AUTHOR CONTRIBUTIONS

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

REFERENCES

1. Whittaker JL, Toomey CM, Nettel-Aguirre A, Jaremko JL, Doyle-Baker PK, Woodhouse LJ, et al. Health-related outcomes after a youth sport-related knee injury. Med Sci Sports Exerc 2019;51:255–63.
2. Ardern CL, Webster KE, Taylor NF, Feller JA. Return to sport following anterior cruciate ligament reconstruction surgery: a systematic review and meta-analysis of the state of play. Br J Sports Med 2011;45:596–606.
3. Ardern CL, Taylor NF, Feller JA, Webster KE. Fifty-five percent return to competitive sport following anterior cruciate ligament reconstruction surgery: an updated systematic review and meta-analysis including aspects of physical functioning and contextual factors. Br J Sports Med 2014;48:1543–52.
4. Toomey CM, Whittaker JL, Nettel-Aguirre A, Reimer RA, Woodhouse LJ, Ghali B, et al. Higher fat mass is associated with a history of knee injury in youth sport. J Orthop Sports Phys Ther 2017;47:80–7.
5. Snoeker B, Turkiewicz A, Magnusson K, Frobell R, Yu D, Peat G, et al. Risk of knee osteoarthritis after different types of knee injuries in young adults: a population-based cohort study. Br J Sports Med 2020;54:725–30.
6. Poulsen E, Goncalves GH, Bricca A, Roos EM, Thorlund JB, Juhl CB. Knee osteoarthritis risk is increased 4-6 fold after knee injury– a systematic review and meta-analysis. Br J Sports Med 2019;53:1454–63.
7. Harkey MS, Baez S, Lewis J, Grindstaff TL, Hart J, Driban JB, et al. Prevalence of early knee osteoarthritis illness among various

patient-reported classification criteria after anterior cruciate ligament reconstruction. Arthritis Care Res (Hoboken) 2022;74:377–85.

8. Luyten FP, Bierma-Zeinstra S, Dell'Accio F, Kraus VB, Nakata K, Sekiya I, et al. Toward classification criteria for early osteoarthritis of the knee. Semin Arthritis Rheum 2018;47:457–63.
9. Englund M, Roos EM, Lohmander LS. Impact of type of meniscal tear on radiographic and symptomatic knee osteoarthritis: a sixteen-year followup of meniscectomy with matched controls. Arthritis Rheum 2003;48:2178–87.
10. Muller B, Yabroudi MA, Lynch A, Chung-Liang L, van Dijk CN, Fu FH, et al. Defining thresholds for the patient acceptable symptom state for the IKDC subjective knee form and KOOS for patients who underwent ACL reconstruction. Am J Sports Med 2016;44:2820–6.
11. Liu JN, Gowd AK, Redondo ML, Christian DR, Cabarcas BC, Yanke AB, et al. Establishing clinically significant outcomes after meniscal allograft transplantation. Orthop J Sports Med 2019;7: 2325967118818462.
12. Ingelsrud LH, Granan LP, Terwee CB, Engebretsen L, Roos EM. Proportion of patients reporting acceptable symptoms or treatment failure and their associated KOOS values at 6 to 24 months after anterior cruciate ligament reconstruction: a study from the Norwegian Knee Ligament Registry. Am J Sports Med 2015;43:1902–7.
13. Vega JF, Jacobs CA, Strnad GJ, Farrow L, Jones MH, Miniaci A, et al. Prospective evaluation of the patient acceptable symptom state to identify clinically successful anterior cruciate ligament reconstruction. Am J Sports Med 2019;47:1159–67.
14. Pedersen JR, Roos EM, Thorlund JB, Terluin B, Ingelsrud LH. Cutoff values to interpret short-term treatment outcomes after arthroscopic meniscal surgery measured with the Knee Injury and Osteoarthritis Outcome Score. J Orthop Sports Phys Ther 2021;51:281–8.
15. Beletsky A, Naami E, Lu Y, Polce EM, Nwachukwu BU, Okoroha KR, et al. The patient acceptable symptomatic state in primary anterior cruciate ligament reconstruction: predictors of achievement. Arthroscopy 2021;37:600–5.
16. Maheshwer B, Wong SE, Polce EM, Paul K, Forsythe B, Bush-Joseph C, et al. Establishing the minimal clinically important difference and patient-acceptable symptomatic state after arthroscopic meniscal repair and associated variables for achievement. Arthroscopy 2021;37:3749–86.
17. Mahmoudian A, Lohmander LS, Mobasheri A, Englund M, Luyten FP. Early-stage symptomatic osteoarthritis of the knee- time for action. Nat Rev Rheumatol 2021;17:621–32.
18. Whittaker JL, Jaremko JL, Emery CA. The applicability of criteria for early osteoarthritis in 'at-risk' post-traumatic knee injured and uninjured youth. Osteoarthritis Cartilage 2019;27:S261–62.
19. Devji T, Carrasco-Labra A, Qasim A, Phillips M, Johnston BC, Devasenapathy N, et al. Evaluating the credibility of anchor based estimates of minimal important differences for patient reported outcomes: instrument development and reliability study. BMJ 2020;369.
20. Macri EM, Young JJ, Ingelsrud LH, Khan KM, et al. Meaningful thresholds for patient-reported outcomes following interventions for anterior cruciate ligament tear or traumatic meniscus injury: an OPTIKNEE systematic review. Br J Sports Med. Under review.
21. MacKay C, Jaglal SB, Sale J, Badley EM, Davis AM. A qualitative study of the consequences of knee symptoms: 'It's like you're an athlete and you go to a couch potato.' BMJ Open 2014;4:e006006.
22. Truong LK, Mosewich AD, Miciak M, et al. Balance, reframe, and overcome: the attitudes, priorities, and perceptions of exercise-based activities in youth 12-24 months after a sport-related ACL injury. J Orthop Res doi: 10.1002/jor.25064 2021. E-pub ahead of print.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 3, March 2022, pp 340–348
DOI 10.1002/acr.24497

2015 American College of Rheumatology Workforce Study and Demand Projections of Pediatric Rheumatology Workforce, 2015–2030

Colleen K. Correll,[1] Marcia M. Ditmyer,[2] Jay Mehta,[3] Lisa F. Imundo,[4] Marisa S. Klein-Gitelman,[5] Seetha U. Monrad,[6] and Daniel F. Battafarano[7]

Objective. To describe the character and composition of the 2015 pediatric rheumatology workforce in the US, evaluate current workforce trends, and project future supply and demand of the pediatric rheumatology workforce through 2030.

Methods. The American College of Rheumatology created the workforce study group to study the rheumatology workforce. The workforce study group used primary and secondary data to create a representative workforce model. Pediatric rheumatology supply and demand was projected through 2030 using an integrated data-driven framework to capture a more realistic clinical full-time equivalent (FTE) and produce a better picture of access to care issues in pediatric rheumatology.

Results. The 2015 pediatric rheumatology workforce was estimated at 287 FTEs (300 providers), while the estimated excess demand was 95 (33%). The projected demand will continue to increase to almost 100% (n = 230) by 2030 if no changes occur in succession planning, new graduate entrants into the profession, and other factors associated with the workforce.

Conclusion. This study projects that the pediatric rheumatology workforce gap will continue to worsen significantly from the 2015 baseline, and by 2030 the demand for pediatric rheumatologists will be twice the supply. Innovative strategies are needed to increase the workforce supply and to improve access to care.

INTRODUCTION

The relative lack of pediatric rheumatologists to treat the approximately 300,000 children in the US with chronic arthritis and other rheumatic diseases has been a recognized problem for decades (1). In the late 1990s Cassidy and Athreya reported that the number of practicing pediatric rheumatologists had grown from 27 in 1976 to 178 (121 board-certified) in 1996 (2). Although the 7-fold growth of the specialty over those 20 years seemed promising, it remained concerning that more than one-third of 125 pediatric academic centers did not have a pediatric rheumatologist faculty member. In 2006, the American Board of Pediatrics (ABP) published data showing that there were 200 board-certified pediatric rheumatologists, with a clear increasing trend in the number of pediatric rheumatology fellows over 10 years (3). However, the same study demonstrated that there were only 3 pediatric rheumatologists per million children in the US, and 14 states had no practicing pediatric rheumatologists. The American College of Rheumatology (ACR) workforce study published 1 year later predicted a pediatric rheumatology deficit of 33 providers by 2025 (4). In response to these findings, a series of policy recommendations, focused on training and economics, health care delivery, and global outreach, were published to aid in increasing the pediatric rheumatology workforce (5–7). Despite these studies and policy recommendations, a significant deficit in the pediatric rheumatology workforce remains.

To understand the full extent of this workforce gap, in 2015 the ACR created the workforce study group. The purpose of the workforce study group was to evaluate the changes in the adult and pediatric workforce through 2030 and to provide potential

[1]Colleen K. Correll, MD, MPH: University of Minnesota, Minneapolis; [2]Marcia M. Ditmyer, PhD, MBA, MS: University of Nevada, School of Dental Medicine, Las Vegas; [3]Jay Mehta, MD: Children's Hospital of Philadelphia, Philadelphia, Pennsylvania; [4]Lisa F. Imundo, MD: Columbia University, New York, New York; [5]Marisa S. Klein-Gitelman, MD: Ann and Robert H. Lurie Children's Hospital of Chicago, Chicago, Illinois; [6]Seetha U. Monrad, MD: University of Michigan, Ann Arbor; [7]Daniel F. Battafarano, DO: San Antonio Military Medical Center, San Antonio, Texas.

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

Address correspondence to Colleen K. Correll, MD, MPH, Pediatric Rheumatology, University of Minnesota, 2450 Riverside Avenue South, East Building Room 668, Minneapolis, MN 55410. Email: corr0250@umn.edu.

Submitted for publication May 20, 2020; accepted in revised form October 22, 2020.

SIGNIFICANCE & INNOVATIONS

- A shortage currently exists of the US pediatric rheumatology workforce to treat children with rheumatic diseases.
- Some geographic regions in the US, especially the South and Southwest, have a severe shortage in pediatric rheumatology providers, and this gap is expected to worsen if interventions do not occur.
- The overall shortage of the workforce is predicted to worsen so that by 2030 demand for pediatric rheumatology providers will be twice the supply.
- Strategies are needed to recruit rheumatologists, physician assistants, and nurse practitioners to pediatric rheumatology and to augment the provider network.

solutions to be addressed by the ACR and other stakeholders. The goals of the pediatric arm of the workforce study group were to describe the current state of the pediatric rheumatology workforce as compared to the previous ACR workforce study (4), project a succession plan as rheumatologists near retirement, develop assumptions regarding the key factors affecting the supply of and demand for rheumatologists, create a patient-centered approach to providing quality care to all patients with rheumatic conditions, and conduct a sensitivity analysis of this workforce model to determine the potential best- and worst-case scenarios. Results from the 2015 adult rheumatology arm of this study have been published previously (8). Here we present the pediatric rheumatology workforce study findings. From these findings, we propose solutions to improve the supply of pediatric rheumatology providers.

MATERIALS AND METHODS

Workforce study group. The workforce study group was composed of a diverse membership group of volunteer rheumatology specialists, including pediatric rheumatologists. There were 5 members of the core leadership, 3 of whom are co-authors on this study (MMD, SUM, and DFB). Two of the core leaders (SUM and DFB) were adult rheumatologists and the other 3 had expertise in workforce and academic leadership. An additional 9 members belonged to the core group. The core membership group included 2 pediatric rheumatologists (co-authors LFI and MSK-G), 1 fellow in adult and pediatric rheumatology, 1 physician assistant (PA), and 1 nurse practitioner (NP). Among both groups, there were 4 division directors (2 adult, 2 pediatric) and 2 adult program directors. Group members came from a variety of geographic locations in the US. Full details of the workforce study group can be found in Appendix A of the 2015 workforce study document (9). Additional focus groups were used to ensure that members of the pediatric rheumatology workforce not represented in the workforce study group were able to provide their perspectives. The workforce study group provided input into the secondary data collection procedures, provided guidance in the primary data collection methods of ACR/Association of Rheumatology Professionals (ARP) members, identified critical factors affecting supply and demand for rheumatology services, approved the workforce study modeling process, and accepted the final workforce study findings. The University of Michigan Institutional Review Board reviewed the study and determined it to be exempt from ongoing review (exemption #2 of the 45 CFR 46.101.[b]; HUM00104523).

Data collection. A mixed-methods approach (both primary and secondary data) was used to identify and evaluate workforce issues. These issues informed the model used to help predict the future pediatric rheumatology workforce. Data were collected from many secondary sources (e.g., American Medical Association, ABP, Rheumatology Nurses Society, National Commission Certification of Physician Assistants) (9). Primary data were collected through electronic surveys distributed to the ACR membership, current rheumatology fellows-in-training, and a group of rheumatology patients identified by the Arthritis Foundation (9). These data were supplemented by data collected through pediatric focus groups and personal interviews. Volunteers were recruited through the ACR to participate in focus groups, both in-person at the ACR Annual Meeting, and via teleconference, for a total of 8 focus groups that included 5–10 participants in each group. Information from these interviews was integrated into the workforce study.

Workforce study modeling. The workforce study model was a critical focus of the workforce study group. The challenge was developing a model that would ensure translating population needs into the appropriate provider supply. The workforce study group selected an integrated workforce framework model that combined socioeconomic and epidemiologic factors along with utilization rates that incorporated the current use of health care services. The first step was to determine the number of pediatric rheumatology providers in the workforce. This step was done by reviewing the number of providers that were ABP board-certified and was supplemented by reviewing pediatric providers in the ACR website and by reviewing responses to the workforce study survey. Pediatric providers included physicians, NPs, and PAs. The next step was to define the pediatric rheumatology workforce that provided direct patient care at the time of the study (2015), defined as the clinical full-time equivalent (FTE). Because of the changing demographics and pattern trends identified, understanding the actual number of practitioners was clearly not sufficient to determine the workforce supply.

The clinical FTE, which is the ratio of units that equate to the number of practitioners seeing patients full-time, was subsequently identified, and used to provide a realistic level of effort devoted to direct patient care. For example, a clinical FTE of 0.5

(or 50%) means that a provider spends half of their time in patient care. Therefore, 2 providers with 0.5 FTE would equate to 1 clinical FTE. After careful assessment and consensus discussion among pediatric rheumatologists in the workforce, the clinical FTE definition for pediatric rheumatology used in the workforce model was 1.0 clinical FTE for physicians in nonacademic settings (approximately 5% workforce) and 0.8 clinical FTE for those working in academic settings (approximately 95% workforce). The pediatric academic FTE was unique from the adult academic FTE, which was estimated at 0.5. This was because compared to pediatric academic rheumatologists, adult academic rheumatologists spent a greater amount of time in scholarly activities and less time in patient care (8). The nonphysician providers (NPs and PAs) were defined as 0.9 clinical FTE regardless of setting.

Workforce study supply and demand assumptions. Factors influencing supply included geographic domestic patterns of population distribution and density (geographic mobility, net migration, and micropolitan statistical areas), practice setting and productivity, succession trends, sex and generational breakdown, and demographic breakdown of new graduates entering the rheumatology workforce (Table 1). The base model assumed no geographic changes over 10 years, that providers working in micropolitan statistical areas worked 15% less than those who worked outside those areas, and that on average, pediatric rheumatologists worked 55 hours per week.

Factors influencing demand included health care utilization patterns, the prevalence of disease, changes in patient demographics, and gross domestic product (GDP) per capita income

Table 1. 2015 ACR workforce study supply and demand base-model assumptions*

	Base-model assumptions
Supply factors	
Geographic	No geographic changes in the model over next 10 years
	Physicians practicing in MSAs work on average 15% fewer hours than those not working in these areas
	On average pediatric rheumatologists work 55 hours per week
Productivity (RVUs)	Pediatric subspecialties saw an increase by 8.0% for compensation per work RVU in 2013
	The work RVU changed by 6.1%, resulting in an increase in compensation of 1.0%
Succession planning	Approximately 32% of pediatric rheumatologists plan to retire in the next 5–10 years
	Approximately 80% of those who plan to retire anticipate a decrease in their patient load by 25%; therefore we factored a three-quarter FTE for those who plan to retire
Sex	In 2015, 68% were female and 32% male
	Females are reported to work 7 fewer hours each week on average
	Females treated approximately 30% fewer than their male counterpart
Full-time vs. part-time employment	Assumed 17.5% work part-time
	Part-time were then assumed to work 0.5 FTE
Practice setting	Approximately 5% nonacademic settings and 95% academic medical center
	One pediatric rheumatologist in nonacademic settings would equal 1 FTE
	One pediatric rheumatologist in an academic medical center would equal 0.8 FTE
New graduate entrants	Approximately 25 graduates annually; 3.9% do not graduate
	Approximately 42.6% are IMGs; approximately 23.9% of the IMGs will practice outside the US
	Approximately 18% will work part-time; approximately 90% of those working part-time are female
	All entering fellows are millennials
Nonphysician providers (NPs/PAs)	Approximately 25% increase in NPs and 25% increase in PAs between 2015 and 2030
Demand factors	
Aging population	Population of children expected to increase by approximately 3% between 2015 and 2030
Prevalence of disease	Females 2.5 times more likely to have rheumatic disease than males
Per capita income	Approximately 1.5% increase
Medicaid expansion	Approximately 30% by 2030 for eligible Medicaid beneficiaries

* Sources: American College of Rheumatology (ACR), 2015 (9); American College of Rheumatology Committee on Rheumatology Training and Workforce Issues, 2013 (10), US Census Bureau (15); Health Resources and Services Administration, 2016 (11); American Board of Pediatrics, 2015 (12); Association of American Medical Colleges, 2016 (13); Accreditation Council for Graduate Medical Education, 2015 (14). FTE = full-time equivalent; IMG = international medical graduate; MSA = micropolitan statistical areas; NPs = nurse practitioners; PAs = physician assistants; RVU = relative value unit.

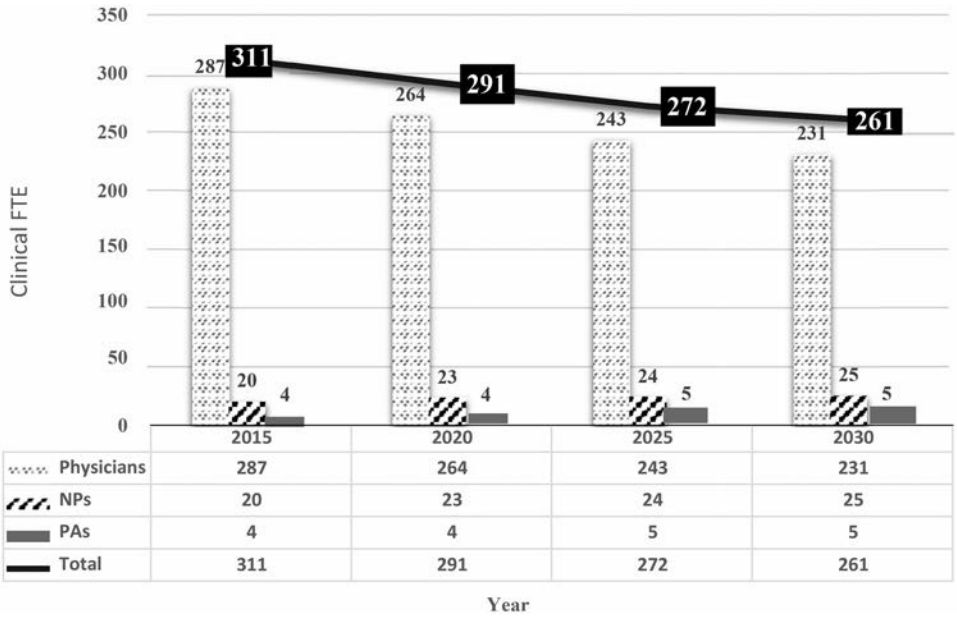


Figure 1. Projected pediatric rheumatology clinical full-time equivalent (FTE) from 2015 through 2030. NPs = nurse practitioners; PAs = physician assistants.

overall and by region (10–14). While the projected population increase in children was anticipated to be relatively small (approximately 3–4%) from 2015–2030, this change was factored in the demand model (15). While the projected effect of the aging US population was far less on pediatric rheumatology than on adult rheumatology, the cost of rheumatology care and GDP per capita income impact was also evaluated. In the 2015 workforce study, a sample of patients was queried to evaluate perceived need and access, which added a new perspective to the supply and demand modeling.

Based on the information collected, the workforce study identified shifts in the demographic breakdown (e.g., sex and

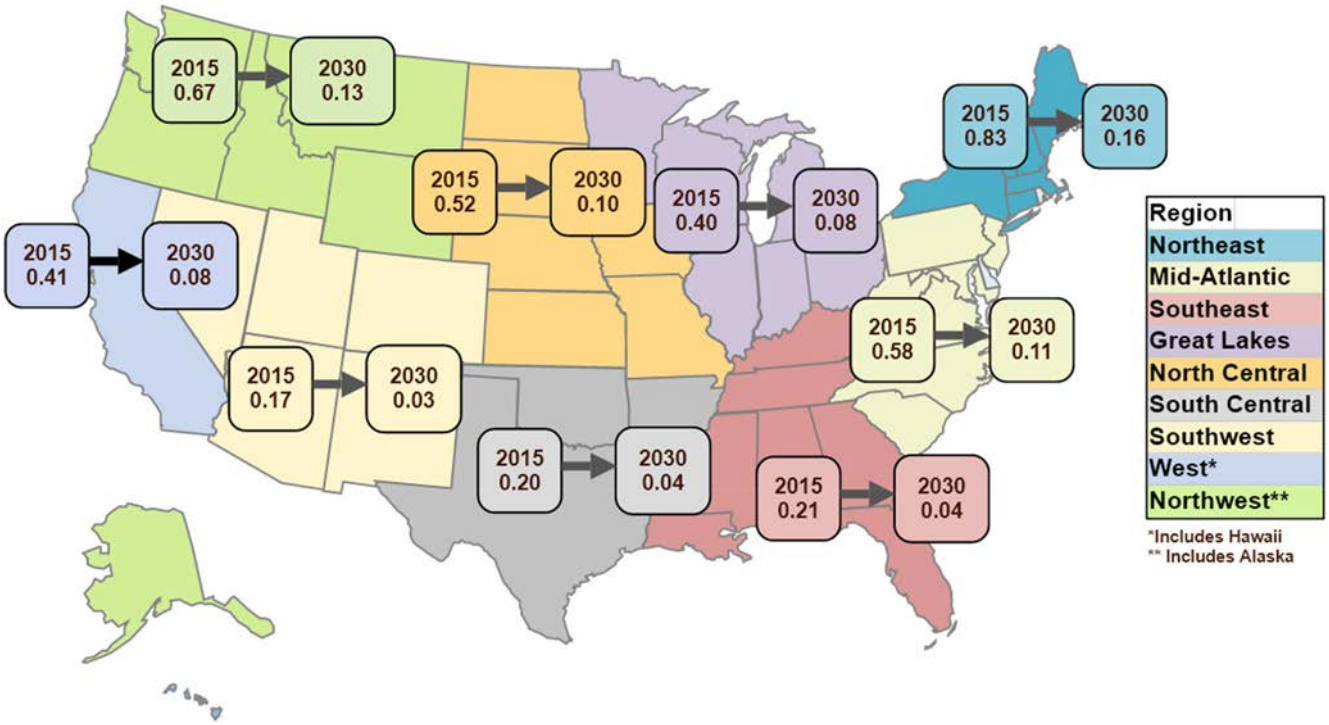
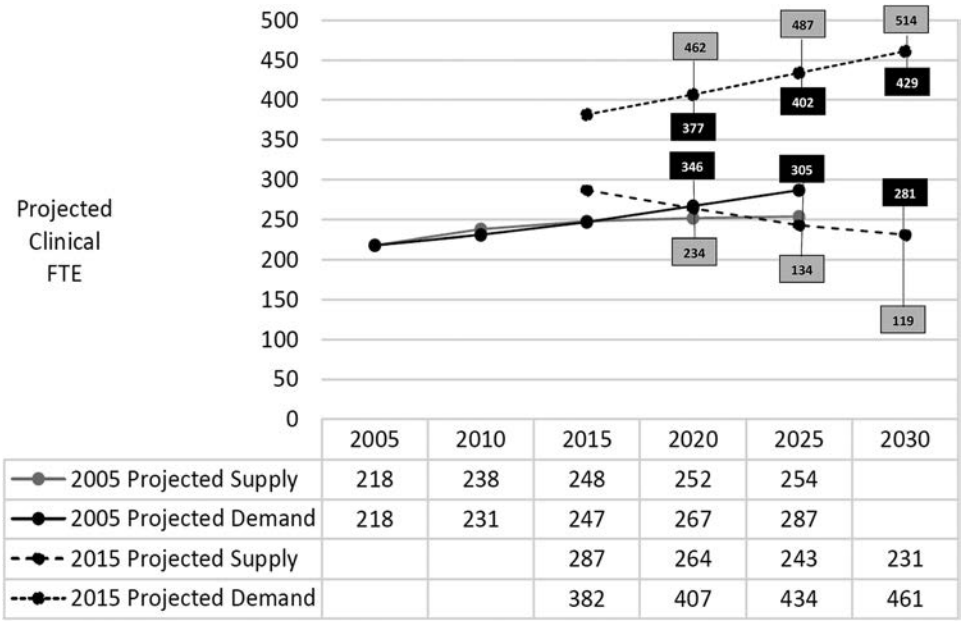


Figure 2. Pediatric rheumatology distribution rate per 100,000 children (2015 versus 2030).



	2005	2010	2015	2020	2025	2030
2005 Projected Supply	218	238	248	252	254	
2005 Projected Demand	218	231	247	267	287	
2015 Projected Supply			287	264	243	231
2015 Projected Demand			382	407	434	461

Figure 3. Projected gap between supply of rheumatologists and demand. This figure includes the previously predicted projection from the 2005 workforce study (4). FTE = full-time equivalent.

generational differences), geographic distribution trends, and practice patterns that indicated a much larger decline in the supply of pediatric rheumatology effort than projected in the 2005 workforce study (4). This decline in supply was theorized to be multifactorial, with an increased number of retiring rheumatology providers, the expansion of part-time providers in the workforce, and the increased number of rheumatology graduates seeking part-time employment. Multivariate and logistic regression with backward stepwise analysis was used to determine factors that contributed significantly to the model for pediatric rheumatology services ($F = 39.06$, $P < 0.001$; $R^2 = 0.37$). Goodness-of-fit tests were used to determine model fit.

Sensitivity testing. To address the variability in the results from the base-model, sensitivity analyses were conducted. Sensitivity testing is an analytic methodology used to build confidence in results. It allows for alternate models to be used in conjunction with a base-case model that incorporates best-estimated values of all selected parameters (16). Sensitivity testing was used to ascertain a best-case and worst-case scenario, providing an estimated range of supply for and demand of services through 2030 (see Supplementary Table 1, available on the Arthritis Care & Research website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24497/abstract). Sensitivity testing is critical to provide for a range in variability that can occur when making future projections.

RESULTS

Baseline rheumatology workforce. Pediatric rheumatology providers were defined as rheumatologists, NPs, and PAs who specialized in treating pediatric patients. Calculations were conducted based on the estimated time providers spent treating patients (referred to as clinical FTE). Figure 1 depicts the pediatric rheumatology workforce supply projections in provider clinical FTEs, including PAs and NPs, from 2015 through 2030. The projections anticipate a 16% decrease between 2015 and 2030.

Demand and supply factors. The factors that were used to assess the future demand of the pediatric rheumatology services included changes in population demographics, health care utilization patterns, practice trends, GDP per capita income, and net migration/geographic trends. Unlike the adult rheumatology workforce, aging was not a major driving force, because according to the US Census Bureau, the population of children age >18 years was not expected to increase significantly between 2014 and 2030, remaining at approximately 74 million by 2020 and 76 million by 2030 (15,17,18). Therefore, population demographics and geographic trends played less of a role in the demand in pediatric rheumatology compared to adult rheumatology. Based on GDP per capita compound growth from 2010 to 2015 and the forecasted value for 2020, an estimated compound growth for 2015–2030 would be approximately 2.5%, up 1.5% from the 2005 study (19–21). In 2015, the growth of the real GDP per capita in the US was approximately 1.5% compared to the previous year. While the GDP per capita continues to rise, the percentage of increase is expected to decrease beginning in 2018 through 2022 (22). Of the factors used to assess future supply for pediatric rheumatology specialists, 3 major drivers included workforce practice trends, access to care/geographic distribution

Table 2. Potential solutions to increase the supply of the pediatric rheumatology workforce*

Increase recruitment of physicians and nonphysician providers
Create a 2-year fellowship for pediatric rheumatologists seeking a clinical-focused career
Implement initiatives to expose more PAs and NPs to rheumatology and increase their recruitment to the field
Increase exposure to pediatric rheumatology in medical school and residency
Give financial incentives (higher salary and/or loan forgiveness)
Optimize the geographic distribution of rheumatologists to improve access to quality care
Extend the use of telemedicine
Providers have significant experience since the COVID-19 pandemic
Reduce referrals of patients with nonrheumatic diseases
Expand rheumatology training in primary care residencies and continuing medical education
Improve rheumatology quality care initiatives in primary care

* PAs = physician assistants; NPs = nurse practitioners.

of rheumatology services, and changes in the demographic breakdown of the new graduates entering the workforce (Table 1) (23–25).

Current workforce practice trends. Given the aging pediatric rheumatology workforce and taking into consideration the current low numbers of pediatric rheumatology providers in the US, succession patterns (e.g., retirement, anticipated changes in workload, etc.) are critical. Labor workforce participation rates for providers of a given age, sex, and international medical graduate status from year to year were reflected in the projections. In addition, sex and millennial workforce practice trends were also included.

Income variability and access to rheumatology workforce. Access to care was defined as physician per population and geographic trends/net migration. While the overall trends show an increase, income varies widely between demographics within the US (26). The poverty rate in the US in 2015 was approximately 15% (15). Poverty rates are persistently higher in rural and inner-city parts of the country as compared to suburban areas. Moreover, 29 states had lower median income, and 18 states had higher median income. When reviewing geographic trends of pediatric providers, there were 3 geographic areas of particular concern. The Southeast had only 0.21 providers per 100,000 children, with a projection of 0.04 per 100,000 in 2030, the South Central region had 0.2 providers per 100,000 children, with a projection of 0.04 per 100,000 in 2030, and the Southwest had 0.17 providers per 100,000, with a projection rate of only 0.03 in 2030 (Figure 2).

New graduates entering the workforce and succession planning. When considering the future supply of pediatric rheumatologists, graduating fellows who enter the workforce were an important factor in the model. The calculated number depended on available fellowship positions, the fill rate of those positions, graduation rates, and number of international medical graduates who anticipate remaining in the US. Other factors that contributed to the entering workforce calculations included sex shifts. Overall, 68% of the pediatric rheumatology workforce was female. Our model assumed that 18% of new graduates entering the workforce would work part-time and 90% of those were female. Our workforce study group survey indicated that 32% of pediatric rheumatologists planned to retire within the next 10 years. Moreover, approximately 80% of those who plan to retire anticipate a decrease in their patient load by 25%. There were approximately 25 pediatric rheumatology fellows graduating each year. Our model predicts that by 2025 there will be an overall loss between retirees and new fellow graduate entrants of 27 providers.

Supply-demand projections. The supply and demand projections of pediatric rheumatology services included NPs and PAs. Figure 1 compares the total number of rheumatology providers (physician and nonphysician) to the projected clinical FTEs of all providers from 2015 to 2030. Figure 3 shows the projected gap between supply of rheumatologists and demand. This figure includes the previously predicted projection from the 2005 workforce study (4). By 2030, the projected supply of pediatric rheumatologist clinical FTEs is 231 compared to a projected demand of 461, thus projecting a net deficit of 230 clinical FTEs.

DISCUSSION

The pediatric rheumatology workforce shortage has been a recognized problem for decades. Although pediatric rheumatology has grown substantially (10-fold) since its beginnings in the 1970s (27), the workforce is approximately 300 providers in the US, which is still a major shortage. The aim of our study was to reassess the trends in supply and demand for pediatric rheumatology care. Notably, the ABP also conducted a pediatric rheumatology workforce study (2018) (28). However, this study primarily used board-certification status as a proxy for clinically available pediatric rheumatology providers, whereas our study attempted to define clinical FTE to more accurately reflect clinically available rheumatologists. This difference is important because most pediatric rheumatologists work in academic settings, and clinical FTE in academia is typically less than that of community practice. At

the time of the workforce study, the academic clinical FTE was determined to be 0.8, based on the fact that anecdotally, most pediatric rheumatologists held clinical educator positions with 0.8 clinical FTE. However, a more recent survey conducted by the American Academy of Pediatrics in 2018 demonstrated that most pediatric rheumatologists self-reported spending only 54% of their time (0.54 FTE) in direct patient care (29). Therefore, currently, the clinical FTE may be closer to 0.5 or 0.6, resulting in an even greater workforce gap than this model predicted. Our study also estimated that approximately 30% of practicing pediatric rheumatologists will retire in the next 10 years. Supportive of this projection, data from the ABP show that approximately 35% of board certified pediatric rheumatologists are age >50 years (28).

The shortage of providers most certainly affects the quality of care of children with rheumatic diseases, as primary care providers refer children to nonrheumatologist pediatric subspecialists and adult rheumatologists (30–32). To provide the highest quality of care, children should be treated by providers with specialized training in pediatric rheumatology and who understand the unique challenges of evaluating and treating a growing child. Given the prediction of a significant workforce shortage, several strategies must be considered to address this problem, including increasing recruitment of physicians and nonphysicians into pediatric rheumatology, promoting changes in the geographical distribution of providers, extending the use of telemedicine, and improving quality of care initiatives in primary care (Table 2).

The ACR and the Childhood Arthritis and Rheumatology Research Alliance both have programs aimed to improve recruitment of pediatric residents into the specialty, and as these programs mature, they should be assessed to determine whether these interventions have been effective (33,34). There are several recognized barriers to recruiting physicians into pediatric rheumatology. These include resident debt, lack of exposure in medical school and residency, concern about being the only specialist in a state or hospital, lower salary than other pediatric specialties and length of training (3-year pediatric fellowship without the 2-year option offered as in adult rheumatology fellowship) (5,6). With only 20–30 new fellows graduating each year, substantial recruitment efforts are needed. The majority of pediatric rheumatologists work in academic institutions in which there is an expectation that academic work requires additional training. Therefore, 3-year fellowships have been the norm in pediatric subspecialties. Few pediatric rheumatologists work in community practice, so a possible solution might be to create a 2-year fellowship for physicians seeking to work in community practice and/or creating strictly clinical positions within academic institutions.

Other measures to improve supply must include recruiting and training more PAs and NPs into the pediatric rheumatology workforce; they have been effectively used to treat adult rheumatology patients (35,36). Financial incentive programs, including medical student loan debt relief are also important. Loan repayment programs have been employed to increase primary care providers in underserved areas (37). A similar loan repayment program for pediatric rheumatology has been introduced to the US Senate but to date has not moved (38).

An important aspect of the workforce supply issue is not only having too few pediatric rheumatologists but also the imbalanced geographical distribution of providers. According to the ABP, there are 9 states (Alaska, Idaho, Montana, New Hampshire, New Mexico, Oklahoma, South Dakota, West Virginia, and Wyoming) without a practicing board-certified pediatric rheumatologist. Several of these states have coverage by outreach programs from other states (28). However, an equally important problem is that several states with large populations of children (e.g., Texas) have only a few pediatric rheumatologists to treat them. Telemedicine has been considered an important possible solution to the geographic barriers to augment timely consultation, reduce patient travel costs and provide access to care, and modify medical management for diagnosed patients. Prior to the COVID-19 pandemic, few pediatric rheumatology telemedicine programs existed (39,40). However, after the COVID-19 pandemic, use of telemedicine skyrocketed across health care in the US, including pediatric rheumatology, and so we will likely see a continuation in telehealth care (41). Anecdotally, we have found patient and provider satisfaction with telemedicine, but studies are needed to optimally assess quality of care in this setting, with a particular emphasis on the quality of the joint examination in telemedicine.

More efforts are needed to reduce the demand on pediatric rheumatologists. Education for primary care providers in conducting musculoskeletal examinations and ordering of rheumatology tests may help reduce referrals of patients with nonrheumatic diseases (42,43). Such training has been successful in adult medicine (44).

A strength of this study was that it used an integrative approach to assess not only the changes in pediatric rheumatology workforce over time, but also integrated changes in the US population, economy, and geographic distribution of providers. Sensitivity testing was used to ascertain best- and worst-case scenarios to establish a range of supply and demand. Importantly, this study also included the patient's perspective on barriers to access to care, and patients reported substantial direct and indirect costs for them when trying to access this care (45). The lack of workforce supply is not limited to pediatric rheumatology; adult rheumatology and several pediatric specialties face similar workforce supply challenges (8,28,46). We believe that this study can serve as a model for assessing workforce problems in other specialties as well.

There were several limitations that are important to highlight. First, it was difficult to determine accurately the number of providers in the workforce who actually treat patients, the ratio of nonacademic and academic providers, the number of medicine/pediatric subspecialists, and how they were documented to

ensure they were not being counted twice. Second, the clinical FTE was selected based on the limited information that was available at the time and cannot be considered 100% accurate. Next, the primary data collection was conducted using the ACR membership, which may limit the generalizability to the overall rheumatology workforce. Notably, the findings from this 2015 workforce study demonstrate a significant worsening in the workforce gap compared to the 2005 study. The supply and demand model is complex, taking into account several population-level factors, in addition to direct rheumatology practice measures such as FTE and disease prevalence. Although great attention was taken in creating the model assumptions, some of the assumptions were possibly inaccurate and thus overestimated the workforce gap, in comparison to the 2005 study. However, the primary purpose of these projections is to demonstrate important trends in workforce gaps and to identify access to care concerns for pediatric rheumatology care with potential solutions for the future.

In conclusion, this ACR/ARP workforce study has demonstrated that the pediatric rheumatology workforce is not meeting demand, and projections show that this excess demand is increasing significantly. Based on our model, by 2030, we are likely to have only half the supply of pediatric rheumatology care needed to meet the demand. Innovative strategies are needed to increase the workforce supply and to improve access to care for pediatric rheumatology patients.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. Dr. Correll 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. Ditmyer, Imundo, Klein-Gitelman, Monrad, Battafarano.

Acquisition of data. Ditmyer.

Analysis and interpretation of data. Correll, Ditmyer, Mehta, Monrad, Battafarano.

REFERENCES

1. Sacks JJ, Helmick CG, Luo YH, Ilowite NT, Bowyer S. Prevalence of and annual ambulatory health care visits for pediatric arthritis and other rheumatologic conditions in the United States in 2001–2004. Arthritis Rheum 2007;57:1439–45.
2. Cassidy JT, Athreya B. Pediatric rheumatology: status of the subspecialty in United States medical schools [letter]. Arthritis Rheum 1997; 40:1182.
3. Althouse LA, Stockman JA. Pediatric workforce: a look at pediatric rheumatology data from the American Board of Pediatrics. J Pediatr 2006;149:869–70.
4. Deal CL, Hooker R, Harrington T, Birnbaum N, Hogan P, Bouchery E, et al. The United States rheumatology workforce: supply and demand, 2005–2025. Arthritis Rheum 2007;56:722–9.
5. Henrickson M. Policy challenges for the pediatric rheumatology workforce. Part I. Education and economics. Pediatr Rheumatol Online J 2011;9:23–4.
6. Henrickson M. Policy challenges for the pediatric rheumatology workforce. Part II. Health care system delivery and workforce supply. Pediatr Rheumatol Online J 2011;9:24.
7. Henrickson M. Policy challenges for the pediatric rheumatology workforce. Part III. The international situation. Pediatr Rheumatol Online J 2011;9:26.
8. Battafarano DF, Ditmyer M, Bolster MB, Fitzgerald JD, Deal C, Bass AR, 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. American College of Rheumatology. 2015 workforce study of rheumatology specialists in the United States. URL: https://www.rheumatology.org/portals/0/files/ACR-Workforce-Study-2015.pdf.
10. American College of Rheumatology Committee on Rheumatology Training and Workforce Issues. Regional distribution of adult rheumatologists. Arthritis Rheum 2013;65:3017–25.
11. Health Resources and Services Administration. Health professional shortage areas. 2016. URL: https://bhw.hrsa.gov/shortage-designation/hpsas.
12. American Board of Pediatrics. Data and workforce. 2015. URL: https://www.abp.org/content/data-and-workforce.
13. Association of American Medical Colleges. Data and reports. 2016. URL: https://www.aamc.org/data-reports.
14. Accreditation Council for Graduate Medical Education. Data collection systems. 2015. URL: https://www.acgme.org/Data-Collection-Systems/Overview.
15. United States Census Bureau. Income and poverty in the United States: 2015. URL: https://www.census.gov/content/dam/Census/library/publications/2016/demo/p60-256.pdf.
16. Ding P, VanderWeele TJ. Sensitivity analysis without assumptions. Epidemiology 2016;27:368–77.
17. National Commission on Certification of Physician Assistants. 2016 statistical profile of certified physician assistants by specialty. URL: https://prodcmsstoragesa.blob.core.windows.net/uploads/files/2016StatisticalProfilebySpecialty.pdf.
18. Ewert S. U.S. Census Bureau data on the foreign born. URL: https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.10/2016/mtg2_WS/4_Ewert_USA.pdf.
19. American Association of Colleges of Nursing. Core competencies for interprofessional collaborative practice. 2011. URL: http://apps.who.int/iris/bitstream/handle/10665/70185/WHO_HRH_HPN_10.3_eng.pdf;jsessionid=F8FB7FB7088BFE33FAA1BD17E034069D?sequence=1.
20. Framework for Action on Interprofessional Education and Collaborative Practice Health Professions Networks Nursing and Midwifery Human Resources for Health. 2010. URL: http://www.who.int/hrh/nursing_midwifery/en/.
21. The WHO STEPwise approach to chronic disease risk factor surveillance. 2005. URL: https://apps.who.int/iris/bitstream/handle/10665/43376/9241593830_eng.pdf.
22. Statista. Gross domestic product (GDP) per capita in the United States in current prices from 1986 to 2026. URL: https://www.statista.com/statistics/263601/gross-domestic-product-gdp-per-capita-in-the-united-states/.
23. U.S. Bureau of Labor Statistics. Occupational outlook handbook: physicians and surgeons. URL: https://www.bls.gov/ooh/healthcare/physicians-and-surgeons.htm.
24. U.S. Bureau of Labor Statistics. Career outlook. Older workers: labor force trends and career options. URL: https://www.bls.gov/careeroutlook/2017/article/older-workers.htm.
25. Association of American Medical Colleges. Research shows shortage of more than 100,000 doctors by 2030. URL: https://www.aamc.org/news-insights/research-shows-shortage-more-100000-doctors-2030.

26. U.S. Bureau of Labor Statistics. Employment projections. Education pays: earnings and unemployment rates and earnings by educational attainment, 2020. URL: https://www.bls.gov/emp/chart-unemployment-earnings-education.htm.
27. Lehman TJ, Magilavy DB, Warren R. Pediatric rheumatology manpower and training: planning for the 1990s. Pediatrics 1989;84: 567–8.
28. The American Board of Pediatrics. Pediatric physicians workforce data book, 2018-2019. 2018. URL: https://www.abp.org/sites/abp/files/workforcedata2018-2019.pdf.
29. Rimsza ME, Ruch-Ross HS, Clemens CJ, Moskowitz WB, Mulvey HJ. Workforce trends and analysis of selected pediatric subspecialties in the United States. Acad Pediatr 2018;18:805–12.
30. Correll CK, Spector LG, Zhang L, Binstadt BA, Vehe RK. Barriers and alternatives to pediatric rheumatology referrals: survey of general pediatricians in the United States. Pediatr Rheumatol Online J 2015; 13: 32.
31. Mayer ML, Sandborg CI, Mellins ED. Role of pediatric and internist rheumatologists in treating children with rheumatic diseases. Pediatrics 2004;113:e173–81.
32. Sherry DD, Wallace CA, Kahn SJ. Pediatric rheumatology in adult rheumatology practices in Washington state. Arthritis Rheum 1996; 39:1218–21.
33. American College of Rheumatology. Pediatric rheumatology residents program. URL: https://www.rheumatology.org/I-Am-A/Student-Resident/Explore-Rheumatology/Pediatric-Rheumatology-Residency-Program.
34. Childhood Arthritis and Rheumatology Research Alliance. Pediatric residents program. URL: https://carragroup.org/meetings-events/past-carra-annual-scientific-meetings/2021-annual-meeting/2021-pediatric-residents-program.
35. Solomon DH, Bitton A, Fraenkel L, Brown E, Tsao P, Katz JN. Roles of nurse practitioners and physician assistants in rheumatology practices in the US. Arthritis Care Res (Hoboken) 2014;66:1108–13.
36. Hooker RS. The extension of rheumatology services with physician assistants and nurse practitioners. Best Pract Res Clin Rheumatol 2008;22:523–33.
37. Scarbrough AW, Moore M, Shelton SR, Knox RJ. Improving primary care retention in medically underserved areas. Health Care Manag (Frederick) 2016;35:368–72.
38. Reed J. All Info - S.2443 - 116th Congress (2019-2020): Investment in tomorrow's pediatric health care workforce act. 2019.
39. Kessler EA, Sherman AK, Becker ML. Decreasing patient cost and travel time through pediatric rheumatology telemedicine visits. Pediatr Rheumatol Online J 2016;14:54.
40. Bullock DR, Vehe RK, Zhang L, Correll CK. Telemedicine and other care models in pediatric rheumatology: an exploratory study of parents' perceptions of barriers to care and care preferences. Pediatr Rheumatol Online J 2017;15:55.
41. Contreras CM, Metzger GA, Beane JD, Dedhia PH, Ejaz A, Pawlik TM. Telemedicine: patient-provider clinical engagement during the COVID-19 pandemic and beyond. J Gastrointest Surg 2020;24: 1692–7.
42. Correll CK, Spector LG, Zhang L, Binstadt BA, Vehe RK. Use of rheumatology laboratory studies among primary pediatricians. Clin Pediatr (Phila) 2016;55:1279–88.
43. McGhee JL, Burks FN, Sheckels JL, Jarvis JN. Identifying children with chronic arthritis based on chief complaints: absence of predictive value for musculoskeletal pain as an indicator of rheumatic disease in children. Pediatrics 2002;110:354–359.
44. Bankhurst A, Olivas C, Larson JH, Bradford A, Fields R, Kalishman S, et al. Rheumatology care in under-resourced areas utilizing the ECHO model . Arthritis Care Res (Hoboken) 2019.
45. Monrad S, Imundo L, Battafarano D DM. Access to care: the patient perspective from the 2015 ACR/ARHP workforce study. Arthritis Rheumatol 2016;68.
46. Mayer ML. Are we there yet? Distance to care and relative supply among pediatric medical subspecialties. Pediatrics 2006;118:2313–2321.

Arthritis Care & Research
Vol. 74, No. 3, March 2022, pp 349–354
DOI 10.1002/acr.24483

American College of Rheumatology
Empowering Rheumatology Professionals

Temporal Relationship Between Juvenile Idiopathic Arthritis Disease Activity and Uveitis Disease Activity

Emily J. Liebling,[1] Walter Faig,[1] Joyce C. Chang,[2] Elizabeth Mendoza,[1] Nicholas Moore,[3] Nahomy Ledesma Vicioso,[4] and Melissa A. Lerman[2]

Objective. To determine whether there is a temporal association between arthritis and uveitis activity among children with juvenile idiopathic arthritis–associated uveitis (JIA-U).

Methods. Uveitis and arthritis data from patients with JIA-U age ≤21 years were collected from July 2013 to December 2019 at a tertiary care center. Arthritis activity was assessed at each rheumatology visit, and the primary outcome was the presence of active uveitis at ophthalmologic examination within 45 days of the rheumatology visit. Repeated-measures logistic regression was used to evaluate the temporal association between any uveitis activity within 45 days of arthritis activity. Models were adjusted for demographic-, disease-, and treatment-related factors.

Results. A total of 98 patients were included: 81 (83%) female, 67 (69%) antinuclear antibody positive, 59 (60%) oligoarticular, and 13 (13%) enthesitis-related arthritis (ERA) subtypes. There were 1,229 rheumatology visits, with a median of 13 visits per patient (interquartile range 7–18). Concordance between arthritis and uveitis activity was observed 73% of the time (694 of 947). There was an independent temporal association between uveitis and arthritis activity (odds ratio 2.47 [95% confidence interval 1.72–3.54]; $P < 0.01$), adjusted for demographic and disease characteristics. Use of combination biologic and nonbiologic disease-modifying antirheumatic drugs, female sex, HLA–B27 positivity, and ERA and polyarticular (rheumatoid factor negative) subtypes were associated with decreased odds of active uveitis at any time point.

Conclusion. In patients with JIA-U, there is a significant temporal association between arthritis and uveitis disease activity. These novel results suggest that an arthritis flare should prompt an expedited referral to the ophthalmologist.

INTRODUCTION

The uveitis associated with juvenile idiopathic arthritis (JIA-U) accounts for an estimated 20–40% of cases of childhood noninfectious uveitis. It is most often insidious in onset and exhibits a chronic course, with JIA subtypes incurring various risks of ocular involvement (1–3). Even after uveitis is identified and treated, patients have a lifetime risk of recurrence and can develop complications well into adulthood (4). Patient- and disease-specific factors associated with uveitis development in JIA have been identified, and those patients with the highest risk undergo frequent ophthalmologic screening (3,5). Specifically, female sex, antinuclear antibody (ANA) positivity, young age of arthritis onset (≤6 years), and oligoarticular subtype fall into this category.

Contrary to the historical dogma of arthritis and uveitis exhibiting distinct and unrelated courses, emerging data over the last 20 years have demonstrated that the arthritis activity of patients with JIA-U may be more persistent than that of JIA patients without uveitis (6,7). Despite the suggestion of a relationship between the 2 entities, little is known about the temporal nature of the association. Rosenberg and Oen's dedicated 1986 study did not establish arthritis and uveitis to run parallel courses, but the generalizability was limited by the small sample size (n = 35), long intervals between examinations (6 months), and lack of defined time window of concomitant disease activity (8). To our

Supported by the Children's Hospital of Philadelphia Rheumatology Research Fund.

[1]Emily J. Liebling, MD, Walter Faig, PhD, Elizabeth Mendoza, BA: The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania; [2]Joyce C. Chang, MD, MSCE, Melissa A. Lerman, MD, PhD, MSCE: The Children's Hospital of Philadelphia and Perelman School of Medicine of the University of Pennsylvania, Philadelphia; [3]Nicholas Moore, BA: Perelman School of Medicine of the University of Pennsylvania, Philadelphia; [4]Nahomy Ledesma Vicioso, BS: Weill Cornell Medical College, New York, New York.

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.24483&file=acr24483-sup-0001-Disclosureform.pdf.

Address correspondence to Emily J. Liebling, MD, The Children's Hospital of Philadelphia, Division of Pediatric Rheumatology, 3401 Civic Center Boulevard, Philadelphia, PA, 19104. Email: lieblinge@email.chop.edu.

Submitted for publication July 1, 2020; accepted in revised form October 6, 2020.

SIGNIFICANCE & INNOVATIONS

- Uncertainty exists as to whether arthritis activity and uveitis activity exhibit parallel courses in juvenile idiopathic arthritis–associated uveitis (JIA-U), and whether children should be evaluated for active uveitis at the time of arthritis flare.
- This is the first large, dedicated cohort study to demonstrate a strong temporal association between arthritis and uveitis activity in the JIA-U disease course.
- The presence of arthritis activity should prompt expedited referral to the ophthalmologist to facilitate early diagnosis and treatment of a uveitis flare.

knowledge, there has since been no large, dedicated systematic evaluation of arthritis activity as it temporally relates to uveitis activity over the JIA-U disease course.

Identification of such a relationship would inform management of children with known JIA-U, as an arthritis flare would prompt immediate referral for assessment of uveitis activity. Here we aimed to determine whether there is a temporal association between arthritis and uveitis disease activity, and which patient- and disease-specific factors are associated with greater uveitis activity over time.

PATIENTS AND METHODS

This is a retrospective, single center, longitudinal cohort study using medical record data of patients with JIA-U treated at The Children's Hospital of Philadelphia between July 1, 2013 and December 1, 2019. To be included, patients must have been age ≤21 years at enrollment, have a physician diagnosis of JIA, of any International League of Associations of Rheumatology subtype, and have a history of uveitis, of any Standardization of Uveitis Nomenclature (SUN) subtype (9,10). Patients were excluded if they had any other autoinflammatory or autoimmune disease. Of the medication exposure during the reporting period, etanercept was excluded due to known inefficacy for the treatment of JIA-U (11). An exemption was granted for this study by The Children's Hospital of Philadelphia Institutional Review Board (13-010355) for the conduct of secondary research for which consent is not required.

Study measures. The primary outcome was the presence of active uveitis at any ophthalmology visit occurring within each 90-day assessment period, up to 45 days before and 45 days after each rheumatology visit. A window of 45 days was chosen to capture the most joint and eye examinations, to produce a robust data set, as most children with JIA-U are evaluated every 3 months by a rheumatologist. Any ophthalmology visit that was not within 45 days of a rheumatology visit was excluded. Uveitis was considered active if at least 1 of the following criteria were met: anterior chamber cell of SUN grade >0.5+, eye examination performed while the patient was administering >2 topical glucocorticoid drops per eye per day, eye examination performed while the patient was treated with oral glucocorticoids, and otherwise defined as active by an ophthalmologist's assessment. This last criterion was established to account for the few examinations performed by ophthalmologists outside of The Children's Hospital of Philadelphia in which SUN grading was not provided. In such cases, uveitis activity was based on the ophthalmologist's assessment. Arthritis was assessed at each rheumatology visit and was considered active in 1 or more joints if there was swelling or tenderness with limited range of motion, or if arthritis was otherwise deemed active by a rheumatologist's assessment.

Patient- and disease-related covariates were considered, including sex, JIA subtype, ANA status, HLA–B27 status, years from uveitis diagnosis, ocular symptoms of pain, redness, vision changes, or photophobia, disease-modifying antirheumatic drug (DMARD), and systemic glucocorticoid use. With respect to JIA subtype, the polyarticular rheumatoid factor (RF) positive and undifferentiated subtypes were combined with those of unknown status as a single category. This combination was performed due to the low number of patients within each of the 3 groups, as analysis of each individual category would not have produced meaningful results.

Statistical analysis. Descriptive statistics of mean ± SD or median (interquartile range [IQR]) were used to summarize demographic and disease characteristics, as well as rheumatology visit characteristics. Repeated-measures logistic regression with generalized estimating equations was used to assess estimates of correlation. For the primary analysis, uveitis status was imputed for rheumatology visits with no eye examination within the proximate 45 days of arthritis assessment, using predicted probabilities from a logistic model of complete cases. All models were adjusted for time, and forward selection methods (*P* value for entry ≤0.2) were used to select additional covariates for inclusion. Sex, ANA status, HLA–B27 status, and JIA subtype were determined a priori to be potential confounders and were forced into the final multivariable model. Chi-square, Student's *t*-test, Kruskal-Wallis test, and Wilcoxon's rank sum tests were performed, as appropriate, to compare covariates across visits by uveitis status. ANA and HLA–B27 status were treated as binary variables (positive versus negative/unknown) in all models, and therefore missingness occurred only with uveitis activity. The distribution of arthritis activity was compared between patients with and without missing eye examinations via chi-square tests to assess patterns of missingness, and a missing-at-random mechanism was assumed. The distributions of other covariates were similarly evaluated and those predictive of missingness were included in the imputation regression.

Several sensitivity analyses were performed to evaluate our cohort definition and to address potential bias due to missing

Table 1. Baseline patient demographic and disease characteristics*

Characteristic	Total (n = 98)
Female	81 (83)
Age, median (IQR) years	
At uveitis diagnosis	5.1 (3.2–7.9)
At JIA diagnosis†	3.3 (2.0–6.7])
JIA subtype	
Oligoarticular	59 (60)
Polyarticular (RF negative)	16 (16)
Enthesitis related	13 (13)
Psoriatic	5 (5)
Other/unknown‡	5 (5)
Uveitis subtype	
Anterior	96 (98)
Panuveitis	2 (2)
ANA positive§	67 (69)
HLA–B27 positive¶	10 (12)
Visits per subject, median (IQR)	13 (7–18)
Medications#	
Biologic DMARD**	69 (70)
Nonbiologic DMARD††	81 (83)
Combination DMARD	64 (65)
Systemic glucocorticoids	10 (10)
Neither	12 (12)

* Values are the number (%) unless indicated otherwise. ANA = antinuclear antibody; DMARD = disease-modifying antirheumatic drug; IQR = interquartile range; JIA = juvenile idiopathic arthritis; RF = rheumatoid factor.
† Two patients have unknown date of JIA diagnosis (n = 96).
‡ Includes polyarticular (RF+) and undifferentiated subtypes.
§ Percentage of patients for whom ANA status is known (n = 97).
¶ Percentage of patients for whom HLA–B27 status is known (n = 83).
# Patients with any exposure to biologic DMARD, nonbiologic DMARD, or both, within the reporting period.
** Infliximab, adalimumab, tocilizumab, rituximab, abatacept, tofacitinib, and ustekinumab.
†† Methotrexate, leflunomide, and mycophenolate mofetil.

data. First, patients with the enthesitis-related arthritis (ERA) subtype of JIA were excluded from the analysis, given the distinct presentation of acute anterior uveitis in this subpopulation. Second, the results from a complete case analysis of only those visits without missing data were compared to the results when including the imputed uveitis status. Lastly, to further address the potential effects of missing data and investigate the robustness of our estimates to possible departures from the missing-at-random assumption, we simulated the range of point estimates that would be obtained if uveitis was active versus inactive at all missing eye examinations. All analyses were performed in SAS software, version 9.4, with a prespecified significance level of 0.05.

RESULTS

Study population. A total of 98 patients were enrolled in the study, with a median of 13 rheumatology visits per patient (IQR 7–18) (Table 1). Of those patients, 83% were female, 60% were oligoarticular subtype, and of the 97 patients for whom ANA status was known, 69% were positive. Over the course of the reporting period, 70% were treated with at least 1 biologic DMARD, 83% with a nonbiologic DMARD, and 65% with combination therapy.

Visit characteristics. There were 1,229 rheumatology visits across the cohort from 2013 to 2019, of which arthritis was active in 17% (Table 2). In total, 18% of all rheumatology visits had active uveitis within 45 days, 59% had inactive uveitis within 45 days, and 23% had no proximate eye exam. Active uveitis occurred proximate to active arthritis at 6% of all rheumatology visits (69 of 1,229), while inactive uveitis occurred proximate to inactive arthritis at 51% of visits (625 of 1,229), resulting in concordance of joint and eye disease activity at 73% of visits (694 of 947) in which uveitis status was known. Univariable associations between patient characteristics and active uveitis are shown in Supplementary Table 1, available on the Arthritis Care & Research website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24483/abstract.

Arthritis and uveitis temporal association. There was a 2.5-fold higher adjusted odds of active uveitis proximate to visits with active arthritis compared to those with inactive arthritis (95% confidence interval [95% CI] 1.72–3.54; $P < 0.01$) (Table 3). The predicted probability of active uveitis within 45 days of active arthritis at the mean of all other covariates was 65% (95% CI 0.49–0.77), as compared to 42% (95% CI 0.28–0.57) in those with no active joints.

Table 2. Visit characteristics at all musculoskeletal exams*

Characteristic	Total (n = 1,229)
Disease status	
Active arthritis	
1–4 joints	198 (16)
≥5 joints	7 (1)
Inactive arthritis	1,024 (83)
Proximate active uveitis†	225 (18)
Active uveitis within 45 days of active arthritis	69 (6)
Proximate inactive uveitis†	722 (59)
Inactive uveitis within 45 days of inactive arthritis	625 (51)
Unable to determine uveitis status (missing)	282 (23)
Age at visit, mean (IQR) years	10.9 (7.4–14.3)
Uveitis symptoms‡	30 (2.4)
Medications§	
Biologic DMARD¶	159 (13)
Nonbiologic DMARD#	263 (21)
Combination DMARD	599 (49)
Systemic glucocorticoids	15 (1.2)
Neither	208 (17)

* Values are the number (%) unless indicated otherwise. DMARD = disease-modifying antirheumatic drug; IQR = interquartile range.
† Uveitis status within 45 days of joint examination.
‡ Eye pain, redness, vision changes, and photophobia.
§ Number of visits with biologic DMARD alone, nonbiologic DMARD alone, or both concurrently.
¶ Infliximab, adalimumab, tocilizumab, rituximab, abatacept, tofacitinib, and ustekinumab.
# Methotrexate, leflunomide, and mycophenolate mofetil.

Table 3. Factors associated with active uveitis over time in an adjusted model*

Variable	Adjusted odds ratio (95% CI)	*P*
Active arthritis†	2.47 (1.72–3.54)	<0.01
Female	0.59 (0.38–0.90)	0.01
Age at visit†	1.04 (0.99–1.09)	0.17
Time from uveitis diagnosis, years†	0.94 (0.89–1.00)	0.04
JIA subtype		
Oligoarticular	–	–
Polyarticular (RF negative)	0.47 (0.29–0.75)	<0.01
Enthesitis-related	0.30 (0.16–0.55)	<0.01
Psoriatic	0.78 (0.42–1.42)	0.41
Other/unknown	0.48 (0.24–0.99)	0.05
ANA positivity	0.95 (0.68–1.32)	0.75
HLA–B27 positivity	0.59 (0.35–1.00)	0.05
Symptoms‡	5.48 (2.39–12.57)	<0.01
DMARD use†		
No DMARD	–	–
Biologic DMARD§	0.62 (0.37–1.05)	0.07
Nonbiologic DMARD¶	0.67 (0.43–1.05)	0.08
Combination DMARD	0.45 (0.30–0.68)	<0.01
Systemic glucocorticoid use†	4.43 (1.36–14.43)	0.01

* Multivariable generalized estimating equations logistic regression model estimating the association between arthritis and uveitis (first row), adjusted for time and demographic-, disease- and treatment-related characteristics. Covariates listed with odds of active uveitis at any time point in the reporting period. 95% CI = 95% confidence interval; ANA = antinuclear antibody; DMARD = disease-modifying antirheumatic drug; JIA = juvenile idiopathic arthritis; RF = rheumatoid factor.
† Indicates time-varying covariate, updated at every visit.
‡ Eye pain, redness, vision changes, and photophobia. Indicates time-varying covariate, updated at every visit.
§ Infliximab, adalimumab, tocilizumab, rituximab, abatacept, tofacitinib, and ustekinumab.
¶ Methotrexate, leflunomide, and mycophenolate mofetil.

Female sex was associated with decreased odds of active uveitis at any time point (odds ratio [OR] 0.59 [95% CI 0.39–0.90], $P = 0.01$), as was use of combination biologic and nonbiologic DMARD (OR 0.45 [95% CI 0.30–0.68], $P < 0.01$). All JIA subtypes other than the psoriatic subtype had a lower odds of active uveitis over time compared to the oligoarticular subtype. HLA–B27 positivity was associated with a decreased odds (OR 0.59 [95% CI 0.35–1.00], $P = 0.05$), though this point estimate resulted from comparison to the combination of patients who were HLA–B27 negative (74%) and those in whom HLA–B27 status was unknown (15%). Years from uveitis diagnosis was also associated with a lower odds of active uveitis over time (OR 0.94 [95% CI 0.89–1.00], $P = 0.04$).

Sensitivity analysis. Excluding patients with ERA subtype from the cohort definition did not significantly change our results (OR 2.28 [95% CI 1.55–3.38], $P < 0.01$). With respect to missing data, complete case analysis yielded a similar estimate for the association between arthritis and uveitis activity compared to multiple imputation (Table 4). In the simulation analysis assuming a range of nonrandom missingness in uveitis activity, the temporal association remained significant; imputing all missing ophthalmology visits as having active uveitis yielded OR 1.61 (95% CI 1.14–2.27; $P = 0.01$), whereas imputing all missing visits as inactive uveitis yielded OR 2.45 (95% CI 1.72–3.49; $P < 0.01$). Variables associated with a higher likelihood of missing eye exams were female sex, nonbiologic DMARD use, lack of systemic glucocorticoid use, and the absence of symptoms (data not shown).

DISCUSSION

To date, this is the largest investigation evaluating the temporal association between arthritis activity and uveitis activity in a cohort of patients with JIA-U. The only prior dedicated study did not demonstrate such a relationship, and subsequently, rheumatologists do not feel compelled to screen for uveitis activity at the time of arthritis flare. However, inherent limitations in patient volume and study design limit the application of those conclusions. Using a 45-day period of continuous uveitis assessment surrounding the time of the rheumatology visit, we demonstrated that the 2 diseases often run parallel courses. The magnitude of this association is striking, with an almost 2.5-fold odds of having active uveitis within 45 days of active arthritis, independent of various patient- and disease-related characteristics. Thus, the predicted probability of having active uveitis within 45 days of active arthritis is 65%, suggesting that active arthritis should prompt earlier and more frequent ophthalmologic evaluation.

Additionally, various patient- and JIA disease–related characteristics associated with the presence of active uveitis were identified in this cohort. Use of combination biologic and nonbiologic DMARD, female sex, ERA, and polyarticular RF negative subtypes were associated with a decreased odds of active uveitis at any time point. Notably, female sex is a risk factor for the development of

Table 4. Sensitivity analyses comparing imputation to complete case analysis*

	Odds ratio (95% CI)	*P*
Imputation analysis†		
All JIA subtypes (1,229 visits)	2.47 (1.72–3.54)	<0.01
ERA subtype excluded (1,062 visits)	2.28 (1.55–3.38)	<0.01
Complete case analysis‡		
All JIA subtypes (947 visits)	2.41 (1.64–3.55)	<0.01
ERA subtype excluded (825 visits)	2.44 (1.60–3.71)	<0.01

* Estimates of the association between active arthritis and active uveitis at any visit from generalized estimating equations logistic regression, adjusted for time, demographic-, disease-, and treatment-related characteristics. 95% CI = 95% confidence interval; ERA = enthesitis related arthritis; JIA = juvenile idiopathic arthritis.
† Regression analysis with imputation of missing uveitis status.
‡ Joint examinations with nonmissing eye examinations within preceding or succeeding 45 days.

uveitis ever in children with JIA, but male sex is typically associated with more aggressive eye disease and worse visual outcomes in the JIA-U population (12–14). Therefore, the observation that male patients in our JIA-U cohort were more likely than female patients to have active uveitis at any visit is consistent with this pattern.

In this population, we did not distinguish between chronic anterior uveitis and acute anterior uveitis, as our experience has shown that even patients with chronic anterior uveitis, typically thought to have silent disease, may exhibit symptoms of eye pain, redness, vision changes, and photophobia. Conversely, the JIA subtypes usually associated with acute anterior uveitis may instead manifest asymptomatic eye disease (15). Our findings indicate that a patient with symptoms at any visit was more likely than an asymptomatic individual to have active uveitis at any time point. Furthermore, the ERA subtype is predisposed to acute uveitis that is considered pathophysiologically distinct from chronic uveitis, and thus providers may be less vigilant about frequent ophthalmologic screening (16,17). However, excluding the ERA subtype did not significantly change the temporal association between arthritis and uveitis activity, nor did it ameliorate the association between eye symptoms and active uveitis. In a similar vein, HLA–B27 positivity was suggestive of a lower likelihood of active uveitis, but 15% of patients with unknown HLA–B27 status obviated a conclusive point estimate. Therefore, an arthritis flare in patients of all JIA subtypes should prompt expedited uveitis screening.

The major limitation of our study was the proportion of missing data, as 23% of rheumatology visits had no proximate eye examination in the reporting period. Among the variables associated with higher likelihood of missing eye examinations was female sex, which was also associated with decreased odds of active uveitis at any time point. Therefore, as 86% of the rheumatology visits without proximate eye examinations had inactive arthritis, the true magnitude of concordance between eye and joint disease was likely even greater than our results demonstrate.

Reassuringly, our findings were robust to several sensitivity analyses designed to address missing data. Both complete-case analysis and multiple imputation demonstrated a significant temporal association of similar magnitude between arthritis and uveitis activity. In addition, we simulated estimates at both extremes, in which the propensity of eye exams to be missing was either perfectly associated with active uveitis or perfectly associated with inactive uveitis. Imputation of all 282 rheumatology visits with either proximate active or inactive uveitis status yielded an OR range of 1.61–2.45. Thus, despite the potential for missing data to change the magnitude of the temporal association, the relationship between eye and joint disease remained clinically significant.

Another limitation is the lack of data regarding time intervals between changes in systemic medication and arthritis assessment. Due to the retrospective nature of this study, ascertainment of the precise time of medication change with respect to the joint examination was difficult, due to variation in provider documentation, as well as the potential lag from the time of medication prescription to the time of actual administration. Our estimates present an average risk at a single time point, on any given medication regimen. Future studies will be needed to assess how the timing of medication changes modifies the association between arthritis and uveitis activity.

Our findings are consistent with previous studies that have alluded to a nonrandom association between arthritis and uveitis disease activity in JIA-U. The Inception Cohort of Newly diagnosed patients with Juvenile Idiopathic Arthritis identified moderate-to-high JIA disease activity as a predictor of uveitis reactivation over a 2-year follow-up period, after achieving sustained uveitis quiescence ≥6 months (18). Along the same lines, in an investigation of young adults with JIA evaluated at an average of 16 years after JIA diagnosis, patients with a history of uveitis ever in their disease course exhibited active arthritis at follow-up more often than those without a history of uveitis (risk ratio 1.38 [95% CI 1.06–1.78], $P < 0.044$); in fact, 4 of 5 individuals with extended oligoarticular JIA had ongoing active uveitis and active arthritis at follow-up (19). These important findings opened the door to further questions about the arthritis-uveitis relationship, for which there is emerging biologic evidence that the 2 share a common inflammatory mechanism (20,21). Tears of children with JIA-U exhibit cytokine signatures more closely associated with inflammatory arthritis than children with idiopathic chronic anterior uveitis (22).

Serum markers of uveitis activity have also been newly recognized, and here we propose that active joint disease may serve as a noninvasive biomarker for active eye disease, for which there is potential even into adulthood (4,19,23). These results hold promise for future evaluation of various JIA-U patient- and disease-related factors that may affect the association. Current American College of Rheumatology/Arthritis Foundation recommendations include standard, 3-month ophthalmic screening for JIA-U patients with controlled eye disease, irrespective of joint disease activity (24). Our current findings, elaborated by future identification of individual factors incurring different probabilities of uveitis activity with respect to arthritis activity, may inform future guidelines.

In conclusion, this novel work has demonstrated an independent temporal association between uveitis and arthritis activity in patients with JIA-U, with a nearly 2.5-fold higher odds of having active uveitis in the context of active arthritis. This important finding challenges the prior paradigm that uveitis runs a separate course from arthritis, and suggests that an arthritis flare should prompt urgent referral to the ophthalmologist in this population.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. Dr. Liebling 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. Liebling, Chang, Lerman.
Acquisition of data. Liebling, Mendoza, Moore, Vicioso.
Analysis and interpretation of data. Liebling, Faig, Chang, Lerman.

REFERENCES

1. Smith JA, Mackensen F, Sen HN, Leigh JF, Watkins AS, Pyatetsky D, et al. Epidemiology and course of disease in childhood uveitis. Ophthalmology 2009;116:1544–51.
2. Holland GN, Denove CS, Yu F. Chronic anterior uveitis in children: clinical characteristics and complications. Am J Ophthalmol 2009;147:667–78.
3. Heiligenhaus A, Niewerth M, Ganser G, Heinz C, Minden K, German Uveitis in Childhood Study Group. Prevalence and complications of uveitis in juvenile idiopathic arthritis in a population-based nation-wide study in Germany: suggested modification of the current screening guidelines. Rheumatology (Oxford) 2007;46:1015–9.
4. Zak M, Fledelius H, Pedersen FK. Ocular complications and visual outcome in juvenile chronic arthritis: a 25-year follow-up study. Acta Ophthalmol Scand 2003;81:211–5.
5. Saurenmann RK, Levin AV, Feldman BM, Laxer RM, Schneider R, Silverman ED. Risk factors for development of uveitis differ between girls and boys with juvenile idiopathic arthritis. Arthritis Rheum 2010; 62:1824–8.
6. Kotaniemi K, Kotaniemi A, Savolainen A. Uveitis as a marker of active arthritis in 372 patients with juvenile idiopathic seronegative oligoarthritis or polyarthritis. Clin Exp Rheumatol 2002;20:109–12.
7. Ezzahri M, Amine B, Rostom S, Rifay Y, Badri D, Mawani N, et al. The uveitis and its relationship with disease activity and quality of life in Moroccan children with juvenile idiopathic arthritis. Clin Rheumatol 2013;32:1387–91.
8. Rosenberg AM, Oen KG. The relationship between ocular and articular disease activity in children with juvenile rheumatoid arthritis and associated uveitis. Arthritis Rheum 1986;29:797–800.
9. Foeldvari I, Bidde M. Validation of the proposed ILAR classification criteria for juvenile idiopathic arthritis: International League of Associations for Rheumatology. J Rheumatol 2000;27:1069–72.
10. Jabs DA, Nussenblatt RB, Rosenbaum JT, Standardization of Uveitis Nomenclature Working Group. Standardization of uveitis nomenclature for reporting clinical data: results of the first international workshop. Am J Ophthalmol 2005;140:509–16.
11. Schmeling H, Horneff G. Etanercept and uveitis in patients with juvenile idiopathic arthritis. Rheumatology (Oxford) 2005;44:1008–11.
12. Zulian F, Martini G, Falcini F, Gerloni V, Zannin ME, Pinello L, et al. Early predictors of severe course of uveitis in oligoarticular juvenile idiopathic arthritis. J Rheumatol 2002;29:2446–53.
13. Edelsten C, Lee V, Bentley CR, Kanski JJ, Graham EM. An evaluation of baseline risk factors predicting severity in juvenile idiopathic arthritis associated uveitis and other chronic anterior uveitis in early childhood. Br J Ophthalmol 2002;86:51–6.
14. Kalinina Ayuso V, Ten Cate HA, van der Does P, Rothova A, de Boer JH. Male gender and poor visual outcome in uveitis associated with juvenile idiopathic arthritis. Am J Ophthalmol 2010;149:987–93.
15. Marino A, Weiss PF, Davidson SL, Lerman MA. Symptoms in noninfectious uveitis in a pediatric cohort: initial presentation versus recurrences. J AAPOS 2019;23:220.
16. Chang JH, McCluskey PJ, Wakefield D. Acute anterior uveitis and HLA-B27. Surv Ophthalmol 2005;50:364–88.
17. Clarke SL, Sen ES, Ramanan AV. Juvenile idiopathic arthritis-associated uveitis. Pediatr Rheumatol Online J 2016;14:27.
18. Heiligenhaus A, Klotsche J, Tappeiner C, Sengler C, Niewerth M, Liedmann I, et al. Predictive factors and biomarkers for the 2-year outcome of uveitis in juvenile idiopathic arthritis: data from the Inception Cohort of Newly diagnosed patients with Juvenile Idiopathic Arthritis (ICON-JIA) study. Rheumatology (Oxford) 2019;58: 975–86.
19. Kotaniemi K, Arkela-Kautiainen M, Haapasaari J, Leirisalo-Repo M. Uveitis in young adults with juvenile idiopathic arthritis: a clinical evaluation of 123 patients. Ann Rheum Dis 2005;64:871–4.
20. Kalinina Ayuso V, de Boer JH, Byers HL, Coulton GR, Dekkers J, de Visser L, et al. Intraocular biomarker identification in uveitis associated with juvenile idiopathic arthritis. Invest Ophthalmol Vis Sci 2013;54: 3709–20.
21. Haasnoot AM, Kuiper JJ, Hiddingh S, Schellekens PA, de Jager W, Imhof SM, et al. Ocular fluid analysis in children reveals interleukin-29/interferon-λ1 as a biomarker for juvenile idiopathic arthritis–associated uveitis. Arthritis Rheumatol 2016;68:1769–79.
22. Angeles-Han ST, Yeh S, Patel P, Duong D, Jenkins K, Rouster-Stevens KA, et al. Discovery of tear biomarkers in children with chronic non-infectious anterior uveitis: a pilot study. J Ophthalmic Inflamm Infect 2018;8:17.
23. Walscheid K, Heiligenhaus A, Holzinger D, Roth J, Heinz C, Tappeiner C, et al. Elevated S100A8/A9 and S100A12 serum levels reflect intraocular inflammation in juvenile idiopathic arthritis-associated uveitis: results from a pilot study. Invest Ophthalmol Vis Sci 2015;56:7653–60.
24. Angeles-Han ST, Ringold S, Beukelman T, Lovell D, Cuello CA, Becker ML, 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;71:703–16.

Arthritis Care & Research
Vol. 74, No. 3, March 2022, pp 355–363
DOI 10.1002/acr.24491

American College of Rheumatology
Empowering Rheumatology Professionals

Assessing the Validity and Reliability of the Effects of Youngsters' Eyesight on Quality of Life Questionnaire Among Children With Uveitis

Amy Cassedy,[1] Mekibib Altaye,[1] Jennifer Andringa,[1] Ashley M. Cooper,[2] Carolyn Drews-Botsch,[3] George Engelhard Jr.,[4] Theresa Hennard,[1] Gary N. Holland,[5] Kirsten Jenkins,[6] Scott R. Lambert,[7] Jessi Lipscomb,[1] Courtney McCracken,[8] Deborah K. McCurdy,[5] Joseph McDonald,[1] Najima Mwase,[1] Sampath Prahalad,[9] Erin Stahl,[2] Virginia Miraldi Utz,[1] A. Adrienne Walker,[10] Steven Yeh,[8] and Sheila T. Angeles-Han[1]

Objective. The Effects of Youngsters' Eyesight on Quality of Life (EYE-Q) questionnaire measures vision-related functioning (VRF) and vision-related quality of life (VRQoL) in children with uveitis. Our aim was to revise the alpha version of the EYE-Q to refine VRF and VRQoL subscales and to assess the validity of the EYE-Q.

Methods. Children with juvenile idiopathic arthritis (JIA), JIA-associated uveitis, and other noninfectious uveitis were enrolled. Patients and parents completed the EYE-Q, Pediatric Quality of Life Inventory (overall quality of life), and Childhood Health Assessment Questionnaire (physical functioning). The development site completed the alpha version of the EYE-Q, and the composite sites completed the beta version. We compared item-subscale correlations, internal consistency, and construct and discriminant validity among the different versions.

Results. Of the 644 patients enrolled, 61.6% completed the alpha version, and 38.4% the beta version of the EYE-Q. Mean ± SD patient age was 11.1 ± 4.2 years, and 70% were female. Fewer White patients (73.5%) completed the alpha version compared to the beta version (86.2%; $P < 0.001$). With the exception of patient-reported VRF, both versions had similar item-subscale correlations. Version comparisons on scale internal consistencies indicated significant differences for parent- and patient-reported VRF, but each scale had a Cronbach's α of >0.80 beta. When data were combined, the EYE-Q showed significant differences between JIA-only and uveitis patients on all parent and patient scores, except for patient-reported VRF.

Conclusion. The EYE-Q appears to be a valid measure of VRF and VRQoL in pediatric uveitis. Our results suggest it may be used as an outcome measure in multicenter pediatric uveitis studies.

INTRODUCTION

Adequate vision is crucial for a child to participate in school and home activities. Vision loss affects social, emotional, mental, and physical well-being, with long-standing effects into adulthood. Pediatric chronic noninfectious uveitis is an inflammatory ocular disease that can lead to sight-threatening complications (1). It can occur in isolation, as in idiopathic uveitis, or be

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

Dr. Angeles-Han's work was supported by the NIH/National Eye Institute (K23-EY-021760 and R01-EY-030521), NIH/National Center for Advancing Translational Sciences (UL1-TR-000077), the Rheumatology Research Foundation, the Childhood Arthritis and Rheumatology Research Alliance, the Arthritis Foundation, the Weill Cornell Medical College Clinical and Translational Science Center (UL1-RR-024996), the Emory Egleston Children's Research Center, and the Knights Templar Eye Foundation.

[1]Amy Cassedy, PhD, Mekibib Altaye, PhD, Jennifer Andringa, BA, Theresa Hennard, BS, Jessi Lipscomb, BS, Joseph McDonald, MD, Najima Mwase, MD, Virginia Miraldi Utz, MD, Sheila T. Angeles-Han, MD, MSc: Cincinnati Children's Hospital Medical Center and University of Cincinnati, Cincinnati, Ohio; [2]Ashley M. Cooper, MD, Erin Stahl, MD: Children's Mercy Hospital, Kansas City, Missouri; [3]Carolyn Drews-Botsch, PhD: Rollins School of Public Health, Emory University, Atlanta, Georgia; [4]George Engelhard Jr., PhD: University of Georgia, Athens; [5]Gary N. Holland, MD, Deborah K. McCurdy, MD: David Geffen School of Medicine at University of California, Los Angeles; [6]Kirsten Jenkins, BS: Children's Healthcare of Atlanta, Atlanta, Georgia; [7]Scott R. Lambert, MD: Stanford University, Stanford, California; [8]Courtney McCracken, PhD, Steven Yeh, MD: Emory University School of Medicine, Atlanta, Georgia; [9]Sampath Prahalad, MD, MS: Children's Healthcare of Atlanta and Emory University School of Medicine, Atlanta, Georgia; [10]A. Adrienne Walker, PhD: Georgia Department of Education, Atlanta.

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

Address correspondence to Sheila T. Angeles-Han, MD, MSc, Division of Rheumatology, Cincinnati Children's Hospital Medical Center, 3333 Burnett Avenue, Cincinnati, OH 45229. Email: sheila.angeles-han@cchmc.org.

Submitted for publication May 1, 2020; accepted in revised form October 15, 2020.

SIGNIFICANCE & INNOVATIONS

- Uveitis can lead to vision-threatening complications in children.
- Few studies examine the impact of uveitis on the quality of life of children and their families due to the lack of a uveitis-specific questionnaire.
- The Effects of Youngsters' Eyesight on Quality of Life (EYE-Q) questionnaire is the only measure of vision-related functioning and quality of life in pediatric uveitis.
- This is the first multicenter validation study of the EYE-Q that supports its use in future studies.

associated with systemic conditions, commonly juvenile idiopathic arthritis (JIA) (2–5). Few studies examine the impact of uveitis on quality of life (QoL) and daily functioning in children (6–10).

Understanding the effects of disease and complex treatment regimens on a child's QoL and vision-related functioning (VRF) can optimize disease management. The World Health Organization defines QoL as an "individual's perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards, and concerns" (11). QoL can be affected by visual impairment, ocular complications, treatment with eye drops or immunosuppressive medications, and frequent subspecialists and laboratory blood draw visits. VRF is a measure of an individual's ability to perform activities of daily living that rely on visual acuity, peripheral vision, contrast, and color vision (12). Visual impairment can lead to difficulty in performing daily tasks that rely on vision. Thus, VRF may be assessed by measuring the degree of visual impairment that one experiences while performing daily tasks. Outcomes of children with JIA-associated uveitis are typically assessed by the ophthalmic examination and questionnaires on physical function and general QoL (13,14). Assessment by ophthalmic examination and questionnaires may result in an underestimation of the effects of uveitis (7,15–17).

In adult uveitis studies, the National Eye Institute Visual Functioning Questionnaire is used to measure outcomes (18–20). Previously, no pediatric vision questionnaires focused on uveitis (21–26). Thus, we developed the Effects of Youngsters' Eyesight on Quality of Life (EYE-Q) questionnaire, the only tool that assesses VRF and vision-related QoL (VRQoL) in this population (27–29). Earlier single-center studies support the validity and the reliability of the EYE-Q. Interviews and focus groups with children, rheumatologists, and ophthalmologists showed face and content validity. Criterion- and construct-related validity were demonstrated as the EYE-Q differentiated children with an underlying uveitis diagnosis, by severity of visual impairment and laterality of eye involvement and was associated with visual acuity and contrast sensitivity. Test–retest reliability has also been established (29). However, unpublished psychometrics on the alpha version of the EYE-Q reported the need to refine the VRF items (see Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24491/abstract), and additional focus groups suggested expanding VRQoL items.

Our aims were to revise the alpha version of the EYE-Q to refine the VRF and VRQoL subscales, to examine the internal consistency of the alpha and beta versions of the EYE-Q subscales, and to assess the construct and discriminant validity of the EYE-Q in a large, multicenter cohort.

PATIENTS AND METHODS

Study design and analysis population. This multicenter study was approved by all institutional review boards, which conformed to the US Health Insurance Portability and Privacy Act requirements. Informed consent/assent was obtained from each participant. The development site testing the alpha version of the EYE-Q was Emory University (EU) in Atlanta, Georgia. The composite sites testing the beta version included Cincinnati Children's Hospital Medical Center (CCHMC) in Cincinnati, Ohio, Children's Mercy Hospital (CMH) in Kansas City, Missouri, and University of California, Los Angeles (UCLA) Mattel Children's Hospital located in Los Angeles, California.

Subjects. Children diagnosed with JIA without uveitis, JIA with uveitis, or noninfectious uveitis of any cause were invited to participate from November 30, 2011 to September 27, 2019. Potential subjects were approached consecutively during their rheumatology or ophthalmology appointments. The cohorts at EU and CCHMC were enrolled in a prospective epidemiology study, and only baseline data were used. Inclusion criteria were: 1) a diagnosis of JIA per International League of Associations for Rheumatology classification (30) with or without uveitis, or noninfectious uveitis of any cause without JIA, 2) age 5–18 years at study visit, and 3) ability to speak English. Exclusion criteria were significant comorbidity unrelated to uveitis (i.e., sickle cell anemia) affecting QoL and function, major development disorders (i.e., cerebral palsy, mental retardation [one patient had Down syndrome but was considered highly functional and was able to complete the questionnaires with the aid of his mother]), inability to speak English, or inability to complete the questionnaire.

Data collection. For disease characteristics, we conducted systematic medical record reviews and administered parent- and patient-based questionnaires at study visits. Data included age, sex, self-described race and ethnicity, JIA subtype, uveitis diagnosis, and ocular examination (best corrected visual acuity [BCVA]). We recorded data from the most recent ophthalmology visit for BCVA in both eyes for CMH and UCLA. Latent variable BCVA data were recorded for all ophthalmology visits for EU and CCHMC. For patient-reported outcome measures, patients (if age appropriate) and parents/guardians completed

patient-reported outcome measures over approximately 20 minutes.

Information on visual functioning and VRQoL was gathered using the EYE-Q, a measure of VRF and VRQoL in children and adolescents ages 5–18 years. It is written at a 3rd-grade reading level and completed by patients and parent proxies. Self-reports are available for patients who are ≥8 years old and able to read at a 3rd-grade level. For patients at a reading level of <3rd grade, the questionnaire was administered by office staff. Paper- and electronic-based formats in large print are available.

The EYE-Q alpha version (see Supplementary Appendix B, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24491/abstract) was completed at the development site. It contained 25 items (19 measured VRF and 6 measured VRQoL) using a 5-point response system with 1 = never or not hard, up to 5 = always/cannot do. Test–retest reliability showed a correlation of 0.75 over a 10-day period, showing no significant changes over that period. Psychometrics for the alpha version using the Rasch partial credit model are reported in Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24491/abstract.

The EYE-Q beta version (see Supplementary Appendix C, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24491/abstract) was created as a result of Rasch analysis and additional focus groups' suggestions to bolster the VRQoL section, with revision of the alpha version. The beta version was completed at composite sites using a 3-point response scale. It contains 24 items, where 16 measure VRF and ophthalmic symptoms (near and far vision, color/night vision, and photosensitivity), and 8 measure VRQoL (feelings about use of medications, about participating in social activities related to vision, and about having uveitis). VRQoL questions were added to identify areas of uveitis that impact on children's socioemotional perceptions. The 3-point response format measures the difficulty in completing VRF tasks (0 = never hard, 1 = sometimes hard, 2 = always hard), and how true a QoL statement is (0 = never, 1 = sometimes, 2 = always). Participants have the option to mark that an item does not apply to them ("does not apply"). A visual analog scale was not included, to remain sensitive to visual impairment.

The scoring system is based on the following process: 1) reverse score each item; 2) sum the items; 3) subtract 2 points for each aid used, up to 4 aids (i.e., using large print books, magnifying glasses, special lighting, or other aids, not including glasses or contact lenses [points were deducted for visual aids because they were viewed as negatively impacting VRF and VRQoL]); 4) divide the raw score by the number of items answered; 5) divide this average score by 2 to create a proportion; and 6) multiply by 100 to get a percent score. Scores in the negative range were excluded since the visual impairment likely prevented the ability to complete VRF tasks. This standardized score allows for easy interpretation and is similar to generic QoL questionnaires such as the Pediatric Quality of Life Inventory (PedsQL). Scores range from 0 to 100, with higher scores indicating better total vision, VRF, and VRQoL. A global assessment of the child's eyesight was also completed by the patient and/or parent. Overall QoL was measured with the PedsQL 4.0, a valid 23-item measure of general health-related QoL in children age 2–18 years that includes 2 summary scores for physical and socioemotional states (31). Scores range from 0 to 100, and higher scores indicate better QoL.

Physical functioning was assessed using the Childhood Health Assessment Questionnaire (C-HAQ), a valid arthritis-specific measure that evaluates functional disability and consists of 20 questions in 8 functional components: 1) dressing and grooming, 2) arising, 3) eating, 4) walking, 5) hygiene, 6) reach, 7) grip, and 8) activities (32). There are 3 parameters within each area: 1) difficulty in performing daily functions, 2) use of special aids or devices, and 3) activities that require assistance from another person. Scores range from 0 to 3, and higher scores indicate worse physical function.

Study variables. Predictor variables were 1) the uveitis group (the presence of JIA-associated uveitis or uveitis of any cause) or the absence of uveitis (JIA without uveitis) and 2) the study site (development or composite). Demographic variables included age, sex, race dichotomized as White versus non-White, and ethnicity: Hispanic versus Non-Hispanic. The clinical variable of BCVA was included to summarize site differences but was not included in the analysis due to missing data. For many patients, eye examinations (where visual acuity was derived) were conducted on different dates than when the EYE-Q questionnaire was administered.

Statistical analysis. Descriptive statistics such as frequency counts and percentages for dichotomous or polytomous variables, and means ± SDs for continuous variables, were used to summarize demographic information, clinical variables, assessment instruments, and outcome measures. To examine the internal consistencies of the alpha and beta version of the EYE-Q subscales (VRF and VRQoL), we completed 2 analyses. The first was item-to-scale correlations to examine the strength of the relationship for each item to each VRF and VRQoL subscale score. An average of the subscale correlations for both the alpha and beta versions was created, and alpha versus beta differences were tested using a Student's *t*-test. Following Nunnally, correlation values were interpreted as weak (≤0.20), fair (0.21–0.40), moderate (0.41–0.60), good (0.61–0.80), and excellent (>0.80) (33). The second analysis was an examination of the internal consistencies of the alpha and beta versions of the EYE-Q VRF and VRQoL subscale scores, conducted by calculating a Cronbach's alpha on the EYE-Q subscale scores for the alpha and then for the beta versions of the EYE-Q. Tavakol and Dennick report the

Table 1. Demographic and clinical characteristics of patients at development (EU) and composite (CCHMC, CMH, UCLA) sites*

	Development (n = 397)	Composite (n = 247)
Age, mean ± SD	11.0 ± 4.5	11.4 ± 3.5
Sex		
Female	270 (68.5)	178 (72.4)
Male	124 (31.5)	68 (27.6)
Race†		
Non-White	105 (26.5)	34 (13.8)
White	292 (73.5)	213 (86.2)
Uveitis group‡		
JIA only	276 (70.1)	128 (51.6)
Uveitis	118 (29.9)	120 (48.4)
Highest visual acuity (worse eye)§		
20/20 to 20/40	116 (61.5)	141 (75.0)
20/50 or worse	72 (38.3)	47 (25.0)
PedsQL, mean ± SD		
Parent total¶	77.6 ± 18.9	83.3 ± 15.4
Parent physical#	75.7 ± 23.8	81.8 ± 19.8
Parent psychosocial**	78.6 ± 18.8	84.0 ± 15.5
Patient total	78.0 ± 17.7	80.9 ± 16.0
Patient physical	78.0 ± 21.5	81.1 ± 19.4
Patient psychosocial	78.0 ± 18.3	80.9 ± 17.0

* Values are the number (%) unless indicated otherwise. CCHMC = Cincinnati Children's Hospital Medical Center; CMH = Children's Mercy Hospital; EU = Emory University; JIA = juvenile idiopathic arthritis; PedsQL = Pediatric Quality of Life Inventory version 4.0; UCLA = University of California, Los Angeles.
† $\chi^2 = 14.5$, $P < 0.001$.
‡ $\chi^2 = 22.2$, $P < 0.001$.
§ 53% of patients were missing data from the development site and 24% were missing data from the composite sites. $\chi^2 = 7.68$, $P = 0.006$.
¶ $t = -3.93$, $P < 0.001$.
# $t = -3.72$, $P < 0.001$.
** $t = -3.35$, $P < 0.001$.

range for Cronbach's α = 0.7–0.9 to be acceptable (34). Test–retest reliability for the beta version was conducted using Pearson's correlation, and paired *t*-tests were used to examine test–retest reliability for the beta versions.

To test for potential site differences in the EYE-Q total, VRF, and VRQoL scores, Student's *t*-tests were calculated. We also tested for site differences on the PedsQL and C-HAQ, to ascertain whether significant differences in EYE-Q scores were due to the uniqueness of the development site and not to the changes from the alpha to the beta version of the EYE-Q.

To assess the construct validity of the EYE-Q, we used the PedsQL and the C-HAQ. We hypothesized that when combining all patients in the development and composite sites, there would be moderate correlations between the PedsQL and EYE-Q total scores, and fair-to-moderate correlations among the PedsQL subscale and the EYE-Q subscale scores. We also hypothesized a poor-to-fair correlation among EYE-Q total and subscale and C-HAQ scores. Pearson's correlation coefficients were calculated between the EYE-Q scores and the PedsQL total, physical health score, and psychosocial health summary, and C-HAQ total scores. Correlations were interpreted as ≤0.20 poor, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 good, and ≥0.81 excellent (33).

To assess the discriminant validity of the EYE-Q, we also used the PedsQL and the C-HAQ. While generic versions of the PedsQL measure overall QoL, and the C-HAQ measures physical limitations, a disease-specific questionnaire is more sensitive to specific conditions that can discriminate between groups within a disease category. To examine group differences between patients with and without a uveitis diagnosis we used general linear models, controlling for demographic covariates as well as sites (composite versus development). We hypothesized that uveitis patients would have lower EYE-Q scores compared to patients without uveitis, but there would be no significant group differences on the PedsQL or C-HAQ.

The Cohen's *d* statistic was used to demonstrate discriminant validity of the EYE-Q by estimating effect sizes between groups. The following scale was used to interpret the Cohen's *d*: 0.20 = weak, 0.50 = moderate, and 0.80 = strong (35). An alpha of *P* less than 0.05 was used to determine statistical significance. All data, except where open-source programing was employed, were analyzed using SAS software, version 9.4.

RESULTS

Patient characteristics. Of 644 patients, 397 (61.6%) were enrolled in the development site (EU) and 247 (38.4%) in

Table 2. Item-to-scale correlations and Cronbach's alpha for EYE-Q total, vision-related functioning (VRF) and vision-related quality of life (VRQoL) scores by version*

	Item-to-scale correlation, mean ± SD		Cronbach's alpha, coefficient (95% CI)	
Version	Alpha	Beta	Alpha	Beta
Parent responses				
VRF	0.72 ± 0.06	0.68 ± 0.09	0.96 (0.95–0.96)	0.94 (0.92–0.95)
VRQoL	0.52 ± 0.15	0.51 ± 0.14	0.77 (0.72–0.80)	0.80 (0.75–0.84)
Patient responses				
VRF	0.62 ± 0.08	0.51 ± 0.10	0.93 (0.91–0.94)	0.87 (0.81–0.89)
VRQoL	0.49 ± 0.14	0.46 ± 0.13	0.75 (0.70–0.80)	0.76 (0.70–0.81)

* Parent alpha version n = 332; patient alpha version n = 223; parent beta version n = 201; patient beta version n = 179. 95% CI = 95% confidence interval; EYE-Q = Effects of Youngsters' Eyesight on Quality of Life.

Table 3. Site comparisons on the EYE-Q, PedsQL, and C-HAQ scores*

	Alpha version site	Beta version sites	*t* score	*P*
Parent responses				
EYE-Q				
Total	87.0 ± 13.2	89.2 ± 14.7	–1.81	0.071
VRF	91.4 ± 13.3	90.7 ± 16.6	0.49	0.623
VRQoL	70.7 ± 22.6	82.6 ± 19.0	–6.00	<0.001
PedsQL				
Total	77.6 ± 18.9	83.3 ± 15.4	–3.93	<0.001
Physical	75.7 ± 23.8	81.8 ± 19.8	–3.35	0.001
Psychosocial	78.6 ± 18.8	84.0 ± 15.5	–3.72	<0.001
C-HAQ total	0.4 ± 0.5	0.3 ± 0.4	2.64	0.009
Patient responses				
EYE-Q				
Total	86.5 ± 12.6	86.2 ± 13.5	0.23	0.820
VRF	89.7 ± 12.3	86.8 ± 14.2	2.15	0.032
VRQoL	74.4 ± 21.6	81.9 ± 19.2	–3.36	<0.001
PedsQL				
Total	78 ± 17.7	80.9 ± 16.0	–1.95	0.051
Physical	78.0 ± 21.5	81.1 ± 19.4	–1.72	0.085
Psychosocial	78.0 ± 18.3	80.9 ± 17.0	–1.80	0.072
C-HAQ total	0.4 ± 0.5	0.3 ± 0.4	3.20	0.004

* Values are the mean ± SD unless indicated otherwise. C-HAQ = Childhood Health Assessment Questionnaire; EYE-Q = Effects of Youngsters' Eyesight on Quality of Life; PedsQL = Pediatric Quality of Life Inventory version 4.0; VRF = vision-related functioning; VRQoL = vision-related quality of life.

the composite sites of CCHMC (n = 210 [85%]), CMH (n = 16 [6.5%]), and UCLA (n = 21 [8.5%]) (Table 1). Data from the composite sites were combined for the validity analyses. Mean ± SD patient age was 11.1 ± 4.2 years, and 70% (n = 448) were females. Of 505 White patients (78.5%), the development site had significantly fewer White patients (73.5%) compared to the composite sites (86.2%; $\chi^2 = 14.5$, $P < 0.001$). The development site (n = 118 [29.9%]) had significantly fewer uveitis patients than the composite sites (n = 120 [48.4%]; $\chi^2 = 22.2$, $P < 0.001$), but had significantly more visual impairment (worse eye) during any examination (BCVA 20/50 or worse) compared to the composite sites (72 [38.3%] versus 47 [25.0%]; $\chi^2 = 7.68$, $P = 0.006$).

Comparison of the item-level to subscale correlation of the EYE-Q between the alpha (development site) and beta (composite sites) versions. Using the item-to-scale correlations, we tested for differences between the alpha and beta versions. With the exception of patient-reported VRF, both alpha and beta versions had similar item-to-scale correlations, with all of the mean item-to-scale correlations indicating moderate-to-strong associations (Table 2). While the item-to-scale correlations for patient-reported VRF were significantly higher for the alpha version (mean ± SD 0.62 ± 0.08) compared to the beta version (mean ± SD 0.51 ± 0.10) (*t* score = –3.62, $P = 0.001$), the higher mean correlation for the alpha version was aided by 2 items with a strong correlation of 0.68 ("How hard is it for you to see the steps so that you do not fall when going down the stairs?" and "How hard is it for you to see the steps so that you do not trip when going up the stairs?").

Comparison of internal consistency of the EYE-Q subscales scores between the alpha and beta versions. To determine whether the internal consistencies of the alpha and beta versions were similar, we compared the Cronbach's alpha for each subscale. The alpha and beta versions of the VRF had a Cronbach's $\alpha = 0.87$ or higher for both the parent- and patient-reported scores. (Table 2). The coefficients for VRQoL ranged between 0.75 and 0.80. A total of 48 parents and patients completed the test–retest task. Test–retest reliability showed a correlation of 0.79 for the parents and 0.83 for patients over a 10-day period, with paired *t*-test on the 2 sets of EYE-Q scores showing no significant changes.

Table 4. Correlations between EYE-Q and PedsQL/C-HAQ*

	EYE-Q total	EYE-Q VRF	EYE-Q VRQoL
Parent responses			
PedsQL			
Total	0.42	0.41	0.30
Physical	0.27	0.27	0.18
Psychosocial	0.47	0.44	0.33
C-HAQ total	–0.27	–0.25	–0.23
Patient responses			
PedsQL			
Total	0.54	0.52	0.41
Physical	0.37	0.37	0.27
Psychosocial	0.57	0.55	0.45
C-HAQ total	–0.41	–0.42	–0.32

* All correlations were significant at the $P < 0.001$ level. C-HAQ = Childhood Health Assessment Questionnaire; EYE-Q = Effects of Youngsters' Eyesight on Quality of Life; PedsQL = Pediatric Quality of Life Inventory version 4.0; VRF = vision-related functioning, VRQoL = vision-related quality of life.

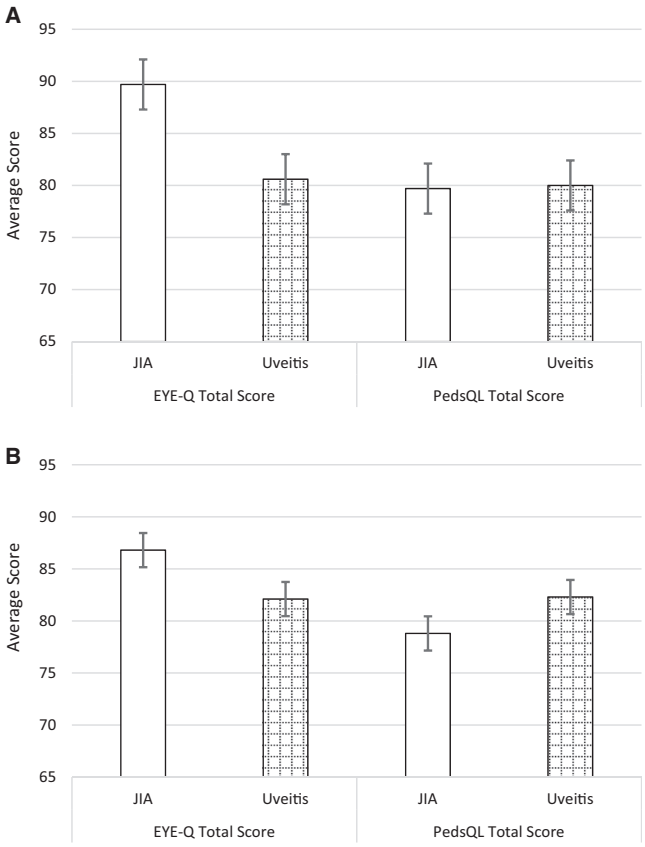


Figure 1. **A**, Parent-reported scores. **B**, Patient-reported scores. Whiskers indicate standard error. EYE-Q = Effects of Youngsters' Eyesight on Quality of Life; JIA = juvenile idiopathic arthritis; PedsQL = Pediatric Quality of Life Inventory.

Comparisons of the EYE-Q versions. While there were significant differences between the parent-report VRQoL scores (mean ± SD alpha version 70.7 ± 22.6 versus beta version 82.6 ± 19.0; $t = -6.0$, $P < 0.001$), significant alpha versus beta version differences were also found in parent-reported PedsQL and C-HAQ scores (Table 3). Results from the patient-reported measures were mixed. Patients completing the alpha version reported better VRF compared to patients taking the beta version (mean ± SD 89.7 ± 12.3 versus 86.8 ± 14.2; $t = 2.15$, $P = 0.03$), but worse VRQoL (mean ± SD 74.4 ± 21.6 versus 81.9 ± 19.2; $t = -3.36$, $P < 0.001$). Additionally, patients completing the alpha version of the EYE-Q reported worse physical functioning (mean ± SD 0.40 ± 0.5) compared to C-HAQ scores for beta version patients (mean ± SD 0.3 ± 0.4; $t = 3.2$, $P = 0.004$).

Construct validity of the EYE-Q. We evaluated the Pearson's correlations between the EYE-Q and PedsQL, with the C-HAQ. The associations between parent-reported EYE-Q scores and the PedsQL and C-HAQ ranged from a low correlation of 0.18 between VRQoL and PedsQL physical health, to a moderate correlation of 0.47 between EYE-Q total and PedsQL (Table 4). Patient reports indicated a stronger association between EYE-Q scores and the PedsQL and C-HAQ, with the correlation between EYE-Q total and PedsQL psychosocial summary scores being the highest (r = 0.57).

Viability of the EYE-Q to assess the overall level of uveitis disease complexity (discriminant validity). Figures 1A and 1B show group differences on the EYE-Q and PedsQL. When data from the alpha and beta versions were combined, the EYE-Q scores showed differences across disease severity groups (JIA without uveitis group, or uveitis group) in the expected direction. On parent and patient scores, except for patient-reported VRF, the uveitis group reported lower EYE-Q

Table 5. Mean scores for EYE-Q and PedsQL by uveitis groups*

	JIA only	Uveitis	*t* score	*P*	Cohen's *d*
Parent responses	n = 332	n = 201			
EYE-Q					
Total	89.7 ± 0.9	80.6 ± 1.0	7.71	<0.001	0.69
VRF	91.4 ± 1.0	84.6 ± 1.1	5.33	<0.001	0.47
VRQoL	82.0 ± 1.6	67.6 ± 1.5	7.25	<0.001	0.69
PedsQL					
Total	78.3 ± 1.2	81.4 ± 1.3	−1.91	0.057	0.16
Psychosocial	79.7 ± 1.2	80.0 ± 1.3	−0.21	0.836	0.02
Physical	75.5 ± 1.5	83.9 ± 1.7	−4.14	<0.001	0.35
Patient responses	n = 223	n = 179			
EYE-Q					
Total	86.8 ± 1.1	82.1 ± 1.1	4.28	<0.001	0.46
VRF	87.8 ± 1.1	85.3 ± 1.1	1.76	0.079	0.18
VRQoL	81.1 ± 1.9	71.1 ± 1.8	4.28	<0.001	0.46
PedsQL					
Total	78.8 ± 1.2	82.3 ± 1.3	−2.19	0.029	0.19
Psychosocial	79.1 ± 1.3	80.7 ± 1.4	−0.96	0.337	0.09
Physical	78.3 ± 1.4	85.4 ± 1.5	−3.78	<0.001	0.34

* Values are the mean ± SD unless indicated otherwise. Models controlled for sex, age, race, and composite site. EYE-Q = Effects of Youngsters' Eyesight on Quality of Life; JIA = juvenile idiopathic arthritis; PedsQL = Pediatric Quality of Life Inventory version 4.0; VRF = vision-related functioning, VRQoL = vision-related quality of life.

scores than the JIA without uveitis group, with corresponding moderate effect sizes ranging between 0.46 and 0.69 (Table 5). Results for the parent- and patient-reported PedsQL demonstrated weak effect sizes for both total and psychosocial subscale (0.02–0.19) scores. Parents and patients in the JIA without uveitis group reported lower physical health summary scores compared to the uveitis group (Cohen's *d*: 0.35 and 0.34, respectively), which may be related to underlying arthritis.

DISCUSSION

Our results suggest that the alpha (original) and beta (revised) versions of the EYE-Q are similar, and that the current version of the EYE-Q is a valid measure of VRF and VRQoL in children with noninfectious uveitis. All mean item-to-scale correlations were in the moderate range, with only the patient-reported VRF scale showing significant differences between the alpha and beta versions. Additionally, internal consistency measures for VRF scores were above 0.90 in both the alpha and beta versions. The VRQoL also had an acceptable internal consistency, with a Cronbach's α of no lower than 0.75.

Differences in racial composition between the patients/parents taking the alpha version and the beta version may have depressed the EYE-Q scores, especially the VRQoL subscale scores. Non-White patients rate their VRQoL lower than White patients. Additionally, parents and patients completing the alpha version of the EYE-Q reported lower scores on the PedsQL. Therefore, their EYE-Q VRQoL scores might also be expected to be lower. While these results are consistent with prior literature (36), we must be aware that QoL scores, including vision-related scores, may also be detecting problems associated with larger societal issues as well as health- or vision-related QoL. Further, we previously showed at the development site that Black children with uveitis had worse BCVA and ocular complications compared to White children (37).

Given that the EYE-Q was able to detect differences in VRF and VRQoL, this finding supports the discriminant validity of the EYE-Q, and the need for a disease-specific measure. The need for a disease-specific measure is endorsed further when results from the EYE-Q are compared to the PedsQL, similar to a study in adults with JIA and uveitis by Haasnoot et al where uveitis had a negative effect on VRQoL, even with good visual acuity (17). Additionally, general QoL measures failed to distinguish between patients with and without uveitis.

No standardized protocol to assess uveitis outcomes in children exists, although efforts are underway to validate outcome measures (38,39). Usual measures include: 1) ophthalmic examination to measure BCVA for visual functioning, the presence of cells, or protein flare in the anterior chamber per Standardization of Uveitis Nomenclature criteria (40) for activity, and ocular complications for damage, and 2) general QoL patient-reported outcome measures. Although questionnaires such as the Juvenile Arthritis Multidimensional Assessment Report assess the impact of JIA, data are lacking on the effects of uveitis on function and health-related QoL in children (9,41).

We previously showed that the PedsQL may not accurately measure the impact of uveitis (29). Further, current pediatric vision questionnaires may not be ideal because they are not culturally optimal, not uveitis-focused, use parent-proxy reports that may not reflect the child's perspective, or focus on function but not on QoL (21–25). Although the National Eye Institute Visual Functioning Questionnaire shows clinically meaningful improvement in visual functioning in adults with uveitis, adult-based tools are not valid for children (18), as they contain irrelevant items related to cooking, driving, and working. Pediatric uveitis–specific questions may reveal areas of vision that are not quantified by general measures or the ophthalmic examination.

Expert groups acknowledge the need for a holistic approach to uveitis outcomes that includes patient-reported outcome measures combined with clinical examination findings (7,39). To date, pediatric uveitis clinical trials do not incorporate uveitis-specific patient-reported outcome measures. Inclusion of ophthalmic examinations and patient-reported outcome measures that reflect changes in the quality of vision, despite normal BCVA or inactive disease, will provide new information on factors to be targeted to improve health-related QoL. Adding the EYE-Q to studies that measure pediatric visual outcomes may provide ways to better support children and families. A potential exists for use in clinical trials as a treatment outcome measure. Work is underway to investigate the use of the EYE-Q as a measure of change in ocular disease status. Early detection of a change in disease status leading to timely intervention can prevent lifelong visual disability.

Ophthalmic examinations did not always occur on the same date as questionnaire completion, so that mapping examination information onto the EYE-Q scores was difficult. This problem could impact VRF scores, as we previously showed that visual impairment measured by BCVA impacts VRF. Some medical records missed ophthalmic examination data. We found that VRQoL was related to the racial composition of the respondent. While health-related QoL is often affected by socioeconomic status (42), consideration of how race might affect VRQoL is necessary. The VRQoL was an indicator of how well, or poorly, a patient's QoL was impacted by uveitis, but within each racial group. As in general and other disease populations, children of different races and ethnicities may have varied health-related QoL. Studies on the impact of uveitis on VRQoL in racially and ethnically diverse populations are needed. While racial composition can be used as an indicator of socioeconomic status, the inclusion of caregiver educational attainment status or family income would have strengthened these control variables. Due to enrollment requirements of a diagnosis of uveitis at CMH and UCLA, more patients with uveitis were recruited at the composite sites. Finally, there are potential treatment or physician practice

differences between the development and composite sites. At CMH and UCLA, recruitment was initially focused on patients with JIA-associated uveitis with a visual acuity worse than 20/50. The BCVA posed a barrier to recruitment, so we expanded recruitment to include normal BCVA.

In conclusion, for pediatric uveitis, clinical measures are primarily used to evaluate visual outcomes, hence the need for a comprehensive evaluation that takes into account child and parent perspectives. The EYE-Q is a reliable and valid instrument that could complement ophthalmic examinations and global health measures and augment the assessment of the impact of disease and utility of treatment.

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. Angeles-Han 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. Cassedy, Altaye, Drews-Botsch, Engelhard, Holland, Lambert, McCracken, Prahalad, Walker, Yeh, Angeles-Han.

Acquisition of data. Cooper, Hennard, Holland, Jenkins, McCurdy, Mwase, Stahl, Miraldi Utz, Angeles-Han.

Analysis and interpretation of data. Cassedy, Altaye, Andringa, Drews-Botsch, Engelhard, Lipscomb, McCracken, McDonald, Walker, Angeles-Han.

REFERENCES

1. Thorne JE, Woreta F, Kedhar SR, Dunn JP, Jabs DA. Juvenile idiopathic arthritis-associated uveitis: incidence of ocular complications and visual acuity loss. Am J Ophthalmol 2007;143:840–6.
2. Ravelli A, Martini A. Juvenile idiopathic arthritis. Lancet 2007;369: 767–78.
3. Cassidy JT, Petty RE, Laxer RM, Lindsley CB. Textbook of pediatric rheumatology. Philadelphia: Elsevier Saunders; 2005.
4. Kump LI, Cervantes-Castaneda RA, Androudi SN, Foster CS. Analysis of pediatric uveitis cases at a tertiary referral center. Ophthalmology 2005;112:1287–92.
5. Kanski JJ, Shun-Shin GA. Systemic uveitis syndromes in childhood: an analysis of 340 cases. Ophthalmology 1984;91:1247–52.
6. Maca SM, Amirian A, Prause C, Gruber K, Mejdoubi L, Barisani-Asenbauer T. Understanding the impact of uveitis on health-related quality of life in adolescents. Acta Ophthalmol 2013;91:e219–24.
7. Solebo AL, Barry RJ, Keane PA, Rahi JS, Denniston AK. Under-utilisation of reproducible, child appropriate or patient reported outcome measures in childhood uveitis interventional research. Orphanet J Rare Dis 2019;14:125.
8. Ezzahri M, Amine B, Rostom S, Rifay Y, Badri D, Mawani N, et al. The uveitis and its relationship with disease activity and quality of life in Moroccan children with juvenile idiopathic arthritis. Clin Rheumatol 2013;32:1387–91.
9. Sestan M, Grguric D, Sedmak M, Frkovic M, Kifer N, Grubic M, et al. Quality of life in children suffering from juvenile idiopathic arthritis-associated uveitis. Rheumatol Int 2020;40:1117–21.
10. Taha R, Papadopoulou M, Zetterberg M, Oskarsdottir S, Gronlund MA. Visual function and quality of life in a cohort of Swedish children with juvenile idiopathic arthritis. Clin Ophthalmol 2019;13:2081–91.
11. World Health Organization Quality of Life Group. The development of the World Health Organization Quality of Life assessment instrument (the WHOQOL). Berlin: Springer-Verlag; 1994.
12. Frick KD, Drye LT, Kempen JH, Dunn JP, Holland GN, Latkany P, et al. Associations among visual acuity and vision- and health-related quality of life among patients in the Multicenter Uveitis Steroid Treatment trial. Invest Ophthalmol Vis Sci 2012;53:1169–76.
13. Moorthy LN, Peterson MG, Harrison MJ, Onel KB, Lehman TJ. Physical function assessment tools in pediatric rheumatology. Pediatr Rheumatol Online J 2008;6:9.
14. Gutierrez-Suarez R, Pistorio A, Cespedes Cruz A, Norambuena X, Flato B, Rumba I, et al. Health-related quality of life of patients with juvenile idiopathic arthritis coming from 3 different geographic areas: the PRINTO Multinational Quality of Life cohort study. Rheumatology (Oxford) 2007;46:314–20.
15. Angeles-Han ST. Quality-of-life metrics in pediatric uveitis. Int Ophthalmol Clin 2015;55:93–101.
16. Parker DM, Angeles-Han ST, Stanton AL, Holland GN. Chronic anterior uveitis in children: psychosocial challenges for patients and their families. Am J Ophthalmol 2018;191:vxi–xxiv.
17. Haasnoot AJ, Sint Jago NF, Tekstra J, de Boer JH. Impact of uveitis on quality of life in adult patients with juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2017;69:1895–902.
18. Mangione CM, Lee PP, Gutierrez PR, Spritzer K, Berry S, Hays RD. Development of the 25-item National Eye Institute Visual Function Questionnaire. Arch Ophthalmol 2001;119:1050–8.
19. Bertrand PJ, Jamilloux Y, Kodjikian L, Errera MH, Perard L, Pugnet G, et al. Quality of life in patients with uveitis: data from the ULISSE study (Uveitis: cLInical and medico-economic evaluation of a Standardised Strategy for the Etiological diagnosis). Br J Ophthalmol 2021;105:935–40.
20. Sugar EA, Burke AE, Venugopal V, Thorne JE, Holbrook JT, Multicenter Uveitis Steroid Treatment Trial and Follow-up Study Research Group. Responsiveness of vision-specific and general quality of life metrics to ocular and systemic events in patients with uveitis. Ophthalmology 2020;127:1710–8.
21. Leske DA, Hatt SR, Castaneda YS, Wernimont SM, Liebermann L, Cheng-Patel CS, et al. Validation of the Pediatric Eye Questionnaire in children with visual impairment. Am J Ophthalmol 2019;208:124–32.
22. Tadic V, Robertson AO, Cortina-Borja M, Rahi JS, Child Vision Patient-Reported Outcome Measures Group. An age- and stage-appropriate patient-reported outcome measure of vision-related quality of life of children and young people with visual impairment. Ophthalmology 2020;127:249–60.
23. Birch EE, Cheng CS, Felius J. Validity and reliability of the Children's Visual Function Questionnaire (CVFQ). J AAPOS 2007;11:473–9.
24. Gothwal VK, Lovie-Kitchin JE, Nutheti R. The development of the LV Prasad-Functional Vision Questionnaire: a measure of functional vision performance of visually impaired children. Invest Ophthalmol Vis Sci 2003;44:4131–9.
25. Khadka J, Ryan B, Margrain TH, Court H, Woodhouse JM. Development of the 25-item Cardiff Visual Ability Questionnaire for Children (CVAQC). Br J Ophthalmol 2010;94:730–5.
26. Cochrane GM, Marella M, Keeffe JE, Lamoureux EL. The impact of vision impairment for children (IVI_C): validation of a vision-specific pediatric quality-of-life questionnaire using Rasch analysis. Invest Ophthalmol Vis Sci 2011;52:1632–40.
27. Angeles-Han ST, Griffin KW, Lehman TJ, Rutledge JR, Lyman S, Nguyen JT, et al. The importance of visual function in the quality of life of children with uveitis. J AAPOS 2010;14:163–8.
28. Angeles-Han ST, Griffin KW, Harrison MJ, Lehman TJ, Leong T, Robb RR, et al. Development of a vision-related quality of life instrument for children ages 8–18 years for use in juvenile idiopathic arthritis–associated uveitis. Arthritis Care Res (Hoboken) 2011;63:1254–61.

29. Angeles-Han ST, Yeh S, McCracken C, Jenkins K, Stryker D, Myoung E, et al. Using the effects of youngsters' eyesight on quality of life questionnaire to measure visual outcomes in children with uveitis. Arthritis Care Res (Hoboken) 2015;67:1513–20.
30. Petty RE, Southwood TR, Manners P, Baum J, Glass DN, Goldenberg J, et al. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. J Rheumatol 2004;31:390–2.
31. Varni JW, Seid M, Kurtin PS. PedsQL 4.0: reliability and validity of the Pediatric Quality of Life Inventory version 4.0 generic core scales in healthy and patient populations. Med Care 2001;39:800–12.
32. Singh G, Athreya BH, Fries JF, Goldsmith DP. Measurement of health status in children with juvenile rheumatoid arthritis. Arthritis Rheum 1994;37:1761–9.
33. Nunnally JC. Psychometric theory. New York: McGraw-Hill; 1978.
34. Tavakol M, Dennick R. Making sense of Cronbach's alpha. Int J Med Educ 2011;2:53–5.
35. McLeod S. What does effect size tell you? 2019. URL: https://www.simplypsychology.org/effect-size.html.
36. Wallander JL, Fradkin C, Elliott MN, Cuccaro PM, Tortolero Emery S, Schuster MA. Racial/ethnic disparities in health-related quality of life and health status across pre-, early-, and mid-adolescence: a prospective cohort study. Qual Life Res 2019;28:1761–71.
37. Angeles-Han ST, McCracken C, Yeh S, Jenkins K, Stryker D, Travers C, et al. The association of race with childhood uveitis. Am J Ophthalmol 2015;160:919–28.
38. Foeldvari I, Klotsche J, Simonini G, Edelsten C, Angeles-Han ST, Bangsgaard R, et al. Proposal for a definition for response to treatment, inactive disease and damage for JIA associated uveitis based on the validation of a uveitis related JIA outcome measures from the Multinational Interdisciplinary Working Group for Uveitis in Childhood (MIWGUC). Pediatr Rheumatol Online J 2019;17:66.
39. Heiligenhaus A, Foeldvari I, Edelsten C, Smith JR, Saurenmann RK, Bodaghi B, et al. Proposed outcome measures for prospective clinical trials in juvenile idiopathic arthritis–associated uveitis: a consensus effort from the Multinational Interdisciplinary Working Group for Uveitis in Childhood. Arthritis Care Res (Hoboken) 2012;64:1365–72.
40. Jabs DA, Nussenblatt RB, Rosenbaum JT. Standardization of uveitis nomenclature for reporting clinical data: results of the first international workshop. Am J Ophthalmol 2005;140:509–16.
41. Lovell DJ, Brunner HI, Ringold S, Weiss PF, Martin N, Consolaro A, et al. The American English version of the Juvenile Arthritis Multidimensional Assessment Report (JAMAR). Rheumatol Int 2018;38 Suppl 1:35–42.
42. Drotar D. Measuring health-related quality of life in children and adolescents: implications for research and practice. Mahwah (NJ): Lawrence Erlbaum; 1998.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 3, March 2022, pp 364–370
DOI 10.1002/acr.24499

BRIEF REPORT

Underdetection of Interstitial Lung Disease in Juvenile Systemic Sclerosis

Ivan Foeldvari,[1] Jens Klotsche,[2] Bernd Hinrichs,[3] Nicola Helmus,[1] Ozgur Kasapcopur,[4] Amra Adrovic,[4] Flavio Sztajnbok,[5] Maria Teresa Terreri,[6] Jordi Anton,[7] Vanessa Smith,[8] Maria Katsicas,[9] Mikhail Kostik,[10] Natalia Vasquez-Canizares,[11] Tadej Avcin,[12] Brian Feldman,[13] Mahesh Janarthanan,[14] Maria Jose Santos,[15] Sujata Sawhney,[16] Dieneke Schonenberg-Meinema,[17] Walter-Alberto Sifuentes-Giraldo,[18] Ekaterina Alexeeva,[19] Simone Appenzeller,[20] Cristina Battagliotti,[21] Lillemor Berntson,[22] Blanca Bica,[23] Patrícia Costa-Reis,[24] Despina Eleftheriou,[25] Tilmann Kallinich,[26] Thomas Lehman,[27] Edoardo Marrani,[28] Kirsten Minden,[29] Susan Nielsen,[30] Farzana Nuruzzaman,[31] Anjali Patwardhan,[32] Raju Khubchandani,[33] Valda Stanevicha,[34] Yosef Uziel,[35] and Kathryn S. Torok[36]

Objective. Utilizing data obtained from a prospective, international, juvenile systemic sclerosis (SSc) cohort, the present study was undertaken to determine if pulmonary screening with forced vital capacity (FVC) and diffusing capacity for carbon monoxide (DLco) is sufficient to assess the presence of interstitial lung disease (ILD) in comparison to high-resolution computed tomography (HRCT) in juvenile SSc.

Methods. The juvenile SSc cohort database was queried for patients enrolled from January 2008 to January 2020 with recorded pulmonary function tests (PFTs) parameters and HRCT to determine the discriminatory properties of PFT parameters, FVC, and DLco in detecting ILD.

Results. Eighty-six juvenile SSc patients had both computed tomography imaging and FVC values for direct comparison. Using findings on HRCT as the standard measure of ILD presence, the sensitivity of FVC in detecting ILD in juvenile SSc was only 40%, the specificity was 77%, and area under the curve (AUC) was 0.58. Fifty-eight juvenile SSc patients had both CT imaging and DLco values for comparison. The sensitivity of DLco in detecting ILD was 76%, the specificity was 70%, and AUC was 0.73.

Conclusion. The performance of PFTs in juvenile SSc to detect underlying ILD was quite limited. Specifically, the FVC, which is one of the main clinical parameters in adult SSc to detect and monitor ILD, would miss ~60% of children who had ILD changes on their accompanying HRCT. The DLco was more sensitive in detecting potential abnormalities on HRCT, but with less specificity than the FVC. These results support the use of HRCT in tandem with PFTs for the screening of ILD in juvenile SSc.

INTRODUCTION

Juvenile systemic sclerosis (SSc) is an orphan disease with a prevalence of ~3 in 1 million children, and due to this paucity of cases, most general guidelines of care and treatment from pediatric rheumatologists are based on adult-onset SSc experience (1). Interstitial lung disease (ILD) is a major cause of morbidity for both adult and juvenile-onset SSc (Figure 1), occurring in ~35–55% of

The inception cohort project was supported by an unrestricted grant from the Joachim Herz Foundation.

[1]Ivan Foeldvari, MD, Nicola Helmus, BS: Hamburg Centre for Pediatric and Adolescent Rheumatology, Schön Klinik Hamburg Eilbek, Hamburg, Germany; [2]Jens Klotsche, MD: German Rheumatism Research Center, Berlin, Germany; [3]Bernd Hinrichs, MD: Asklepios Klinik Nord – Heidberg, Hamburg, Germany; [4]Ozgur Kasapcopur, MD, Amra Adrovic, MD: Cerrahpasa Medical School, Istanbul University, Istanbul, Turkey; [5]Flavio Sztajnbok, MD: Universidade do Estado, Rio de Janeiro, Brazil; [6]Maria Teresa Terreri, MD: Universidade Federal de São Paulo, Sao Paulo, Brazil; [7]Jordi Anton, MD: Hospital Sant Joan de Déu, Barcelona, Spain; [8]Vanessa Smith, MD: Ghent University, Ghent University Hospital, and VIB Inflammation Research Center, Ghent, Belgium; [9]Maria Katsicas, MD: Hospital de Pediatria, Buenos Aires, Argentina; [10]Mikhail Kostik, MD: Saint Petersburg State Pediatric Medical University and Almazov National Medical Research Centre, St. Petersburg, Russian Federation; [11]Natalia Vasquez-Canizares, MD: Children's Hospital at Montefiore, Bronx, New York; [12]Tadej Avcin, MD: University Children's Hospital, University Medical Center Ljubljana, Ljubljana, Slovenia; [13]Brian Feldman, MD: SickKids, The Hospital for Sick Children, Toronto, Ontario, Canada; [14]Mahesh Janarthanan, MD: Sri Ramachandra University, Chennai, India; [15]Maria Jose Santos, MD: Hospital Garcia de Orta, Almada, Portugal; [16]Sujata Sawhney, MD: Sir Ganga Ram Hospital, Delhi, India; [17]Dieneke Schonenberg-Meinema, MD: Amsterdam University Medical Centers, Amsterdam, The Netherlands; [18]Walter-Alberto Sifuentes-Giraldo, MD: University Hospital Ramón y Cajal, Madrid, Spain; [19]Ekaterina Alexeeva, MD: National Medical Research Center of Children's Health, Sechenov First Moscow State Medical

SIGNIFICANCE & INNOVATIONS

- The discriminatory properties of forced vital capacity (FVC) have low sensitivity, which would miss interstitial lung disease (ILD) in ~2 of 3 juvenile SSc cohort patients.
- The discriminatory properties of the diffusion capacity for carbon monoxide (DLco) had better sensitivity but lower specificity.
- When FVC and DLco are combined, discriminatory properties are improved, but only when they both are above or below the traditional 80% cutoff, which negates several patients who have values that are on conflicting sides of the cutoffs for the 2 pulmonary function testing values.
- Physicians should ascertain the presence of ILD with the combination of high-resolution computed tomography and pulmonary function tests in patients with juvenile SSc.

patients in published juvenile SSc cohorts (2,3). Currently, novel medications that can attenuate inflammatory-driven lung fibrosis are being tested in clinical trials of adult SSc patients with ILD. Resulting from these trials are medications such as nintedanib, an inhibitor of platelet-derived growth factor receptor α and β, which has been recently licensed for SSc-associated ILD (4), and a pediatric study is planned (EMEA-001006-PIP05-18). Responsiveness to such medication relies on early recognition of ILD to minimize fibrotic load and prevent more permanent tissue damage.

Screening for ILD in adult and pediatric SSc patients traditionally includes pulmonary function tests (PFTs), specifically spirometry and single-breath diffusion capacity for carbon monoxide (DLco) at a minimum (5). This allows the detection of a restrictive lung defect (low forced vital capacity [FVC] and restricted pattern on flow volume loop) with associated decreased DLco, reflecting the thickened interstitium from ILD (after parenchymal inflammation and fibrosis) (6–9). Assessment of FVC is fairly standardized from the age of 3 years, with a grading system for the quality of the test session starting at the age of 2 years (10–12). Following FVC and DLco PFT values in a serial manner in adult SSc has proven to be helpful as a biomarker reflecting lung disease status, with a change of 10% in FVC and 15% in DLco regarded as clinically meaningful (6,8,9,13). However, PFT measurements are limited in detecting earlier lung disease in SSc. This has been recognized in both children and adults because measurements of FVC and DLco are associated with considerable variability due to technical factors, diurnal or seasonal variability, or patient-related factors separate from true pathophysiologic changes. A specific technical example in SSc is the ability to fully make a seal on the mouthpiece to obtain accurate readings; often, with tight facial skin and limited oral aperture, this is not possible. The use of high-resolution computed tomography (HRCT) has become the mainstream imaging modality for screening and quantifying ILD in adults (5,7,8,13), with the combination of HRCT and PFTs both to detect and follow ILD progression and regression (6,8).

In children, traditionally, PFTs have been used for ILD screening, with the more guarded use of HRCT for concerns of radiation in children being reserved for the more symptomatic juvenile SSc patients or for those with abnormal FVC and/or DLco values (<80%) (predicted for age, weight, height, sex, and race) (10,11,14). The decreased amount of radiation used in HRCT recently (14), the evidence from adult SSc suggesting that HRCT is more sensitive for earlier ILD (7), the number of asymptomatic children who tend to subconsciously self-limit physical activity or more rigorous exercise to avoid dyspnea, and the poor reliability of the DLco in children <8 years of age (10–12) all support the more current practice of using HRCT in tandem with PFTs for ILD screening in juvenile SSc. This reflects the more recent management style of pediatric rheumatologists enrolling patients into the international juvenile SSc cohort. Therefore, we reviewed the data collected from our prospective international cohort to evaluate the performance parameters of PFTs in the context of concurrently collected HRCT findings.

PATIENTS AND METHODS

The international inception cohort for patients with juvenile SSc is a large multicenter observational study including 25 centers from Europe, 5 from Asia, 6 from North America, and 6 from South America, representing 42 academic institutions (2). The

University of the Ministry of Health of the Russian Federation, Moscow, Russian Federation; [20]Simone Appenzeller, MD: State University of Campinas, Sao Paulo, Brazil; [21]Cristina Battagliotti, MD: Hospital de Niños Dr. Orlando Alassia, Santa Fe, Argentina; [22]Lillemor Berntson, MD: Uppsala University, Uppsala, Sweden; [23]Blanca Bica, MD: Hospital Universitário Clementino Fraga Filho, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil; [24]Patrícia Costa Reis, MD: Hospital de Santa Maria, Universidade de Lisboa, Lisbon, Portugal; [25]Despina Eleftheriou, MD: Great Ormond Street Hospital for Children, NHS Foundation Trust, London, UK; [26]Tilmann Kallinich, MD: Charité University Medicine, Berlin, Germany; [27]Thomas Lehman, MD: Hospital for Special Surgery, New York, New York; [28]Edoardo Marrani, MD: Meyer Children's Hospital, Florence, Italy; [29]Kirsten Minden, MD: German Rheumatism Research Center and Charité University Medicine, Berlin, Germany; [30]Susan Nielsen, MD: Rigshospitalet, Copenhagen, Denmark; [31]Farzana Nuruzzaman, MD: Stony Brook Children's Hospital, Stony Brook, New York; [32]Anjali Patwardhan, MD: University of Missouri, Columbia; [33]Raju Khubchandani, MD: SRCC Children's Hospital Mumbai, Mumbai, India; [34]Valda Stanevicha, MD: Riga Stradins University, University Children's Hospital, Riga, Latvia; [35]Yosef Uziel, MD: Meir Medical Center, Tel Aviv University, Kfar Saba, Israel; [36]Kathryn S. Torok, MD: Children's Hospital of Pittsburgh, Pittsburgh, Pennsylvania.

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

Address correspondence to Ivan Foeldvari, MD, Hamburg Centre for Pediatric and Adolescence Rheumatology, Schön Klinik Hamburg Eilbek, Dehnhaide 120, 22081 Hamburg, Germany (email: sprechstunde@kinderrheumatoliogie.de); or to Kathryn S. Torok, MD, University of Pittsburgh Scleroderma Center, Children's Hospital of Pittsburgh, 4401 Penn Avenue, Pittsburgh, PA 15224 (email: Kathryn.torok@chp.edu).

Submitted for publication May 28, 2020; accepted in revised form October 27, 2020.

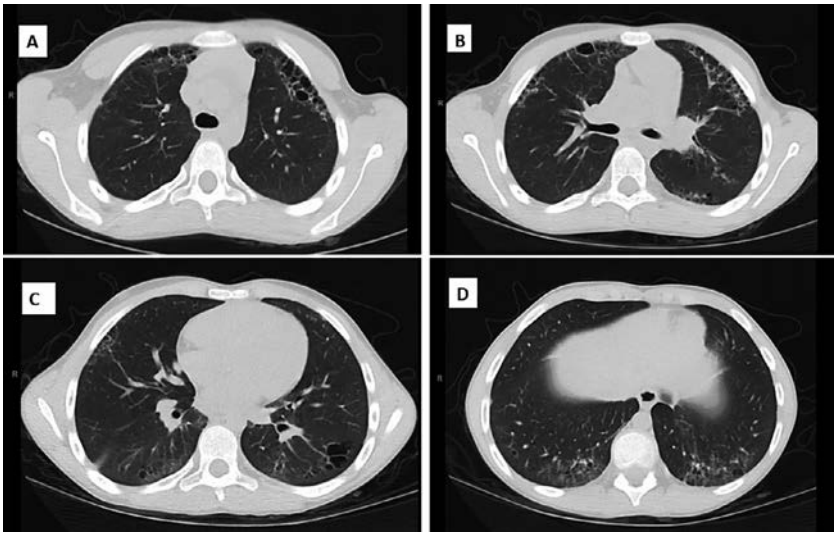


Figure 1. Interstitial lung disease (ILD) in a 9-year-old girl with systemic sclerosis. **A**–**D** (cephalad to caudal lungs) demonstrate high-resolution computed tomography (CT) findings consistent with ILD, including peripheral ground-glass opacities, with distribution in all lobes with hyperdensity in the anterior mid (**A** and **B**) and lower (**C** and **D**) lobes, stacking cystic changes consistent with honeycombing, especially in the anterior upper lobes (**A**), and subpleural cystic changes in the posterior lower lobes (**C** and **D**). Concurrent pulmonary function tests were discordant with the moderate-to-severe degree of pulmonary involvement seen on CT, with a forced vital capacity of 78% predicted, total lung capacity of 96% predicted, and a diffusion capacity for carbon monoxide of 84% predicted. The patient had been treated with cyclophosphamide, hydroxychloroquine, rituximab, glucocorticoids, mycophenolate mofetil, and tocilizumab to help stabilize her ILD, underscoring the severity of her condition.

cohort includes patients with a definite diagnosis of juvenile SSc who are <18 years of age at enrollment. The rheumatologist is requested to report at enrollment (baseline) and every 6 months clinical characteristics, examination findings, laboratory values, evaluations, and treatment related to juvenile SSc management using a standardized case report form. In addition, physician- and patient-reported outcomes of global disease activity and damage are collected. Lung-specific data collection includes PFT parameters (FVC, DLco), which are reported as percentage of the predicted value for the patient's demographic characteristics. A cutoff value of <80% (predicted for age, weight, height, sex, and race) is used to determine an abnormal FVC and DLco as defined as a traditional threshold in healthy children (10–12). Abnormal findings of HRCT examinations were recorded, and findings such as ground-glass opacities, reticulations, and honeycombing consistent for ILD were used to define the presence of ILD (8). PFTs and HRCT were obtained clinically and were suggested to be performed at the study visits; the presence or absence of testing and the results were recorded at each study visit.

Data for this analysis were based on patients enrolled in the juvenile SSc cohort from January 2008 to January 2020 who had recorded PFT parameters and HRCT results to determine the sensitivity of PFT detecting ILD. Standard statistics to evaluate the discriminative ability of FVC and DLco to detect ILD presence or absence, such as sensitivity, specificity, or receiver operating characteristic area under the curve (AUC), were calculated. The proportion of patients with positive findings on HRCT was compared by means of a logistic regression model. Chi-square or *t*-tests were used to determine differences between characteristics of those with and without HRCT performed ($P < 0.05$ was used for significance). To identify clinical characteristics as potential predictors of ILD, a univariable logistic regression of the individual clinical measures was performed, followed by a multivariable model using backward selection ($P < 0.10$) to analyze those variables significant in the univariable model.

RESULTS

Demographic characteristics. Of the 150 patients enrolled in the juvenile SSc cohort at the time of data query, 86 (57%) had both CT imaging and an FVC reading from PFT for direct comparison. Among those, 77% (66 of 86) had diffuse subtype, and 80% were female. Mean ± SD disease duration was 3.1 ± 3.0 years, and mean ± SD age at onset of Raynaud's phenomenon was 10.1 ± 3.9 years. Seventy-nine patients (92%) were >8 years of age at the assessment of PFT and HRCT (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24499/abstract). To assess for bias in those with available studies for the test performance of FVC and DLco, juvenile SSc patients with reported PFT parameters and who underwent an HRCT (n = 86) were compared to patients without these examinations (n = 64) and were found to be comparable across multiple demographic and clinical characteristics, including disease severity assessment, with the only exception being a higher proportion of

Table 1. Diagnostic test properties of forced vital capacity (FVC) and diffusing capacity for carbon monoxide (DLco) as assessment for interstitial lung disease (ILD) detection*

	FVC†		DLco‡	
	<80%	≥80%	<80%	≥80%
ILD on HRCT, no.	15	23	19	6
No ILD on HRCT, no.	11	37	10	23
Sensitivity	39.5 (24.0–56.6)	39.5 (24.0–56.6)	76.0 (54.9–90.6)	76.0 (54.9–90.6)
Specificity	77.1 (62.7–88.0)	77.1 (62.7–88.0)	69.7 (51.3–84.4)	69.7 (51.3–84.4)
Positive predictive value	57.7 (36.9–76.6)	57.7 (36.9–76.6)	65.5 (45.7–82.1)	65.5 (45.7–82.1)
Negative predictive value	61.7 (48.2–73.9)	61.7 (48.2–73.9)	79.3 (60.3–92.0)	79.3 (60.3–92.0)
Area under the curve§	0.58 (0.48–0.68)	0.58 (0.48–0.68)	0.73 (0.61–0.85)	0.73 (0.61–0.85)
Likelihood ratio (+)	1.72 (0.90–3.3)	1.72 (0.90–3.3)	2.51 (1.43–4.40)	2.51 (1.43–4.40)
Likelihood ratio (–)	0.79 (0.58–1.10)	0.79 (0.58–1.10)	0.34 (0.17–0.72)	0.34 (0.17–0.72)

* Values are the % (95% CI) unless indicated otherwise. 95% CI = 95% confidence interval; CT = computed tomography; HRCT = high-resolution CT; PFT = pulmonary function testing.
† N = 86 patients with CT imaging and a FVC reading from PFT.
‡ N = 58 patients with CT imaging and a DLco reading from PFT.
§ Area under the curve estimated for FVC </≥80% and DLco </≥80%.

pulmonary involvement (59% versus 25%; $P < 0.001$) (see Supplementary Table 1, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24499/abstract). Disease prevalence of ILD, defined as HRCT findings consistent with ILD, was 44% (38 of the 86 patients with HRCT results). Clinical characteristics of the 38 children with ILD on HRCT compared to those 48 without evidence on HRCT via univariate analyses demonstrated the association of a higher modified Rodnan skin thickness score (MRSS) and gastrointestinal (GI) involvement with the presence of ILD on HRCT ($P < 0.05$) (see Supplementary Table 2A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24499/abstract), with multivariable model analysis of these factors supporting a trend toward significance of a higher MRSS and significant association between the presence of GI involvement and ILD on HRCT ($P = 0.046$) (see Supplementary Table 2B, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24499/abstract).

Test performance of FVC. Table 1 presents the association between FVC <80%/≥80% and the presence of ILD on HRCT. The discriminative ability (AUC) of FVC as a test to detect ILD was 0.58 (95% confidence interval [95% CI] 0.48–0.68). The sensitivity and specificity were 40% and 77%, respectively.

Test performance of DLco. DLco readings were recorded in 71 juvenile SSc patients from the data query, and 58 of them had accompanying HRCT readings. Disease prevalence of ILD was 43% (25 of 58). The AUC of DLco <80%/≥80% to discriminate between patients with and without ILD on HRCT was 0.73 (95% CI 0.61–0.85) (Table 1). The sensitivity and specificity were 76% and 70%, respectively.

Test performance of FVC and DLco. Both FVC and DLco were reported in 59 patients. Five of the 26 (19%) who had both parameters (FVC and DLco) measuring above the threshold of 80% (normal cutoff) had an abnormal HRCT (see Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24499/abstract). Both FVC and DLco values were recorded as <80% for 12 patients; of these patients, 9 (75%) had ILD on HRCT. The discriminative ability of both DLco and FVC ≥80% versus both DLco and FVC <80% in relation to HRCT findings was 0.76 (95% CI 0.61–0.91), with a sensitivity and specificity of 64% and 88% (see Supplementary Table 4, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24499/abstract). Patients with conflicting PFT results (FVC <80% and DLco ≥80% [n = 5]; FVC ≥80% and DLco <80% [n = 16]) had ~50% detection of ILD on HRCT and were excluded from this combination analysis (see Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24499/abstract). Patients with DLco <80% were more likely to show abnormal findings on HRCT irrespective of FVC (see Supplementary Table 3, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24499/abstract).

Receiver operating curve. The receiver operating characteristic curves for continuous values of FVC, DLco, and the combination of FVC and DLco are shown in Figure 2. The AUC, reflecting the performance of the PFT variables to detect patients with ILD on HRCT, was lower for the FVC (AUC 0.65 [95% CI 0.53–0.77]) compared to the DLco (AUC 0.80 [95% CI 0.67–0.92]) and the combination of FVC and DLco. The minimum of FVC and DLco was selected for combining both tests; in 69% of patients, the DLco was counted because of its lower value compared to FVC (AUC 0.82 [95% CI 0.69–0.93]).

Alternative FVC and DLco cutoffs. In addition to the performance of traditional thresholds for healthy children (80%), alternative thresholds of FVC and DLco of 75%, 85%, and 90% were

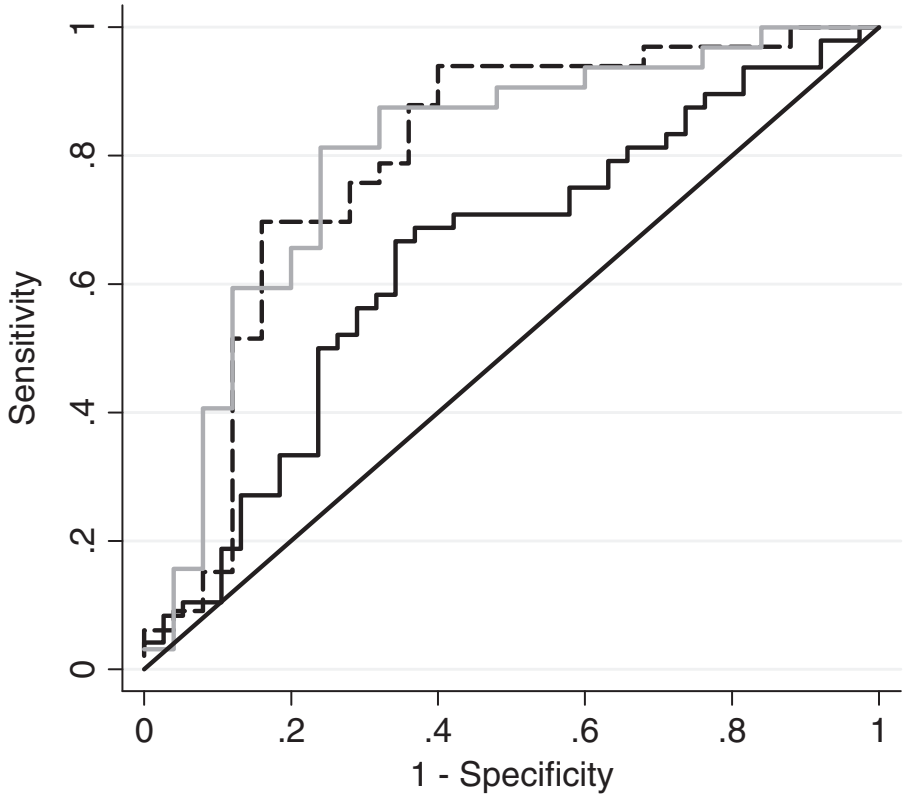


Figure 2. Receiver operating characteristic curves for % predicted forced vital capacity (FVC) (solid line) and diffusion capacity for carbon monoxide (DLco) (broken line) demonstrating the discriminative ability as a test to detect interstitial lung disease on high-resolution computed tomography. The minimum of FVC and DLco (shaded line) was selected for the curve combining both tests; in 69% of patients, the DLco was counted because of its lower value compared to FVC.

examined on an exploratory basis. The sensitivity, specificity, positive predictive value, negative predictive value (NPV), and AUC are reported in Supplementary Tables 5 and 6, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24499/abstract. The combination of increased sensitivity (58% and 84%) without losing too much specificity (71% and 70%) while gaining the maximum AUC (0.64 and 0.77) is reflected in the 85% cutoff for both the FVC and DLco, respectively. DLco continues to have superior discernibility over the FVC using the alternative thresholds.

DISCUSSION

This is the first pediatric evaluation of the discriminatory ability of PFT values (FVC and DLco) to assess ILD in juvenile SSc patients in context with HRCT. The discriminatory properties of the FVC alone using the traditional threshold of 80% was quite poor, with a sensitivity of only 40%; this would have missed the detection of ILD in 60% (23 of 38) of the patients with ground-glass opacities and reticulations on HRCT. The sensitivity of FVC is much lower in juvenile patients compared to adult SSc patients (7,15), where the sensitivity is 0.69 using the same FVC <80% cutoff. The specificity of FVC to detect ILD in juvenile SSc compared to adult SSc patients was similar (0.73 adult SSc [7], and 0.77 juvenile SSc). The limited discrimination of FVC (poor sensitivity, high false negative) may be a particular problem in juvenile SSc because PFTs in children can be more difficult to perform than in adults (10–12). The DLco is even more difficult to perform in children of ages <8 years due to lack of cooperation, and much effort is required during the testing (11). Even though some of these limitations were minimized by our sample of patients being mainly comprised of children >8 years of age (94%), therefore typically obtaining more reliable FVC and DLco results than for younger children, there were still notable limitations of the abilities of these tests to detect lung disease. This may pose even more of a discriminatory problem in younger patients; the general population of the juvenile SSc cohort (n = 150) has ~25% of enrolled patients <8 years of age (2), for whom only FVC can be reliably measured as a screening tool for ILD. The high rate of false negatives (60%) seems to be unacceptable. Even with maximizing the FVC with alternative thresholds, the best sensitivity of 66% obtained at 90% threshold still provides a high rate of false negatives at 35%, which is still problematic for screening purposes in a cohort with potential serious pulmonary involvement.

Applying the DLco measurement does assist in the detection of ILD, with a reasonable sensitivity and specificity of 76% and

70%, respectively, using the standard 80% threshold. Comparable results are also reported in adult SSc (7,13). The ability of PFTs to detect ILD can be improved when DLco is combined with FVC (sensitivity and specificity of 64% and 88%, respectively). Although the combination of FVC and DLco provides good discriminatory properties, the caveat, as mentioned, is the limited ability of children to perform the DLco and the limitation to interpret diagnostic properties of PFTs if the 2 parameters, DLco and FVC, are contradictory (one is normal, and the other is abnormal). The prevalence of ILD in both juvenile SSc and adult SSc is frequent enough that the NPV of the clinical screening test should be high enough to ensure that the SSc patients with a negative screening test, FVC and DLco >80% in this instance, indeed do not have the condition of interest, ILD. The NPV of FVC in juvenile SSc is low at 62%, with a more adequate NPV of the DLco (79%), which is consistent with adult DLco data (7,15). Even with similar NPV values of the DLco, the adult rheumatology community has found PFTs for screening for ILD in SSc patients alone inadequate and endorses baseline HRCT in tandem with PFT given the number of patients with ILD not detected by PFT alone (up to 50% in one study) (9). In agreement with our adult rheumatology colleagues, our data suggest that relying on PFT alone for screening for ILD in juvenile SSc is inadequate and would have missed the detection of ILD in almost two-thirds of the sample cohort, supporting the tandem use of HRCT for detection of ILD in children with juvenile SSc.

Additional clinical information, which may assist in identifying those at risk for ILD in juvenile SSc, increasing the positive predictive value, includes the degree of skin thickening assessed by the MRSS; even more so, from the multivariable model is the association with GI involvement. An average MRSS of 20 compared to 12 was seen in those with ILD compared to those without ILD in the univariate analysis ($P = 0.016$), which supports a possible relationship between the extent of skin thickening with the presence of ILD. This has been established in adult SSc patients, with diffuse cutaneous SSc patients having higher frequency of ILD, especially in earlier disease when skin score is advancing (16). GI involvement remains significant in the multivariable model, providing stronger support for its association with the presence of ILD. Esophageal abnormalities, including dysmotility, low esophageal pressure and contractility, bolus clearance abnormalities, and larger esophageal diameter (dilated, patulous), have been associated with ILD defined by HRCT findings and abnormal PFT readings in adult SSc (17). Although debatable as far as causality of ILD, there is an association between the 2 with the idea that abnormalities in esophageal function and compromised esophageal integrity allow for stasis of esophageal contents, leading to chronic microaspiration, augmenting ILD (17). A recent study in juvenile SSc also supports abnormal esophageal findings on upper GI tests (abnormal esophageal peristalsis or bolus clearance) and increased esophageal diameter on HRCT and suggests a significant association with restrictive lung function, decreased FVC, forced expiratory volume in 1 second, and vital capacity (18). Further investigation of adding such clinical characteristics likely influencing ILD, such as GI involvement, into the PFT sensitivity and specificity model is warranted once a larger number of patients with completed study results is enrolled.

Our study has several limitations. This is an observational clinical study performed across multiple institutions to comprise the international multicenter cohort; therefore, PFT and HRCT findings were collected clinically and not as a requirement to be standardly collected in all patients, reflecting the current standard of care. Eighty-six of 150 registry patients (57%) had available HRCT results to include in the sensitivity and specificity analyses. This may have enriched our population for patients with pulmonary disease. However, the 44% prevalence of pulmonary involvement in our cohort reflects reported prevalence rates in the other published juvenile SSc cohorts (between 36% and 55%) (2,3), and comparison of juvenile SSc patients in our cohort with and without available HRCT results did not reveal any significant demographic, other organ, or overall disease severity differences (see Supplementary Table 1, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24499/abstract). Therefore, we do not project that the standardized screening of all consecutive juvenile SSc patients in our cohort would have resulted in drastically different sensitivity and specificity properties of the PFT in comparison to the HRCT findings. An additional general limitation of our study includes the fact that some of the subgroup analyses were based on a limited number of patients, despite the large cohort of prospectively followed juvenile SSc patients, which must be taken into consideration.

Future directions include confirming our PFT property findings in a validation cohort and evaluating the psychometric properties of sensitivity to change of FVC and ILD in the juvenile SSc cohort longitudinally, as adult SSc studies suggest a decrease of 10% in FVC and 15% in DLco as poor prognostic factors and clinically worsening disease (6,8). Further study in our cohort once a higher number of subjects is enrolled and screened for ILD could include the analysis of the discriminatory properties of these parameters, if the FVC and DLco thresholds were modified from the 80% cutoffs, and the utility of the addition of other clinical characteristics that might predict ILD, such as the MRSS and GI involvement.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Foeldvari, Klotsche, Torok.

Acquisition of data. Foeldvari, Klotsche, Hinrichs, Helmus, Kasapcopur, Adrovic, Sztajnbok, Terreri, Anton, Smith, Katsicas, Kostik, Vasquez-Canizares, Avcin, Feldman, Janarthanan, Santos, Sawhney, Schonenberg-Meinema, Sifuentes-Giraldo, Alexeeva, Appenzeller, Battagliotti, Berntson,

Bica, Reis, Eleftheriou, Kallinich, Lehman, Marrani, Minden, Nielsen, Nuruzzaman, Patwardhan, Khubchandani, Stanevicha, Uziel, Torok.
Analysis and interpretation of data. Foeldvari, Klotsche, Torok.

REFERENCES

1. Smith V, Scire CA, Talarico R, Airo P, Alexander T, Allanore Y, et al. Systemic sclerosis: state of the art on clinical practice guidelines. RMD Open 2018;4 Suppl 1:e000782.
2. Foeldvari I, Klotsche J, Torok KS, Kasapcopur O, Adrovic A, Stanevica V, et al. Characteristics of the first 80 patients at timepoint of first assessment included in the juvenile systemic sclerosis inception cohort. J Scleroderma Relat Disord 2019;4:49–61.
3. Stevens BE, Torok KS, Li SC, Hershey N, Curran M, Higgins GC, et al. Clinical characteristics and factors associated with disability and impaired quality of life in children with juvenile systemic sclerosis: results from the Childhood Arthritis and Rheumatology Research Alliance Legacy Registry. Arthritis Care Res (Hoboken) 2018;70: 1806–13.
4. Distler O, Highland KB, Gahlemann M, Azuma A, Fischer A, Mayes MD, et al. Nintedanib for systemic sclerosis-associated interstitial lung disease. N Engl J Med 2019;380:2518–28.
5. Distler O, Assassi S, Cottin V, Cutolo M, Danoff SK, Denton CP, et al. Predictors of progression in systemic sclerosis patients with interstitial lung disease. Eur Respir J 2020;55:1902026.
6. Volkmann E. Natural history of systemic sclerosis–related interstitial lung disease: how to identify a progressive fibrosing phenotype. J Scleroderma Relat Disord 2020;5:31–40.
7. Showalter K, Hoffmann A, Rouleau G, Aaby D, Lee J, Richardson C, et al. Performance of forced vital capacity and lung diffusion cutpoints for associated radiographic interstitial lung disease in systemic sclerosis. J Rheumatol 2018;45:1572–6.
8. Silver KC, Silver RM. Management of systemic-sclerosis-associated interstitial lung disease. Rheum Dis Clin North Am 2015;41:439–57.
9. Steele R, Hudson M, Lo E, Baron M, on behalf of the Canadian Scleroderma Research Group. Clinical decision rule to predict the presence of interstitial lung disease in systemic sclerosis. Arthritis Care Res (Hoboken) 2012;64:519–24.
10. Quanjer PH, Stanojevic S, Cole TJ, Baur X, Hall GL, Culver BH, et al. Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. Eur Respir J 2012;40: 1324–43.
11. Beydon N, Davis SD, Lombardi E, Allen JL, Arets HG, Aurora P, et al. An official American Thoracic Society/European Respiratory Society statement: pulmonary function testing in preschool children. Am J Respir Crit Care Med 2007;175:1304–45.
12. Graham BL, Steenbruggen I, Miller MR, Barjaktarevic IZ, Cooper BG, Hall GL, et al. Standardization of spirometry 2019 update: an official American Thoracic Society and European Respiratory Society technical statement. Am J Respir Crit Care Med 2019;200:e70–88.
13. Tashkin DP, Volkmann ER, Tseng CH, Kim HJ, Goldin J, Clements P, et al. Relationship between quantitative radiographic assessments of interstitial lung disease and physiological and clinical features of systemic sclerosis. Ann Rheum Dis 2016;75:374–81.
14. Nguyen-Kim TD, Maurer B, Suliman YA, Morsbach F, Distler O, Frauenfelder T. The impact of slice-reduced computed tomography on histogram-based densitometry assessment of lung fibrosis in patients with systemic sclerosis. J Thorac Dis 2018;10:2142–52.
15. Suliman YA, Dobrota R, Huscher D, Nguyen-Kim TD, Maurer B, Jordan S, et al. Pulmonary function tests: high rate of false-negative results in the early detection and screening of scleroderma-related interstitial lung disease. Arthritis Rheumatol 2015;67:3256–61.
16. Walker UA, Tyndall A, Czirják L, Denton C, Farge-Bancel D, Kowal-Bielecka O, et al. Clinical risk assessment of organ manifestations in systemic sclerosis: a report from the EULAR Scleroderma Trials and Research Group database. Ann Rheum Dis 2007;66:754–63.
17. Marie I, Dominique S, Levesque H, Ducrotté P, Denis P, Hellot MF, et al. Esophageal involvement and pulmonary manifestations in systemic sclerosis. Arthritis Rheum 2001;45:346–54.
18. Ambartsumyan L, Zheng HB, Iyer RS, Soares J, Henstorf G, Stevens AM. Relationship between esophageal abnormalities on fluoroscopic esophagram and pulmonary function testing in juvenile systemic sclerosis. Arthritis Care Res (Hoboken) 2019;71:1444–9.

Arthritis Care & Research
Vol. 74, No. 3, March 2022, pp 371–376
DOI 10.1002/acr.24496

American College of Rheumatology
Empowering Rheumatology Professionals

Treatment of Sarcoidosis in US Rheumatology Practices: Data From the American College of Rheumatology's Rheumatology Informatics System for Effectiveness (RISE) Registry

Nevin Hammam,[1] Michael Evans,[1] Esi Morgan,[2] Andreas Reimold,[3] Christine Anastasiou,[1] Julia L. Kay,[1] Jinoos Yazdany,[1] and Gabriela Schmajuk[4]

Objective. Sarcoidosis is often treated with glucocorticoids, although the use of biologics is growing. Prescribing patterns for biologics for patients with sarcoidosis in US rheumatology practices have never been examined. Given that there are no steroid-sparing US Food and Drug Administration–approved therapies for sarcoidosis, we sought to characterize the real-world treatment of sarcoidosis and to assess practice-level variation in prescribing patterns.

Methods. We conducted an observational study of patients with sarcoidosis using data from the Rheumatology Informatics System for Effectiveness (RISE) registry (2014–2018). The RISE registry represents an estimated 32% of the US clinical rheumatology workforce. Adult patients with ≥2 codes for sarcoidosis ≥30 days apart were included. We examined sarcoidosis-specific medication use at any time during the study period. Data were analyzed at the practice level.

Results. A total of 3,276 patients with sarcoidosis from 184 practices were included. Of those patients, 75.1% were women, with a mean age of 59.0 ± 12.5 years; 48.3% were White and 27.6% were Black. Overall, 59.3% of patients were prescribed glucocorticoids, and 24.7% received prolonged glucocorticoid therapy (≥10 mg/day for ≥90 days). In all, 12.1% received a biologic or targeted synthetic disease-modifying antirheumatic drug (tsDMARD), most commonly tumor necrosis factor inhibitors. There was wide practice-level variation among 31 practices with ≥30 patients with sarcoidosis; biologic use ranged from 15.6% to 69.2%. Infliximab represented the most common biologic prescribed.

Conclusion. In a large sample of US rheumatology practices, 12.1% of patients with sarcoidosis received biologics or tsDMARDs. We found high variability in biologic use across practices. The significant use of long-term glucocorticoids suggests unmet therapeutic needs in this patient population.

INTRODUCTION

Sarcoidosis is a rare multisystem disease of unknown etiology with adult-onset typically before the fourth decade. The prevalence of sarcoidosis in the US is estimated to be 35.5 per 100,000 for African Americans and 10.9 per 100,000 for Caucasians (1). The natural history and prognosis of sarcoidosis can vary greatly, from mild and self-limited to severe disease that leads to significant organ impairment and death in 5% of patients (2,3).

The data presented here were supported by the American College of Rheumatology's Rheumatology Informatics System for Effectiveness (RISE) registry. However, the views expressed represent those of the authors and do not necessarily represent the views of the American College of Rheumatology. This study protocol was prepared by and the data analysis was executed by the RISE data analytic center at University of California San Francisco.

The work of Drs. Yazdany and Schmajuk was supported by the Russell/Engleman Medical Research Center for Arthritis. Supported by the NIH (National Institute of Arthritis and Musculoskeletal and Skin Diseases [NIAMS] grants P30-AR-070155 and K24-AR-0074534 to Dr. Yazdany and NIAMS grant 5T32-AR-007304-40 to Dr. Anastasiou).

[1]Nevin Hammam, MD, PhD, Michael Evans, MS, Christine Anastasiou, MD, Julia L. Kay, BA, Jinoos Yazdany, MD, MPH: University of California, San Francisco; [2]Esi Morgan, MD: University of Cincinnati, Cincinnati, Ohio; [3]Andreas Reimold, MD: Dallas Veterans Affairs Medical Center and University of Texas Southwestern University Medical Center, Dallas, Texas; [4]Gabriela Schmajuk, MD, MS: University of California, San Francisco and San Francisco Veterans Affairs Medical Center, San Francisco, California.

Drs. Yazdany and Schmajuk contributed equally to this work.

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

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

Submitted for publication April 6, 2020; accepted in revised form October 22, 2020.

SIGNIFICANCE & INNOVATIONS

- This study provides the first detailed description of national, real-world treatments provided by rheumatologists to patients with sarcoidosis during routine office visits.
- Approximately one-fourth of patients received prolonged, moderate doses of glucocorticoids, and 12.1% of patients with sarcoidosis used biologic or targeted synthetic disease-modifying antirheumatic drugs, despite the absence of US Food and Drug Administration approval for these indications.
- We found wide variations in the patterns of biologic therapies used by US rheumatologists to treat sarcoidosis, likely reflecting the lack of standardized treatment recommendations for this disease.

The highly variable clinical features and disease course, together with a lack of steroid-sparing pharmacologic treatments approved by the US Food and Drug Administration (FDA), explain why treatment for sarcoidosis is not standardized. Current mainstays of treatment include glucocorticoids, disease-modifying antirheumatic drugs (DMARDs), and tumor necrosis factor inhibitors (TNFi) for some manifestations (4). Glucocorticoids are considered first-line treatment for most forms of sarcoidosis but can result in significant cumulative toxicity, even at relatively low doses (5). At present, there is significant variation in medication use for patients with sarcoidosis and, as a result, in medical costs (6,7).

Literature on the real-world treatment of sarcoidosis is limited. One study reported that, among 1,774 patients with sarcoidosis followed in a large university medical center, treatment was prescribed in 61% of patients, with 55.3% and 5.0% of the cohort receiving glucocorticoids and TNFi, respectively (8). This finding contrasts with a more recent study that included all patients with sarcoidosis from a large US insurance claims database in which only 22.8% received any treatment (6).

In this study, we used data from the Rheumatology Informatics System for Effectiveness (RISE) registry, which represents an estimated 32% of the US clinical rheumatology workforce, to examine treatment patterns for sarcoidosis. Understanding real-world medication use can identify areas of unmet therapeutic need and inform future clinical trials.

PATIENTS AND METHODS

Study design and data source. We performed an observational study using data derived from the RISE registry, a national electronic health record (EHR)–enabled rheumatology registry. RISE is a US registry that passively collects data on all patients seen by participating practices, thereby reducing selection bias present in single insurer claims databases (9). Available data are collected through the EHR, mainly from group and private practices across the US, and includes individual demographics, diagnoses, procedures, medications, laboratory test results, and vital signs. Rheumatology practices started contributing data to RISE as early as January 2014, but for some practices, many years of historical data are also available. As of 2018, RISE held validated data from 1,113 providers in 226 practices, with a total of 1,623,504 patients.

Study population and study period. Patients included in this study were age ≥18 years and had ≥2 diagnosis codes for sarcoidosis ≥30 days apart (n = 4,888; International Classification of Diseases, Ninth Revision, Clinical Modification [ICD-9-CM] 135 or ICD-10-CM D86.9) (10). We then excluded subjects who met administrative definitions for other autoimmune conditions (≥2 ICD codes over ≥30 days apart for rheumatoid arthritis, systemic lupus erythematosus, systemic sclerosis, ankylosing spondylitis, inflammatory bowel disease, or psoriatic arthritis; n = 1,612) to increase the specificity of our case definition and because among patients with >1 condition, inferring whether drugs were prescribed for sarcoidosis or for the other rheumatic condition was difficult. This exclusion left 3,276 patients with sarcoidosis in the study sample. The study period included all observation time available in the RISE registry.

Demographic, covariate, and clinical information. We extracted demographic data on participant age, sex, race/ethnicity, and geographic region (East North Central, West North Central, Mid-Atlantic, Mountain, New England, Pacific, South Atlantic, East South Central, and West South Central). Patients were classified as White, Black or African American, or other (which included Hispanic or Latino, Asian, Native Hawaiian or other Pacific Islander, American Indian or Alaska Native, multiracial, and patients with no race classification). Individual comorbidities examined included a diagnosis of diabetes mellitus, asthma, hypertension, and cancer, based on diagnosis codes at any time, as well as Charlson comorbidity index score based on codes from any time during the study period, calculated according to the Deyo modification (11). We enumerated the frequencies of ICD-9-CM/ICD-10-CM codes used to specify sarcoidosis-specific organ involvement for each patient where possible (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24496/abstract).

Medications. Sarcoidosis-specific therapies used during the study period were examined. Systemic glucocorticoids included prednisone and other oral and intravenous steroids. Prolonged glucocorticoid therapy was defined as ≥10 mg of prednisone daily (or its equivalent) for at least 3 months (12). Conventional synthetic DMARDs (csDMARDs) included methotrexate, azathioprine, hydroxychloroquine, leflunomide, sulfasalazine, mycophenolate, cyclosporine, minocycline, and tacrolimus. Biologic DMARDs (bDMARDs) included TNFi (etanercept,

adalimumab, infliximab, golimumab, certolizumab) and non-TNFi biologics (abatacept, rituximab, secukinumab, ustekinumab, omalizumab, anakinra, tocilizumab, sarilumab). Targeted synthetic DMARDs (tsDMARDs) included tofacitinib, baricitinib, and apremilast.

Statistical analysis. Data were presented as the mean ± SD or median with interquartile range (IQR) for numerical variables and as frequency (percentage) for categorical variables. We compared sociodemographic characteristics of patients with sarcoidosis to the overall RISE population. We also compared patients with sarcoidosis with <2 years versus ≥2 years of follow-up using *t*-tests and chi-square tests. The proportion of patients with sarcoidosis by practice was reported out of the total number of patients in the practice. Among practices with ≥30 patients with sarcoidosis (high-volume practices), we calculated the proportion of patients receiving particular medications, out of the total number of patients with sarcoidosis in the practice (practices with <30 patients with sarcoidosis were excluded to reduce the random variation in medication use that would result from small practice sample sizes). Statistical significance was defined as a *P* value of less than 0.05. Data analysis was performed using Stata statistical software, version 15. For privacy protections, we reported no cell sizes <10. This study was approved by the University of California San Francisco and WCG institutional review boards.

RESULTS

Subject characteristics. A total of 3,276 unique patients with sarcoidosis with a mean ± SD age of 59.1 ± 12.5 years were included. Of those patients, 75.1% were female. Most patients (48.3%) were White, and 27.6% were Black. Patients with sarcoidosis represented between 0.2% and 1.8% of patients in the 184 practices included (median 0.2%). The median follow-up time during the study period was 1.9 years (IQR 0.6–4.2). The mean ± SD modified Charlson score was 1.3 ± 0.9. Other characteristics are summarized in Table 1. Age, sex, and geographic distributions of patients with sarcoidosis reflected the underlying population of patients in the RISE registry (see Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24496/abstract). As expected, the proportion of African American patients with sarcoidosis was significantly higher compared to the proportion in RISE overall (27.6% versus 7.2%; $P < 0.001$).

Most patients were identified using codes for "sarcoidosis-unspecified" (51.0%) (Table 2). Among patients with codes for specific clinical manifestations, musculoskeletal involvement (joints, muscles, and bones) was most commonly coded (22.1%), followed by pulmonary (15.1%), ocular (6.3%), renal (6.3%), neurologic (5.3%), and cardiac (1.6%) manifestations. Patients with >2 years of follow-up were more likely to have received codes for specific disease manifestations (see Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24496/abstract). Rheumatoid factor and the anti–cyclic citrullinated peptide antibody were detected in 8.9% and 5.0%, respectively, among patients with results available as structured data.

Table 1. Characteristics of patients with sarcoidosis in the RISE registry*

Characteristic	Total (n = 3,276)
Age, mean ± SD years	59.1 ± 12.5
Female	2,461 (75.1)
Race	
White	1,582 (48.3)
Black or African American	905 (27.6)
Other†	789 (24.1)
Insurance	
Medicare	803 (24.5)
Medicaid	115 (3.5)
Private	1,205 (36.7)
Other‡	151 (4.6)
Missing	1,002 (30.6)
US geographic division	
East North Central	107 (3.3)
West North Central	510 (15.6)
Mid-Atlantic	384 (11.7)
Mountain	299 (9.1)
New England	874 (26.7)
Pacific	396 (12.1)
South Atlantic	321 (9.8)
East South Central	69 (2.1)
West South Central	316 (9.6)
Practice type	
Single-specialty group	2,205 (67.3)
Multispecialty group	555 (16.9)
Solo practitioner	289 (8.8)
Other clinical setting	186 (5.7)
Health system	41 (1.2)
Comorbidities	
Charlson comorbidity index score, mean ± SD	1.3 ± 0.9
Hypertension	545 (16.6)
Asthma	189 (5.8)
Diabetes mellitus	321 (9.8)
Cancer	177 (5.4)
Number of RISE visits, median (IQR)	3.5 (2–6.5)
Follow-up duration, median (IQR) years	1.9 (0.6–4.2)

* Values are the number (%) unless indicated otherwise. IQR = interquartile range; RISE = Rheumatology Informatics System for Effectiveness.
† Other race included Hispanic or Latino, Asian, Native Hawaiian or other Pacific Islander, American Indian or Alaska Native, and multiracial.
‡ Other insurance included veterans and other.

Medication use. Medications prescribed during the study period are shown in Table 3. The majority of patients (59.3%) were prescribed glucocorticoids (at any dose) at some point during the study period; 24.7% of patients received prolonged glucocorticoid therapy (≥10 mg/day for ≥90 days), and 18.2% of patients received glucocorticoid monotherapy (without any DMARDs). Methotrexate and hydroxychloroquine were the most commonly used csDMARDs. Overall, 12.1% received 1 or more bDMARD

Table 2. Frequency of ICD codes for specific clinical manifestations among patients with sarcoidosis in the RISE registry*

Sarcoidosis clinical manifestation	Total (n = 3,276)
Sarcoid, unspecified	1,671 (51.0)
Musculoskeletal	725 (22.1)
Pulmonary (n = 495)	532 (15.1)
Ocular	208 (6.3)
Renal	207 (6.3)
Neurologic	175 (5.3)
Skin	131 (4.0)
Cardiac	54 (1.6)
Lymph	23 (0.7)
RF positivity†	66 (8.9)
Anti-CCP positivity‡	40 (5.0)

* Values are the number (%). International Classification of Diseases (ICD) codes for sarcoidosis clinical manifestations are provided in Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24496/abstract. Anti-CCP = anti–cyclic citrullinated peptide antibody; RF = rheumatoid factor; RISE = Rheumatology Informatics System for Effectiveness.
† Number of patients tested for RF = 745. Patients were categorized as positive for RF at a level of ≥20 units/ml.
‡ Number of patients tested for anti-CCP = 795. Patients were categorized as positive for anti-CCP at a level of ≥20 units/ml.

or tsDMARD, the most common of which were TNFi (10.9%; top drugs were infliximab [6.7%], and adalimumab [4.4%]), followed by rituximab (0.5%), omalizumab (<10), and abatacept (<10). Biologics were usually used in combination with other csDMARDs; only 29 of 387 patients received biologic monotherapy. Among high-volume practices (practices with ≥30 patients with sarcoidosis; n = 31), we found large variations in biologic use, ranging from 15.6% to 69.2% of patients with sarcoidosis (Figure 1). Among these practices, infliximab was used in between 0% and 40% of patients. Similar variability was seen with non-TNF biologics, which were used in between 0% and 50% of patients in each practice.

DISCUSSION

Using data from the RISE registry, we identified 3,276 patients with sarcoidosis, the largest sample for this condition that has been published using real-world EHR data to date. Of these patients, nearly one-fourth were receiving moderate-dose steroids for a prolonged period, and ~12% received bDMARDs or tsDMARDs. We found wide variation in the patterns of biologic use among rheumatology practices. These findings likely reflect the lack of standardized treatment recommendations for this disease.

Rare diseases such as sarcoidosis present unique challenges for researchers working to develop and study treatments or to quantify patient outcomes. For a single center to collect enough patients to make important inferences about effective treatments for the disease is difficult. Disease-specific registries are often collaborations across academic centers and do not provide information about treatment patterns in the community. The RISE registry is unique in its inclusion of US clinical rheumatologists. Participating practices account for an estimated 32% of the clinical rheumatology workforce, and the registry brings the added benefit that patients are not selected based on severity of disease or having been seen at a tertiary care center, where many of the most severe cases may receive care. We found that between 0.2% and 1.8% of patients seen by nonacademic rheumatologists carried a diagnosis of sarcoidosis, and only 31 of 184 practices cared for >30 patients with the condition. Sarcoidosis is relatively rare, and collecting data is also difficult due to the fact that patients may seek care from clinicians across many different specialties, including neurology, dermatology, pulmonology, and rheumatology. This study provides a deeper look into the heterogeneity of this disease among patients seeking care from rheumatologists.

Specific manifestations of sarcoidosis appear to be significantly undercoded in rheumatology EHRs. Not surprisingly for patients seen in rheumatology practices, we found that the most commonly coded organ manifestations were musculoskeletal

Table 3. Medications prescribed by US rheumatologists for patients with sarcoidosis*

Medications	Total patients (n = 3,276)
None recorded	731 (22.3)
Glucocorticoids	
Any prednisone or equivalent†	1,943 (59.3)
Glucocorticoid prolonged‡	811 (24.7)
Glucocorticoid monotherapy	598 (18.2)
Conventional synthetic DMARDs	
Methotrexate	986 (30.1)
Hydroxychloroquine	941 (28.7)
Azathioprine	297 (9.1)
Mycophenolate	187 (5.7)
Leflunomide	146 (4.5)
Sulfasalazine	52 (1.6)
Cyclosporine	<10
Minocycline	36 (1.1)
Tacrolimus	24 (0.7)
Biologics, TNFi	
Infliximab or infliximab biosimilars	219 (6.7)
Adalimumab	145 (4.4)
Etanercept	16 (0.5)
Certolizumab	<10
Golimumab	<10
Biologics, non-TNFi	
Rituximab	17 (0.5)
Omalizumab	<10
Abatacept	<10
Other biologics§	<10
Targeted small molecules	
Tofacitinib	<10
Apremilast	<10

* Values are the number (%). DMARDs = disease-modifying antirheumatic drugs; TNFi = tumor necrosis factor inhibitors.
† Prednisone or equivalent included prednisone and other oral and intravenous steroids.
‡ Prolonged glucocorticoid therapy was defined as ≥10 mg of prednisone daily (or its equivalent) for ≥90 days.
§ Other biologics include belimumab, tocilizumab, anakinra, secukinumab, and ustekinumab.

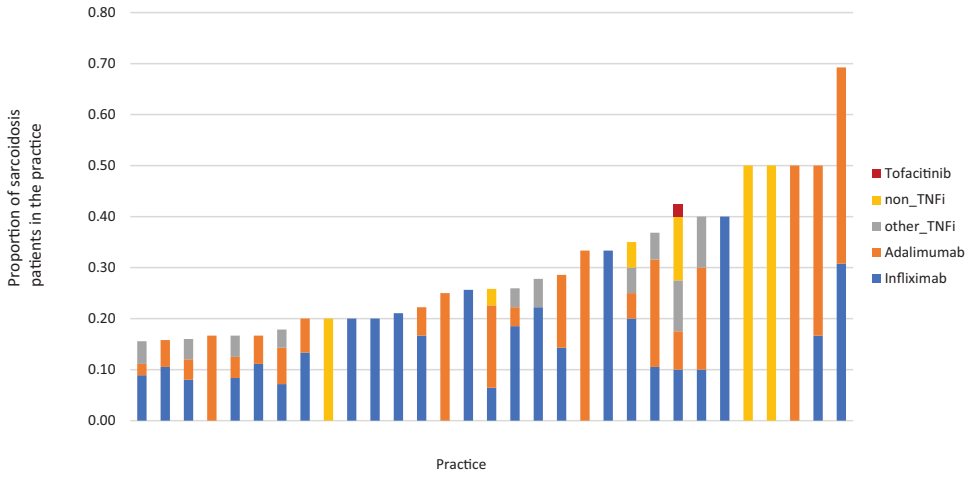


Figure 1. Proportion of patients with sarcoidosis prescribed a biologic or targeted synthetic disease-modifying antirheumatic drug by practice (n = 31) in the Rheumatology Informatics System for Effectiveness registry. Each bar represents a different practice with at least 30 patients with sarcoidosis. The figure shows that biologics use among these practices ranged from 15.6% to 69.2% of patients with sarcoidosis. TNFi = tumor necrosis factor inhibitors.

(22.1%), followed by pulmonary (15.1%). On the contrary, a study of patients with sarcoidosis at a disease-specific clinic at a large university medical center showed that the lung was involved in 89%, followed by the skin (26%) and eyes (23%) (8). Some of this difference could be accounted for by studying a sample of patients who see rheumatologists, but in addition, these differences may be a result of the methods used to extract information about disease manifestations; in the current study, we relied only on ICD codes to identify manifestations, whereas prior cohort studies identified manifestations using radiographic or pathologic reports (13,14). The validation of specific manifestations is not possible when data are extracted from the EHR via problem lists as opposed to detailed case abstraction forms. In the future, using natural language processing approaches to extract information from clinical notes may allow for more detailed identification of specific disease manifestations at scale (15,16).

Glucocorticoids are first-line therapy for sarcoidosis (17), and accordingly, we found that more than half of the patients in this study (59.3%) were treated with glucocorticoids. This finding is similar to previous studies in which prednisone use was reported in 55–65% of patients with sarcoidosis (6,18). These proportions were significantly higher when compared to a recent study using MarketScan data, which found that only 25.5% of patients with sarcoidosis were prescribed prednisone during a single calendar year, although the difference in follow-up time (1 year versus a median of 2 years) may account for theses discrepancies (19). We also found that 24.7% received prolonged glucocorticoids (≥10 mg/day for ≥90 days). This finding is important because a recent expert consensus guideline suggested that a maintenance dose of >10 mg of daily prednisone equivalent was suboptimal and associated with significant side effects (20). The finding that one-fourth of patients exceed this dose for a prolonged period highlights the need for additional and more effective therapies to be developed in this disease.

A total of 12% of patients were prescribed at least 1 biologic or tsDMARD at some point during the study period, presumably to treat signs and symptoms related to their sarcoidosis, since patients with other autoimmune conditions were excluded from the study sample. Infliximab and adalimumab were the most common biologic agents used, which is consistent with prior studies (5), although their overall use was slightly more common than previously reported in the US claims–based data analysis (6), perhaps because patients had more musculoskeletal complaints or because they were all seen by rheumatologists, who were comfortable prescribing these drugs.

We observed meaningful variation in the use of biologics and tsDMARDs across practices. For example, 19% of high-volume practices prescribed infliximab as their only biologic for the treatment of sarcoidosis, while nearly 10% of practices prescribed exclusively non-TNFi biologics such as abatacept. Tofacitinib was only prescribed in 2 practices. Interestingly, trials have shown limited benefit of non-TNFi biologics for sarcoidosis (21–23), and only a few case reports have examined tsDMARDs like tofacitinib in patients with cutaneous sarcoidosis, with some improvement in clinical and histologic remission in the skin disease (24–26). The variation in medication use might partly be explained by patient factors such as race, insurance status, specific manifestations, or disease severity (27) but variation likely also reflects prescribing preferences by clinicians (28). Large studies are clearly needed to evaluate the efficacy and safety of these drugs in patients with sarcoidosis (29).

Using the RISE registry provides a representative sample of nonacademic rheumatology practices across the US and reports on the largest sample of patients with sarcoidosis to date. Despite the strengths of the current study, its limitations should be

addressed. Patients included in this study sought consultation by rheumatologists, and thus, the resulting sample still may not be entirely reflective of the general population of patients with sarcoidosis. Our research examined mainly nonacademic rheumatology practices, so results may not apply to academic health care settings. Although we attempted to be conservative in defining a diagnosis of sarcoidosis by excluding patients with other autoimmune conditions, future work should focus on the validation of codes used to diagnose sarcoidosis and identify its manifestations.

In summary, using data from the RISE registry, we performed the largest study of patients with sarcoidosis to date. We found a significant number of patients were receiving long-term glucocorticoids and a clinically important fraction were receiving biologics. With no FDA-approved drugs available for extrapulmonary sarcoidosis, our findings highlight the need for a greater focus on developing standardized treatments for patients with this disease.

AUTHOR CONTRIBUTIONS

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

Study conception and design. Hammam, Morgan, Reimold, Anastasiou, Yazdany, Schmajuk.

Acquisition of data. Hammam, Evans, Yazdany, Schmajuk.

Analysis and interpretation of data. Hammam, Evans, Kay, Yazdany, Schmajuk.

REFERENCES

1. Cozier YC. Assessing the worldwide epidemiology of sarcoidosis: challenges and future directions. Eur Respir J 2016;48:1545–8.
2. Drent M, Strookappe B, Hoitsma E, De Vries J. Consequences of sarcoidosis. Clin Chest Med 2015;36:727–37.
3. Iannuzzi MC, Fontana JR. Sarcoidosis: clinical presentation, immunopathogenesis, and therapeutics. JAMA 2011;305:391–9.
4. James WE, Baughman R. Treatment of sarcoidosis: grading the evidence. Expert Rev Clin Pharmacol 2018;11:677–87.
5. Beegle SH, Barba K, Gobunsuy R, Judson MA. Current and emerging pharmacological treatments for sarcoidosis: a review. Drug Des Devel Ther 2013;7:325–38.
6. Baughman RP, Field S, Costabel U, Crystal RG, Culver DA, Drent M, et al. Sarcoidosis in America: analysis based on health care use. Ann Am Thorac Soc 2016;13:1244–52.
7. Rice JB, White A, Lopez A, Nelson WW. High-cost sarcoidosis patients in the United States: patient characteristics and patterns of health care resource utilization. J Manag Care Spec Pharm 2017;23:1261–9.
8. Judson MA, Boan AD, Lackland DT. The clinical course of sarcoidosis: presentation, diagnosis, and treatment in a large white and black cohort in the United States. Sarcoidosis Vasc Diffuse Lung Dis 2012;29:119–27.
9. Yazdany J, Bansback N, Clowse M, Collier D, Law K, Liao KP, et al. Rheumatology Informatics System for Effectiveness: a national informatics-enabled registry for quality improvement. Arthritis Care Res (Hoboken) 2016;68:1866–73.
10. Ungprasert P, Matteson EL, Crowson CS. Accuracy of diagnostic coding for sarcoidosis in electronic databases: a population-based study. Lung 2017;195:713–5.
11. Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. J Clin Epidemiol 1992;45:613–9.
12. Korsten P, Strohmayer K, Baughman RP, Sweiss NJ. Refractory pulmonary sarcoidosis: proposal of a definition and recommendations for the diagnostic and therapeutic approach. Clin Pulm Med 2016;23:67–75.
13. Ungprasert P, Carmona EM, Utz JP, Ryu JH, Crowson CS, Matteson EL. Epidemiology of sarcoidosis 1946–2013: a population based study. Mayo Clin Proc 2016;91:183–8.
14. Gribbin J, Hubbard RB, Jeune IL, Smith CJ, West J, Tata LJ. Incidence and mortality of idiopathic pulmonary fibrosis and sarcoidosis in the UK. Thorax 2006;61:980–5.
15. Ananthakrishnan AN, Cai T, Savova G, Cheng SC, Chen P, Perez RG, et al. Improving case definition of Crohn's disease and ulcerative colitis in electronic medical records using natural language processing: a novel informatics approach. Inflamm Bowel Dis 2013;19:1411–20.
16. Jamian L, Wheless L, Crofford LJ, Barnado A. Rule-based and machine learning algorithms identify patients with systemic sclerosis accurately in the electronic health record. Arthritis Res Ther 2019;21:305.
17. Al-Kofahi K, Korsten P, Ascoli C, Virupannavar S, Mirsaeidi M, Chang I, et al. Management of extrapulmonary sarcoidosis: challenges and solutions. Ther Clin Risk Manag 2016;12:1623–34.
18. Judson MA, Baughman RP, Costabel U, Flavin S, Lo KH, Kavuru MS, et al. Efficacy of infliximab in extrapulmonary sarcoidosis: results from a randomised trial. Eur Respir J 2008;31:1189–96.
19. Low D, Simpson KN, Rissmiller R, James E. A new side of sarcoidosis: medication and hospitalization use in a privately insured patient population. Sarcoidosis Vasc Diffuse Lung Dis 2019;36:124–9.
20. Schutt AC, Bullington WM, Judson MA. Pharmacotherapy for pulmonary sarcoidosis: a Delphi consensus study. Respir Med 2010;104:717–23.
21. Judson MA, Baughman RP, Costabel U, Drent M, Gibson KF, Raghu G, et al. Safety and efficacy of ustekinumab or golimumab in patients with chronic sarcoidosis. Eur Respir J 2014;44:1296–307.
22. Bomprezzi R, Pati S, Chansakul C, Vollmer T. A case of neurosarcoidosis successfully treated with rituximab. Neurology 2010;75:568–70.
23. Lower EE, Baughman RP, Kaufman AH. Rituximab for refractory granulomatous eye disease. Clin Ophthalmol 2012;6:1613–8.
24. Damsky W, Thakral D, Emeagwali N, Galan A, King B. Tofacitinib treatment and molecular analysis of cutaneous sarcoidosis. N Engl J Med 2018;379:2540–6.
25. Damsky W, Young BD, Sloan B, Miller EJ, Obando JA, King B. Treatment of multiorgan sarcoidosis with tofacitinib. ACR Open Rheumatol 2020;2:106–9.
26. Damsky W, Thakral D, McGeary MK, Leventhal J, Galan A, King B. Janus kinase inhibition induces disease remission in cutaneous sarcoidosis and granuloma annulare. J Am Acad Dermatol 2020;82:612–21.
27. Bucholc M, O'Kane M, Ashe S, Wong-Lin K. Prescriptive variability of drugs by general practitioners. PLoS ONE 2018;13:e0189599.
28. Yazdany J, Bansback N, Clowse ME, Collier D, Law K, Liao K. Practice-level variation in quality of care in the ACR's Rheumatology Informatics System for Effectiveness (RISE) registry. Arthritis Rheumatol 2016;68:1866–73.
29. Adler BL, Wang CJ, Bui TL, Schilperoort HM, Armstrong AW. Anti-tumor necrosis factor agents in sarcoidosis: a systematic review of efficacy and safety. Semin Arthritis Rheum 2019;48:1093–104.

Arthritis Care & Research
Vol. 74, No. 3, March 2022, pp 377–385
DOI 10.1002/acr.24809

American College of Rheumatology
Empowering Rheumatology Professionals

Prevalence of Early Knee Osteoarthritis Illness Among Various Patient-Reported Classification Criteria After Anterior Cruciate Ligament Reconstruction

Matthew S. Harkey,[1] Shelby Baez,[1] Jordan Lewis,[1] Terry L. Grindstaff,[2] Joseph Hart,[3] Jeffrey B. Driban,[4] Andrew Schorfhaar,[1] and Christopher Kuenze[1]

Objective. To compare the prevalence of participants meeting different patient-reported criteria for early osteoarthritis (OA) illness after anterior cruciate ligament reconstruction (ACLR).

Methods. Participants completed the Knee Injury and Osteoarthritis Outcomes Score (KOOS) at a single time point 5.0–7.9 months post-ACLR. We used established KOOS subscale criteria (i.e., Luyten original and Englund original) to define patient-reported early OA illness. A two-by-two contingency table and McNemar's test were used to compare the prevalence of participants who met the Luyten original versus Englund original KOOS criteria for early OA illness. These analyses were repeated using KOOS subscale thresholds based on established population-specific patient acceptable symptom state (PASS) within the Luyten and Englund KOOS criteria (i.e., Luyten PASS and Englund PASS).

Results. A greater prevalence of participants with ACLR met the Luyten original criteria (n = 165 [54%]) compared to those who met the Englund original criteria (n = 128 [42%]; $\chi^2 = 19.3$, $P < 0.001$]). When using the KOOS subscale PASS as thresholds, a significantly greater prevalence of participants with ACLR met the Luyten PASS criteria (n = 133 [43%]) compared to those who met the Englund PASS criteria (n = 85 [28%]; $\chi^2 = 48.0$, $P < 0.001$). When combining the Luyten and Englund KOOS criteria and using the original/PASS subscale thresholds, respectively, 40%/57% of participants met neither, 24%/15% met only 1, and 36%/28% met both KOOS criteria.

Conclusion. Regardless of the classification criteria used to define early OA illness, it is concerning that 28–54% of patients report considerable symptoms ~6 months post-ACLR. Our findings will improve the classification criteria to define early OA illness, which may raise awareness for the need of population-specific criteria.

INTRODUCTION

Anterior cruciate ligament (ACL) injury and reconstruction (ACLR) occur most commonly in people between 16 and 24 years of age (1,2). ACL injury and the subsequent ACLR are considered inciting events that increase a patient's risk for knee osteoarthritis (OA) (3). To help define knee OA, an Osteoarthritis Research Society International expert working group provided recommendations to the US Food and Drug Administration (4). This expert working group suggested distinguishing OA as a disorder that manifests as: 1) a disease defined by structural changes at a joint and 2) an illness characterized by patient-reported symptoms (4). While this expert working group did not specify which measures to use to define OA, they highlighted the importance of understanding the progression of both OA disease and illness (4). There is substantial evidence that 1 in 3 people will develop radiographic OA disease within a decade post-ACLR (5). When coupling this evidence with the common age of occurrence of ACLR, this means that many young adults will develop OA as early as age 26 to 34 years and then have to live over half their life with significant disease (1,2,6). It is less clear how many people develop early-onset OA illness following ACLR. However, based on a recent national registry, 1 in 3 patients report unacceptable knee-related symptoms (i.e., single-item patient acceptable symptom state [PASS] question) 1 and 2 years post-ACLR (mean ± SD age 29.7 ± 11.5 years) (7). Being able to reliably identify

[1]Matthew S. Harkey, PhD, Shelby Baez, PhD, Jordan Lewis, BS, Andrew Schorfhaar, DO, Christopher Kuenze PhD: Michigan State University, East Lansing; [2]Terry L. Grindstaff, PhD: Creighton University, Omaha, Nebraska; [3]Joseph Hart, PhD: University of Virginia, Charlottesville; [4]Jeffrey B. Driban, PhD: Tufts Medical Center, Boston, Massachusetts.

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.24809&file=acr24809-sup-0001-Disclosureform.pdf.

Address correspondence to Matthew S. Harkey, PhD; Department of Kinesiology, Michigan State University, 308 W Circle Drive, East Lansing, MI, 48824. Email: harkeym1@msu.edu.

Submitted for publication May 21, 2021; accepted in revised form October 19, 2021.

SIGNIFICANCE & INNOVATIONS

- At 5–7 months post–anterior cruciate ligament reconstruction (ACLR), 54% and 42% of people report a level of self-reported disability that meets the Luyten original and Englund original Knee Injury and Osteoarthritis Outcomes Score (KOOS) criteria for early knee osteoarthritis (OA) illness, respectively.
- After refining the Luyten and Englund KOOS criteria by using post-ACLR specific patient acceptable symptom state (PASS) as KOOS subscale thresholds, 43% and 28% of people, respectively, present with self-reported early OA illness.
- By combining the Luyten original and Englund original KOOS criteria into a single composite early knee OA illness variable and using the original/PASS subscale thresholds, respectively, we identified that 36%/28% of participants meet both, 24%/15% meet 1, and 40%/57% meet neither of the original/PASS KOOS criteria.
- Regardless of the criteria or threshold, it is concerning that at least 1 in 4 participants meet a patient-reported definition of early knee OA illness at a time when these patients post-ACLR are nearing a return to unrestricted physical activity.

early OA illness may help clinicians develop targeted strategies to reduce symptom burden and limit OA progression (2,8).

Current OA management focuses on the treatment of symptoms in people with established, radiographic disease rather than interventions that slow OA progression in people with early OA illness (9). Early OA illness is operationally defined as people with knee symptoms without radiographic disease who are at high risk for developing established OA (4,8,10). While the field is in the early phases of establishing the optimal approach to identifying people with early knee OA, patient-reported outcome measures are proposed outcomes in the assessment of early OA illness (2,8). OA illness is easily detected in a clinical setting and often precedes the development of radiographic OA disease (11,12). Therefore, identifying people with early OA illness prior to irreversible joint damage will allow OA management to shift to a more proactive approach that emphasizes targeting interventions to patients who may be more amenable to treatment (2,8,9). The Knee Injury and Osteoarthritis Outcome Score (KOOS) was developed to assess patients across the spectrum of OA, including from post-injury to established knee OA (13). Despite the importance of symptoms in defining early knee OA illness, symptoms experienced by patients over the first year post-ACLR are often considered a normal consequence of surgery rather than chronic symptoms that reflect early knee OA illness (14,15). Yet, there is evidence from longitudinal studies post-ACLR that all KOOS subscale scores remain steady from 6 to 12 months post-ACLR (mean age at ACLR 25.8–28.0 years) (14,16). Additionally, KOOS subscale scores at 1 year post-ACLR remain steady up to 5 and 10 years after surgery (mean age at ACLR 24.0–25.3 years) (17,18). Therefore, a patient's status at 5–7 months post-ACLR may reflect potential long-term outcomes rather than unresolved postsurgical symptoms.

Recently, multiple classification criteria have been developed in multiple studies to define patient-reported OA illness using thresholds for the KOOS subscales of symptoms, pain, quality of life (QoL), function during activities of daily living (ADL), and function during sport and recreation activities (sport) (Table 1) (10,19). Englund et al defined a criterion for knee symptoms (Englund original) that was considered significant enough to seek medical attention based on thresholds of the KOOS subscales in older individuals following meniscal tear (mean $\pm$ SD age 54.3 $\pm$ 11.9 years) (Table 1) (19). Subsequently, Luyten et al proposed a classification criterion for defining early knee OA illness (Luyten original) based on expert consensus using different thresholds for the KOOS subscales and without a specific patient population considered during development (Table 1) (10). Even though these KOOS criteria were not developed for patients with ACLR, the Luyten and Englund KOOS criteria have been applied to patients with ACLR. Furthermore, these criteria are often interpreted to represent early OA illness among patients post-ACLR (20–22). However, the differences in the subscale thresholds and the number of subscales used to define each criterion indicate that meeting the Luyten versus the Englund KOOS criteria may represent varying levels of symptom burden in patients post-ACLR (Table 1).

Table 1. KOOS PASS and subscale thresholds for early knee OA illness criteria*

	KOOS subscale threshold, Englund†		KOOS subscale threshold, Luyten‡	
KOOS subscale	Original	PASS§	Original	PASS§
QoL	87.5	62.5	85.0	62.5
Pain	85.7	88.9	85.0	88.9
Symptoms	86.1	57.1	85.0	57.1
ADL	86.8	100.0	85.0	100.0
Sport and recreation	85.0	75.0	–	–

* Values are the percent of participants. ADL = activities of daily living; KOOS = Knee Injury and Osteoarthritis Outcome Score; OA = osteoarthritis; PASS = patient acceptable symptom state; QoL = quality of life.
† Englund early OA illness KOOS criterion included being below the quality-of-life subscale threshold and below 2 of the other 4 KOOS subscale thresholds (ref. 19).
‡ Luyten early OA illness KOOS criterion included being below 2 of the 4 KOOS subscale thresholds (ref. 10).
§ Developed in individuals 1 to 5 years after a primary anterior cruciate ligament reconstruction (ref. 23).

When using these KOOS criteria, 21% of people at 5 years post-ACLR met the Luyten original criteria for early knee OA illness (21) while 43% and 39% of people met the Englund original criteria at 2 and 6 years post-ACLR (22), respectively. Additionally, the prevalence of people meeting the Luyten original and Englund original criteria for early OA illness from 5 to 7 months post-ACLR is currently unknown. This time period post-ACLR is clinically important because it is when many patients are engaging with a health care professional to return to unrestricted activity and could be screened for early knee OA illness.

The Luyten original and Englund original KOOS criteria represent a key step toward defining early knee OA illness; however, the previous authors emphasized that the proposed classification strategies are a starting point that needs further refinement (10). One approach to refine these KOOS classification criteria to be more applicable for detecting early knee OA illness in participants post-ACLR is to set the KOOS subscale thresholds to previously published PASS (23). Using the ACLR-specific PASS as KOOS subscale thresholds in the early OA illness criteria (i.e., Luyten PASS and Englund PASS) may represent a more patient population–specific approach for defining meaningful OA-related symptoms in patients post-ACLR, because these subscale thresholds indicate clinically significant symptoms specifically for patients with ACLR (23).

Therefore, the purpose of the present study was to compare the prevalence of participants at 5 to 7 months post-ACLR who meet the Luyten original and Englund original KOOS criteria for early OA illness. We hypothesized that a greater prevalence of participants would meet the Luyten original compared to the Englund original KOOS criteria for early OA illness. Additionally, we also refined the early OA illness criteria by repeating these analyses with the updated Luyten PASS and Englund PASS KOOS criteria that use previously established PASS for each KOOS subscale in people post-ACLR. We hypothesized that a smaller proportion of participants would meet the Luyten PASS and Englund PASS KOOS criteria than the original criteria because we are using a more patient population–specific approach for defining early knee OA illness. This was a necessary next step to refine the previously recommended KOOS criteria to highlight the utility of these patient-reported outcomes criteria for identifying those with early knee OA illness in a group of participants at increased risk for being diagnosed with knee OA. Identifying people with early knee OA illness following ACLR would provide clinicians and researchers a strategy for targeting patients who may need further care (e.g., self-management guidance, education, additional therapeutic exercise) while they are still engaged in the health care system.

MATERIALS AND METHODS

Study design. This was a secondary analysis of patient-reported outcomes collected for ongoing research at Michigan State University (n = 123), the University of Virginia (n = 55), and Creighton University (n = 128). These analyses rely on cross-sectional data from a single visit in ongoing larger observational studies at each site, which relied on similar eligibility criteria.

Participants. We included participants between 13 and 35 years of age who underwent a primary, unilateral ACLR 5.0 to 7.9 months prior to study enrollment. We selected 5 to 7 months post-ACLR because this is an important period when patients are in the terminal phases of rehabilitation, cleared for unrestricted physical activity, and when KOOS scores have potentially stabilized (14,16,24). Participants were excluded if they had a multiligament reconstruction or a neurologic, cardiovascular, or other medical condition that prevented safe study participation in the ongoing larger studies at each site that included biomechanics and strength testing. Participants 18 years of age or older (≥19 years of age at the Creighton University site) provided informed written consent, while participants <18 years of age (<19 years of age at the Creighton University site) provided informed written assent and a legal guardian's written consent. All experimental procedures were approved by each institution's institutional review board.

KOOS. The KOOS questionnaire was developed for patients across the spectrum of OA (e.g., from post-injury to established knee OA) and has been used in men and women ranging from 14 to 79 years of age with varying knee pathologies (13). Participants completed the KOOS questionnaire to evaluate the following 5 subscales: symptoms, pain, QoL, ADL, and sport (13). The subscales have multiple 5-point Likert scale questions with responses ranging from 0 (i.e., no dysfunction) to 4 (i.e., worst dysfunction). Each subscale was converted to a score from 0 to 100, with 100 indicating no patient-reported disability (25,26). A PASS threshold for each KOOS subscale has previously been established in participants post-ACLR: symptoms >57.1%, pain >88.9%, QoL >62.5, ADL >100.0%, and sports >75.0 (Table 1) (23). These KOOS subscale PASS thresholds were developed in people who were at the 1 to 5 year time point after a primary ACLR, using a receiver operator curve analysis to identify participants that answered yes to the following question: "Taking into account all the activity you have during your daily life, your level of pain, and also your activity limitations and participation restrictions, do you consider the current state of your knee satisfactory?" (23).

Luyten original KOOS criterion for defining early knee OA illness. The Luyten original KOOS criterion operationally defines a participant as having early knee OA illness if they score ≤85% on at least 2 of the following 4 KOOS subscales: symptoms, pain, QoL, or ADL (Table 1) (10). The Luyten original KOOS criterion omits the KOOS sport subscale in the early OA

illness definition and has been used to define early knee OA illness in a previous study of people post-ACLR (21).

Englund original KOOS criterion for defining early knee OA illness. Englund et al defined the level of self-reported knee symptoms that were considered significant enough to seek medical attention in people following meniscal injury based on the KOOS (19). The Englund original KOOS criterion operationally defines a participant as having significant knee symptoms if they meet the KOOS subscale threshold of QoL ≤87.5% and 2 of the 4 following thresholds: symptoms ≤85.7%, pain ≤86.1%, ADL ≤86.8%, or sport ≤85.0% (Table 1) (8). These Englund KOOS subscale thresholds represent the score of a subscale if participants report a 1 on a scale of 0–4 for at least 50% of the questions per subscale. For example, if a participant responded 0 out of 4, 0 out of 4, 1 out of 4, and 1 out of 4 for the 4 KOOS QoL questions, their KOOS QoL subscale score would be 87.5%. This criterion was initially introduced in a study of participants following meniscal injury but has also been used to define significant knee symptoms in studies of people post-ACLR (20,22).

Using the KOOS PASS to update the subscale thresholds. In addition to using the previously defined KOOS subscale thresholds for the Luyten and Englund KOOS criteria, we also separated our cohort by the Luyten PASS and Englund PASS KOOS criteria using KOOS subscale thresholds based on previously established PASS in participants post-ACLR (Table 1) (23). Using the Luyten PASS and Englund PASS criteria allowed us to couple the KOOS criteria logic created by the previous authors with patient population–specific KOOS subscale thresholds that reflected meaningful symptoms post-ACLR.

Statistical analysis. A two-by-two contingency table was created to determine how many people met none, only 1, or both the Luyten original and Englund original KOOS criteria. A McNemar's test was then used to compare the prevalence of participants who met the Luyten original versus the Englund original KOOS criteria. Mean values and SDs for continuous variables or frequencies and percentages for categorical variables were calculated for the demographic, injury, or KOOS subscales for participants who met: 1) neither the Luyten original nor the Englund original KOOS criteria, 2) 1 of either the Luyten original or Englund original KOOS criteria, or 3) both the Luyten original and Englund original KOOS criteria. These analyses were repeated to compare the prevalence of participants who met the Luyten PASS and Englund PASS KOOS criteria, which utilizes the PASS as the threshold for each of the KOOS subscales. Because time since ACLR may influence patient-reported outcomes, we performed a post hoc stratified analysis that repeated the analyses in participants who were at the 5-month time point post-ACLR (n = 109), as well as in participants who were at the 6- or 7-month time point post-ACLR (n = 197). We combined participants at 6 and 7 months post-ACLR due to smaller sample size of participants at 7 months post-ACLR (n = 41). We used SAS version 9.4 for all statistical analyses and set the alpha level a priori at $P < 0.05$.

RESULTS

The prevalence of participants who had undergone ACLR who met the Luyten original KOOS criteria of early knee OA illness (n = 165 [54%]) was significantly greater ($\chi^2 = 19.3$, $P < 0.0001$) than those who met the Englund original KOOS criteria (n = 128 [42%]) (Table 2). Across both original KOOS criteria, 40% of participants met neither the Luyten original nor the Englund original KOOS criteria, 24% met only 1 of the Luyten original or Englund original KOOS criteria, and 36% met both the Luyten original and the Englund original KOOS criteria (Table 2). Participant demographic characteristics for each of these groups are in Table 3.

When using the KOOS subscale PASS as the criteria thresholds, the prevalence of participants post-ACLR who met the Luyten PASS KOOS criteria (n = 133 [43%]) was significantly greater ($\chi^2 = 48.0$, $P < 0.001$) than those who met the Englund PASS KOOS criteria (n = 85 [28%]) (Table 4). Across both PASS KOOS criteria, 57% of participants met neither the Luyten PASS nor the Englund PASS KOOS criteria, 16% met only 1 of the Luyten PASS or Englund PASS KOOS criteria, and 28% met both the Luyten PASS and Englund PASS KOOS criteria (Table 4). Thirty-five percent of participants were classified differently between the PASS KOOS criteria that use the patient population–specific KOOS subscale thresholds and the original KOOS criteria subscale thresholds (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24809/abstract). Since both the Luyten

Table 2. Prevalence of the Luyten original and Englund original KOOS criteria for early knee OA illness in participants post-ACLR*

	Englund original KOOS criteria		
	No early OA illness	Early OA illness	Total
Luyten original KOOS criteria			
No early OA illness	124 (40)	17 (6)	141 (46)
Early OA illness	54 (18)	111 (36)	165 (54)
Total	178 (58)	128 (42)	306 (100)

* Values are the number (%). ACLR = anterior cruciate ligament reconstruction (see Table 1 for other definitions).

Table 3. Demographic characteristics across each combination of meeting the Luyten original and Englund original KOOS criteria for early OA illness*

Demographic characteristics	Overall cohort	Did not meet Luyten or Englund criteria	Met Luyten/did not meet Englund criteria; Did not meet Luyten/met Englund criteria	Met Luyten and Englund criteria
No. of participants	306	124	71	111
Female, no. (%)	162 (53)	68 (55)	56 (50)	56 (50)
Age, years	20.0 ± 4.9	18.7 ± 3.8	19.3 ± 4.5	21.8 ± 5.6
BMI (kg/m^2)	24.9 ± 4.5	24.6 ± 4.1	24.2 ± 4.5	25.7 ± 4.9
Time post-ACLR, months	6.2 ± 0.6	6.2 ± 0.7	6.3 ± 0.6	6.1 ± 0.6
Preinjury Tegner score (range 0–10), median (IQR)†	9 (7–9)	9 (8–10)	9 (7–10)	9 (7–9)
IKDC (range 0–100)	81.7 ± 11.7	88.6 ± 7.6	83.3 ± 8.8	72.9 ± 11.4
KOOS subscales (range 0–100)‡				
QoL	66.2 ± 19.2	77.4 ± 17.2	66.5 ± 15.1	53.7 ± 15.7
Pain	92.1 ± 8.4	96.4 ± 4.4	94.3 ± 5.7	85.8 ± 9.5
Symptoms	77.6 ± 15.4	89.6 ± 11.3	71.4 ± 11.2	68.1 ± 12.6
ADL	97.5 ± 6.6	99.5 ± 1.1	98.6 ± 2.9	94.4 ± 9.8
Sport	83.3 ± 15.0	91.0 ± 9.0	89.9 ± 10.3	70.6 ± 14.4

* Values are the mean ± SD unless indicated otherwise. ACLR = anterior cruciate ligament reconstruction; BMI = body mass index; IKDC = International Knee Documentation Committee Subjective Knee Form (lower score indicates worse function); IQR = interquartile range (see Table 1 for other definitions).
† Based on the Tegner Activity Scale, which includes a one-item score, grading based on work and sports activities.
‡ For the Knee Osteoarthritis Outcome Score (KOOS), lower score indicates worse outcome.

PASS and Englund PASS KOOS criteria are using the same KOOS subscale thresholds, it is impossible for a participant to meet the Englund PASS criteria but not the Luyten PASS KOOS criteria; therefore, we omitted this column in Table 4. Table 5 provides the demographic characteristics and KOOS subscale scores across the groups that met neither, 1, or both the Luyten PASS and Englund PASS KOOS criteria. In the post hoc stratified analysis that separated the group by time since ACLR, the prevalence of participants meeting the different KOOS classification criteria was similar between a group of participants at 5 months post-ACLR (n = 109) compared to participants at 6 or 7 months post-ACLR (n = 197) (see Supplementary Table 2, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24809/abstract).

DISCUSSION

This study highlights that 54% and 42% of participants at 5–7 months post-ACLR met the previously established Luyten original or Englund original KOOS criteria for early knee OA illness, respectively. Our results also highlight the discrepancy between the Luyten original and the Englund original KOOS criteria when classifying people with early knee OA illness, as 24% of the cohort met either the Luyten original or Englund original criteria but not both (Table 2; met Luyten/did not meet Englund criteria = 18%, did not meet Luyten/met Englund criteria = 6%); however, when we adopted ACLR-specific KOOS subscale PASS thresholds for the Luyten PASS and Englund PASS KOOS criteria, we found that 43% and 28% of participants, respectively, met the criteria for early knee OA illness. Using the ACLR-specific PASS as KOOS subscale thresholds may represent a more patient population–specific approach for defining meaningful OA-related symptoms in patients post-ACLR because these subscale thresholds indicate clinically significant symptoms specifically for patients with this pathology (23). Regardless of the criteria or threshold, it is concerning that 28% to 54% of participants met these criteria for early knee OA illness at 5 to 7 months post-ACLR.

While the Luyten and Englund KOOS criteria represent similar constructs that use OA-related symptoms across multiple

Table 4. Prevalence of the Luyten and Englund PASS KOOS criteria for early knee OA illness in participants post-ACLR using PASS as thresholds for each KOOS subscale*

	Englund PASS KOOS criteria		
	No early OA illness	Early OA illness	Total
Luyten PASS KOOS criteria			
No early OA illness	173 (57)	0 (0)	173 (57)
Early OA illness	48 (15)	85 (28)	133 (43)
Total	221 (72)	85 (28)	306

* Values are the number (%) of participants. ACLR = anterior cruciate ligament reconstruction; KOOS = Knee Injury and Osteoarthritis Outcome Score; OA = osteoarthritis; PASS = patient acceptable symptom state.

Table 5. Demographic characteristics across each combination of meeting the Luyten PASS and Englund PASS KOOS criteria for early OA illness*

Characteristics	PASS KOOS criteria		
	Did not Luyten or Englund criteria	Met Luyten, did not meet Englund criteria	Met Luyten and Englund criteria
No. of participants	173	48	85
Female sex, no. (%)	94 (54)	24 (50)	44 (52)
Age, years	19.1 ± 4.4	19.8 ± 4.5	21.8 ± 5.5
BMI (kg/m^2)	24.7 ± 4.5	25.2 ± 4.5	25.3 ± 4.6
Time post-ACLR, months	6.2 ± 0.7	6.2 ± 0.6	6.1 ± 0.6
Preinjury Tegner (range 0–10), median (IQR)†	9 (8–10)	8.5 (7–9)	9 (7–9)
IKDC (range 0–100)	87.4 ± 8.1	80.7 ± 7.7	70.5 ± 11.6
KOOS subscales (range 0–100)‡			
QoL	75.8 ± 16.1	62.7 ± 15.0	48.8 ± 13.3
Pain	96.6 ± 3.4	90.6 ± 7.4	83.8 ± 9.5
Symptoms	82.3 ± 13.9	74.9 ± 15.1	69.4 ± 14.9
ADL	99.6 ± 1.5	95.8 ± 13.0	94.1 ± 5.9
Sport	90.1 ± 9.6	86.4 ± 11.0	67.9 ± 14.8

* Values are the mean ± SD unless indicated otherwise. ACLR = anterior cruciate ligament reconstruction; ADL = activities of daily living; BMI = body mass index; IKDC = International Knee Documentation Committee Subjective Knee Form (lower score indicates worse function); IQR = interquartile range; QoL = quality of life.
† Based on the Tegner Activity Scale, which includes a one-item score, grading based on work and sports activities.
‡ For the Knee Osteoarthritis Outcome Score (KOOS), lower score indicates worse outcome.

KOOS subscales, the present study highlights the difference between the 2 criteria in categorizing people as having early knee OA illness post-ACLR (Table 2). This difference occurs because the Englund KOOS criteria require the KOOS QoL subscale and 2 other KOOS subscales to meet respective thresholds (19), while the Luyten KOOS criteria only require any 2 KOOS subscales to be below their threshold (10). Due to the discrepancy between the Luyten and Englund original KOOS criteria highlighted in Table 2, we proposed an exploratory approach to identify the patients with the most concerning symptoms by creating a single variable to define early knee OA illness. In order to do so, we separated people into the following 3 groups based on the original and PASS KOOS criteria: 1) no early knee OA illness (does not meet either the Luyten or Englund criteria), 2) possible early knee OA illness (meets only 1 of the Luyten or Englund criteria), and 3) probable early knee OA illness (meets both Luyten and Englund criteria).

When combining the Luyten original and Englund original KOOS criteria, 40%, 24%, and 36% of participants met the definition of no, possible, and probable early knee OA illness, respectively (Table 2). When combining the Luyten PASS and Englund PASS KOOS criteria, 28%, 15%, and 57% of participants met the definition of no, possible, and probable early knee OA illness, respectively (Table 4). Tables 3 and 5 highlight the demographic characteristics for participants who were considered to have no, possible, and probable OA-related symptoms using the original and PASS KOOS criteria, respectively. Participants who were considered to have probable early knee OA illness were similar to the other 2 groups in sex, body mass index, time post-ACLR, and preinjury physical activity, but were on average 2–3 years of age older (Tables 3 and 5). We propose that this novel composite variable offers a tiered approach to define early knee OA illness, which will allow future researchers to easily identify people with probable OA-related symptoms based on various combinations of subthreshold KOOS subscales. Interestingly, the percentage of participants from this study who fit into the probable early knee OA illness category using the original KOOS criteria (36%) is similar to the prevalence of magnetic resonance imaging evidence of early knee OA at 1 year post-ACLR (31%) (27), radiographic knee OA at 10 years post-ACLR (36%) (5), and unacceptable symptoms at 1, 2, and 6 years post-ACLR (33–43%) (7,22). An important next step will be to determine if the presence of probable early knee OA illness is associated with a greater risk of developing incident OA–related structural pathology.

This analysis also identified 24 participants (15%) who presented with possible early knee OA illness using the original/PASS KOOS classification criteria (i.e., meeting only 1 the Luyten or Englund KOOS criteria) (Table 2). This finding indicates that participants may be differentially classified as having OA-related symptoms that is solely dependent on which method the investigators selected to define early knee OA illness. Therefore, people with possible knee OA illness may be missed if researchers or clinicians opt to define OA-related symptoms using the Luyten original versus the Englund original KOOS criteria (and vice versa). Even though these individuals with possible early knee OA illness may not have the same magnitude of symptoms as those with probable early knee OA illness, they still meet an established criterion for defining OA-related symptoms at 5 to 7 months post-ACLR. These individuals with possible early knee OA illness may represent a clinically relevant group who present with an earlier

stage of OA illness that may be ideal for interventions prior to the development of more symptomatic disability. Identifying patients with probable knee OA illness may represent a conservative way to target those in most need of intervention but identifying people with possible knee OA illness may ensure that we capture all individuals with relevant symptoms while they are still engaging with the health care system. Due to the cross-sectional nature of the present study, it is unclear whether participants with possible early knee OA illness were on a trajectory of symptom improvement or a trajectory leading to developing probable early knee OA illness or clinical diagnosis of knee OA.

The creators of the Luyten original and Englund original KOOS criteria emphasize that their proposed classification strategies for defining early knee OA are starting points that need further refinement (10). One way to refine the proposed classification strategies is to update the KOOS subscale thresholds to be more specific to various patient populations. The Englund original KOOS criteria was created in patients following meniscal tear who were an average of 54 years of age (19). The Luyten original KOOS criteria was created by an international expert working group and did not justify why ≤85% was selected as the threshold for all KOOS subscales (10). Our study refined these definitions of early knee OA illness for people post-ACLR by creating the Luyten PASS and Englund PASS KOOS criteria, adjusting the KOOS subscale thresholds to match their respective PASS using values from a prior study of patients post-ACLR (23).

This approach allowed us to pair the KOOS criteria logic created by the previous authors with patient population–specific KOOS subscale thresholds that reflect meaningful symptoms post-ACLR. A lower prevalence of definite and probable early knee OA illness when using the PASS KOOS criteria (28% and 15%) compared to the prevalence using the original KOOS criteria (36% and 24%; Table 2), respectively, is shown in Table 4. Additionally, 35% of people were classified differently between the original KOOS criteria and our refined PASS KOOS criteria (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24809/abstract). Specifically, 22% of participants characterized as having possible/probable early knee OA illness with the original KOOS criteria were characterized as having no early knee OA illness when using the PASS KOOS criteria. Thus, updating the KOOS subscale thresholds to post-ACLR PASS scores for the Luyten PASS and Englund PASS KOOS criteria may offer a more patient population–specific method for defining early knee OA illness in people post-ACLR.

While the present study provides important information to refine and highlight the prevalence of early knee OA illness in people post-ACLR, there are some limitations. The cross-sectional design of this study only allowed us to highlight cross-sectional differences in the prevalence of these different early knee OA illness definitions at an average of 6 months post-ACLR; we were unable to definitively state whether participants reached a stable symptom state. Additionally, we were unable to confirm where the participants were in their rehabilitation plan or if they were cleared for unrestricted physical activity. However, our post hoc stratified analysis highlights that a similar proportion of participants met the various early OA classification criteria at 5 and 6 or 7 months post-ACLR (see Supplementary Table 2, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24809/abstract). Future longitudinal studies need to determine if meeting a symptom criterion for early OA illness criteria at 5–7 months post-ACLR represents a risk factor for future clinical or structural knee OA progression or whether these symptoms represent ongoing recovery from ACLR. Since the present study utilized multiple different variations of the Englund and Luyten early OA illness criteria, future studies need to determine if the original KOOS subscales or the patient population–specific PASS subscales have a greater prognostic potential for predicting OA progression.

The clinical utility of the present study is that once we establish which method of defining early knee OA illness is most prognostic of future OA progression, we will be able to detect the people at greatest risk and selectively target interventions that prevent or slow the progression of OA. Additionally, once there is an accepted definition of early knee OA illness for people post-ACLR, future work in larger cohorts may benefit from determining if sex, age, and other factors are risk factors for early knee OA illness after ACLR. Our study focused on defining early knee OA illness based solely on patient-reported outcome measures, as this method represents a cost-efficient, clinically relevant option for defining early OA–related symptoms, which may be a precursor to a clinical diagnosis of knee OA. However, OA is a multifactorial disorder that manifests as the patient's experiences living with OA (i.e., OA illness) and structural changes occurring within the joint (i.e., OA disease) (4). Therefore, using other criteria, like clinical and imaging examinations, may be needed to specifically identify people presenting with both the illness and disease components of early knee OA (2,8,10). Since the current study did not include an imaging examination, we cannot confirm whether or not participants had radiographic evidence of knee OA. However, due to the relatively young age and the relatively short time period following surgery (5 to 7 months post-ACLR), the majority of participants likely do not have radiographic knee OA. Furthermore, our focus was the manifestation of early knee OA illness after an ACLR, which is focused on a patient's perception.

The present study highlights the prevalence and various refinements of criteria used to define early knee OA illness post-ACLR. At 5 to 7 months post-ACLR, 54% and 42% of people meet the Luyten original and Englund original KOOS criteria, respectively, which indicates a level of self-reported disability that may reflect early knee OA illness. We created a novel strategy for defining patient-reported early knee OA illness by combining the Luyten and Englund KOOS criteria into a single composite variable to identify that 36% and 28% of participants present with probable (i.e., meeting both criteria) and early knee OA illness

based on the original and PASS KOOS subscales, respectively. Using the Luyten PASS and Englund PASS KOOS criteria, which use ACLR-specific PASS as the KOOS subscale thresholds, offer a more patient population-specific method for defining early knee OA illness post-ACLR. Regardless of the early knee OA illness definition used, it is concerning that 28–54% of patients are reporting considerable symptoms at 5 to 7 months post-ACLR. Therefore, we need to pay more attention to these patient-reported outcomes during this critical time after an ACLR and try to address these outcomes or teach patients what to do if these symptoms persist. This study represents an important step to better define the criteria to define early OA illness, which may raise awareness for research and clinicians to screen for these people and to develop treatment strategies specific for this patient population.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. Dr. Harkey 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. Harkey, Baez, Driban, Kuenze.

Acquisition of data. Harkey, Lewis, Grindstaff, Hart, Schorfhaar, Kuenze.

Analysis and interpretation of data. Harkey, Baez, Grindstaff, Driban, Hart, Kuenze.

REFERENCES

1. Sanders TL, Maradit Kremers H, Bryan AJ, Larson DR, Dahm DL, Levy, BA, et al. Incidence of anterior cruciate ligament tears and reconstruction: a 21-year population-based study. Am J Sports Med 2016;44:1502–7.
2. Watt FE, Corp N, Kingsbury SR, Frobell R, Englund M, Felson DT, et al. Towards prevention of post-traumatic osteoarthritis: report from an international expert working group on considerations for the design and conduct of interventional studies following acute knee injury. Osteoarthritis Cartilage 2019;27:23–33.
3. Lohmander LS, Englund PM, Dahl LL, Roos EM. The long-term consequence of anterior cruciate ligament and meniscus injuries: osteoarthritis. Am J Sports Med 2007;35:1756–69.
4. Lane NE, Brandt K, Hawker G, Peeva E, Schreyer E, Tsuji W, et al. OARSI-FDA initiative: defining the disease state of osteoarthritis. Osteoarthritis Cartilage 2011;19:478–82.
5. Luc B, Gribble PA, Pietrosimone BG. Osteoarthritis prevalence following anterior cruciate ligament reconstruction: a systematic review and numbers-needed-to-treat analysis. J Athl Train 2014;49:806–19.
6. Ackerman IN, Bucknill A, Page RS, Broughton NS, Roberts C, Cavka B, et al. The substantial personal burden experienced by younger people with hip or knee osteoarthritis. Osteoarthritis Cartilage 2015;23:1276–84.
7. Ingelsrud LH, Granan LP, Terwee CB, Engebretsen L, Roos EM. Proportion of patients reporting acceptable symptoms or treatment failure and their associated KOOS values at 6 to 24 months after anterior cruciate ligament reconstruction: a study from the Norwegian Knee Ligament Registry. Am J Sports Med 2015;43:1902–7.
8. Emery CA, Whittaker JL, Mahmoudian A, Stefan Lohmander L, Roos EM, Bennell KL, et al. Establishing outcome measures in early knee osteoarthritis. Nat Rev Rheumatol 2019;15:438–48.
9. Chu CR, Williams AA, Coyle CH, Bowers ME. Early diagnosis to enable early treatment of pre-osteoarthritis. Arthritis Res Ther 2012;14:212.
10. Luyten FP, Bierma-Zeinstra S, Dell'Accio F, Kraus VB, Nakata K, Sekiya I, et al. Toward classification criteria for early osteoarthritis of the knee. Semin Arthritis Rheum 2018;47:457–63.
11. Case R, Thomas E, Clarke E, Peat G. Prodromal symptoms in knee osteoarthritis: a nested case-control study using data from the Osteoarthritis Initiative. Osteoarthritis Cartilage 2015;23:1083–9.
12. Davis J, Eaton CB, Lo GH, Bing L, Price LL, McAlindon TE, et al. Knee symptoms among adults at risk for accelerated knee osteoarthritis: data from the Osteoarthritis Initiative. Clin Rheumatol 2017;36:1083–9.
13. Roos EM, Lohmander LS. The Knee injury and Osteoarthritis Outcome Score (KOOS): from joint injury to osteoarthritis. Health Qual Life Outcomes 2003;1:64.
14. Teitsma XM, van der Hoeven H, Tamminga R, de Bie RA. Impact of patient sex on clinical outcomes: data from an anterior cruciate ligament reconstruction registry, 2008-2013. Orthop J Sports Med 2014;2:2325967114550638.
15. Hill GN, O'Leary ST. Anterior cruciate ligament reconstruction: the short-term recovery using the Knee Injury and Osteoarthritis Outcome Score (KOOS). Knee Surg Sports Traumatol Arthrosc 2013;21:1889–94.
16. Frobell RB, Roos EM, Roos HP, Ranstam J, Lohmander LS. A randomized trial of treatment for acute anterior cruciate ligament tears. N Engl J Med 2010;363:331–42.
17. Ahlden M, Samuelsson K, Sernert N, Forssblad M, Karlsson J, Kartus J. The Swedish National Anterior Cruciate Ligament Register: a report on baseline variables and outcomes of surgery for almost 18,000 patients. Am J Sports Med 2012;40:2230–5.
18. Spindler KP, Huston LJ, Chagin KM, Kattan MW, Reinke EK, Amendola A, et al. Ten-year outcomes and risk factors after anterior cruciate ligament reconstruction: a MOON longitudinal prospective cohort study. Am J Sports Med 2018;46:815–25.
19. Englund M, Roos EM, Lohmander LS. Impact of type of meniscal tear on radiographic and symptomatic knee osteoarthritis: a sixteen-year followup of meniscectomy with matched controls. Arthritis Rheum 2003;48:2178–87.
20. Pietrosimone B, Blackburn JT, Padua DA, Pfeiffer SJ, Davis HC, Luc-Harkey BA, et al. Walking gait asymmetries 6 months following anterior cruciate ligament reconstruction predict 12-month patient-reported outcomes. J Orthop Res 2018;36:2932–40.
21. Arhos EK, Thoma LM, Grindem H, Logerstedt D, Risberg MA, Snyder-Mackler L. Association of quadriceps strength symmetry and surgical status with clinical osteoarthritis 5 years after anterior cruciate ligament rupture. Arthritis Care Res (Hoboken) 2022;74:386–391.
22. Wasserstein D, Huston LJ, Nwosu S, MOON Group, Kaeding CC, Parker RD, et al. KOOS pain as a marker for significant knee pain two and six years after primary ACL reconstruction: a Multicenter Orthopaedic Outcomes Network (MOON) prospective longitudinal cohort study. Osteoarthritis Cartilage 2015;23:1674–84.
23. Muller B, Yabroudi MA, Lynch A, Lai CL, Niek van Dijk C, Fu FH, et al. Defining thresholds for the patient acceptable symptom state for the IKDC Subjective Knee Form and KOOS for patients who underwent ACL reconstruction. Am J Sports Med 2016;44:2820–6.
24. Wright RW, Haas AK, Anderson J, Calabrese G, Cavanaugh J, Hewett TE, et al. Anterior cruciate ligament reconstruction rehabilitation: MOON guidelines. Sports Health 2015;7:239–43.
25. Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynnon BD. Knee Injury and Osteoarthritis Outcome Score (KOOS)–development of a

self-administered outcome measure. J Orthop Sports Phys Ther 1998;28:88–96.

26. McHugh M, Droy E, Muscatelli S, Gagnier JJ. Measures of adult knee function. Arthritis Care Res (Hoboken) 2020;72:219–49.

27. Culvenor AG, Collins NJ, Guermazi A, Cook JL, Vicenzino B, Khan KM, et al. Early knee osteoarthritis is evident one year following anterior cruciate ligament reconstruction: a magnetic resonance imaging evaluation. Arthritis Rheumatol 2015;67:946–55.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 3, March 2022, pp 386–391
DOI 10.1002/acr.24479

Association of Quadriceps Strength Symmetry and Surgical Status With Clinical Osteoarthritis Five Years After Anterior Cruciate Ligament Rupture

Elanna K. Arhos,[1] Louise M. Thoma,[2] Hege Grindem,[3] David Logerstedt,[4] May Arna Risberg,[5] and Lynn Snyder-Mackler[1]

Objective. The objective of this study was to examine the association of quadriceps strength symmetry and surgical status (anterior cruciate ligament [ACL] reconstruction or nonoperative management) with early clinical knee osteoarthritis (OA) 5 years after ACL injury or reconstruction.

Methods. In total, 204 of 300 athletes were analyzed 5 years after ACL injury or reconstruction. Quadriceps strength was measured and reported as a limb symmetry index. We identified participants with early clinical knee OA using the criteria that 2 of 4 Knee Injury and Osteoarthritis Outcome Score (KOOS) subscales score ≤85%. We calculated odds ratios (ORs) and 95% confidence intervals (95% CIs) using logistic regression and adjusted for age, sex, meniscal injury, and body mass index to examine the associations of quadriceps strength and surgical status with clinical knee OA.

Results. In all, 21% of participants met the KOOS criteria for clinical knee OA. For every 1% increase in quadriceps limb symmetry index, there was a 4% lower odds of clinical OA (adjusted OR [OR_{adj}] 0.96 [95% CI 0.93–0.99]) at 5 years. Surgical status was not associated with clinical knee OA (OR_{adj} 0.58 [95% CI 0.23–1.50]).

Conclusion. More symmetric quadriceps strength, but not surgical status, 5 years after ACL injury or reconstruction was associated with lower odds of clinical knee OA.

INTRODUCTION

After anterior cruciate ligament (ACL) rupture, the chances of developing knee osteoarthritis (OA) increase rapidly. Patients after ACL rupture have a 60–90% chance of developing knee OA, with ~50% of individuals developing OA within 5–10 years of ACL reconstruction (ACLR) (1,2). Posttraumatic knee OA (PTOA), which may develop after a joint injury, has devastating effects, including lower participation in activity, more pain, and worse self-reported quality of life compared to non-injured individuals (3). Strategies are needed to identify individuals after ACL rupture who would benefit from targeted intervention to mitigate the disease process before symptoms and joint damage occur.

OA is traditionally diagnosed radiographically by the presence of osteophytes and the loss of joint space, but we know that the OA disease starts early, long before it is visible on radiographs (4). Secondary prevention (i.e., implementing interventions after the initiating trauma has occurred) may be the most ideal time to intervene clinically. However, we lack targets for intervention at this level. Secondary prevention depends on the detection and treatment of risk factors for development and progression of OA. The ability to clinically detect early OA prior to widespread joint damage would allow for early clinical intervention and education after ACL rupture. Identifying modifiable risk factors for posttraumatic OA, such as muscle weakness, where we can clinically intervene at an earlier time point may delay OA progression.

Previously identified modifiable factors associated with early development of PTOA after ACL rupture and reconstruction include asymmetric knee walking mechanics, poor functional performance, and changes in knee joint loading (5–8). Although quadriceps weakness has not been directly implicated in the

Supported by the NIH (grants R37-HD-037985 and F32-AR-073090) and the University of Delaware (Department of Physical Therapy PT PhD Endowment grant).

[1]Elanna K. Arhos, PT, DPT, Lynn Snyder-Mackler, PT, ATC, ScD, SCS, FAPTA: University of Delaware, Newark; [2]Louise M. Thoma, PT, DPT, PhD: University of North Carolina at Chapel Hill; [3]Hege Grindem, PT, PhD: Oslo Sports Trauma Research Center, Oslo, Norway, and Stockholm Sports Trauma Research Center, Karolinska Institutet, Stockholm, Sweden; [4]David Logerstedt, PhD, MPT, MA: University of the Sciences in Philadelphia, Philadelphia, Pennsylvania; [5]May Arna Risberg, PT, PhD: Oslo University Hospital and Norwegian School of Sports Sciences, Oslo, Norway.

Dr. Snyder-Mackler has received consulting fees from Flexion Therapeutics (less than $10,000). No other disclosures relevant to this article were reported.

Address correspondence to Elanna K. Arhos, PT, DPT, 540 South College Avenue, Suite 210 Z, Newark, DE 19713. Email: earhos@udel.edu.

Submitted for publication April 14, 2020; accepted in revised form October 1, 2020.

SIGNIFICANCE & INNOVATIONS

- More symmetric quadriceps strength, but not management (anterior cruciate ligament [ACL] reconstruction or rehabilitation alone), 5 years after ACL rupture was associated with lower odds of early clinical knee osteoarthritis (OA).
- Our results provide a potential target (i.e., quadriceps strength symmetry) for secondary OA prevention programs after ACL rupture, allowing clinicians the possibility of intervening with quadriceps strengthening to potentially delay symptomatic knee OA.

development of PTOA after ACL rupture, loss of quadriceps strength over time is associated with symptomatic knee OA long term (9). Furthermore, quadriceps weakness has been shown to be associated with an increased risk of developing knee OA (10). Patients with quadriceps weakness also show movement asymmetries at the knee joint and poor functional performance and patient-reported outcomes (11). Quadriceps weakness contributes to altered walking mechanics early after ACLR (12). Similarly, quadriceps weakness is associated with radiographic knee OA and lower patient-reported outcomes early after ACLR (13). After meniscectomy, weaker quadriceps are associated with more severe radiographic OA changes of the operated and contralateral knee at the 11-year follow-up visit (14).

Quadriceps strength is a critical objective measurement used to make clinical decisions throughout rehabilitation (15,16). Limb symmetry indexes are often used to express quadriceps strength and are calculated as the involved limb's strength value as a percentage of the uninvolved limb value. For safe return to sport and prior activity level, quadriceps strength symmetry is highly recommended to be >90% (17). Similarly, individuals who present with minimal strength deficits (quadriceps strength symmetry >90%) at the time of return to sport perform similarly functionally compared to uninjured individuals (18). Quadriceps strength may therefore be an important modifiable component of a clinical evaluation to identify individuals at risk for developing PTOA.

Recent criteria, proposed by Luyten et al (19) after an international workshop and consensus process, were developed to classify early-stage OA in a primary care setting. Development of consensus classification criteria is an important step toward detecting early knee OA; however, the validity, sensitivity, and responsiveness of these criteria are not yet known. The classification criteria included scores from 2 of the 4 Knee Injury and Osteoarthritis Outcome Score (KOOS) subscales and a clinical examination of joint line tenderness and crepitus in those with no radiographic OA (i.e., Kellgren/Lawrence grade of 0–1). As a first step toward classifying presence of clinical knee OA, the KOOS component of the Luyten et al (19) criteria may offer a strategy to monitor patient symptoms remotely, with no procedures such as an office visit or imaging required. These clinical criteria have the potential to be used to identify individuals who may be candidates for secondary prevention strategies during rehabilitation after ACL injury, as well as during postoperative rehabilitation, and may bridge the gap between primary and tertiary levels of prevention.

The purpose of this study was to examine the association of quadriceps strength with presence of early clinical knee OA 5 years after ACL rupture. We hypothesized that more symmetric quadriceps strength would be associated with lower odds of clinical knee OA. Secondarily, we hypothesized that surgical status was not associated with clinical knee OA.

PATIENTS AND METHODS

This was a secondary analysis of prospectively collected data from an international cohort study through the University of Delaware (Newark, Delaware) and the Oslo University Hospital (Oslo, Norway). The University of Delaware Institutional Review Board and the Regional Committee for Medical Research Ethics of South-Eastern Norway approved this study. All participants provided written informed consent prior to inclusion.

Only participants between 13 and 60 years old who were regularly (>50 hours per year) involved in level I (e.g., soccer, basketball) or II (e.g., racket sports, baseball) sports (20) at the time of ACL rupture were eligible. Participants were excluded if they had a concomitant grade III ligamentous injury, articular cartilage lesions of >1 cm^2, bilateral lower extremity injury (e.g., contralateral ligamentous injury), or an obviously repairable meniscal tear. ACL ruptures were confirmed with magnetic resonance imaging and a side-to-side difference of ≥3 mm measured with a KT-1000 arthrometer (MEDmetric). The cohort included those who elected either ACLR or rehabilitation alone as management after ACL injury. This decision was made via a shared decision-making process during a period in which the participant performed a prehabilitation program (21). Those who selected ACL reconstruction had a bone–patellar tendon–bone autograft, single-bundle hamstring autograft, or double-bundle hamstring autograft. For this analysis, we included 204 of 300 individuals who returned for their 5-year follow-up data collection and had complete data sets, including KOOS score and bilateral quadriceps strength measures at 5 years (Figure 1).

Quadriceps femoris muscle strength testing. Quadriceps femoris muscle strength was measured using a Kin-Com electromechanical dynamometer (DJO Global; Biodex) during a maximal voluntary isometric contraction knee extension test in Delaware. Participants were seated with their knees and hips flexed to 90 degrees. The dynamometer's axis of rotation was aligned with the axis of rotation of the knee, and the leg was strapped in at the pelvis, thigh, and shank during testing to minimize accessory motion. Each participant completed 3 submaximal practice trials and then 3 maximal effort trials for the uninvolved

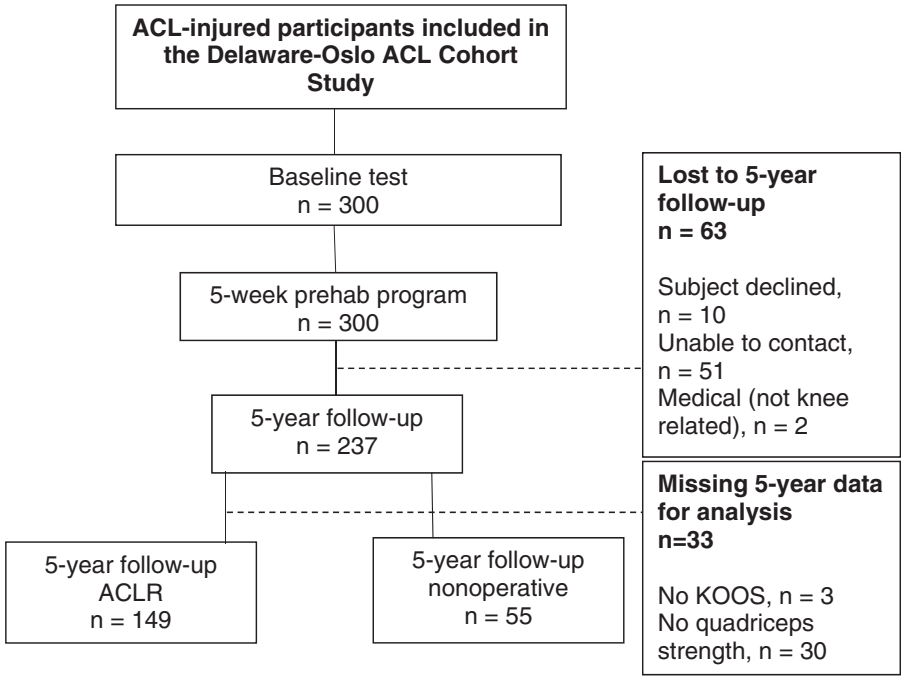


Figure 1. Delaware–Oslo Anterior Cruciate Ligament (ACL) cohort study consort diagram. ACLR = ACL reconstruction; KOOS = Knee Injury and Osteoarthritis Outcome Score.

limb, followed by the involved limb. In Oslo, Norway, quadriceps strength was measured with a Biodex 6000 isokinetic dynamometer (Biodex Medical Systems). Participants performed 4 submaximal practice trials, a 1-minute rest, and then 5 recorded maximal effort repetitions for the uninvolved limb, followed by the involved limb. For the purposes of this study, we defined strength as symmetric peak strength values between limbs. Quadriceps strength was reported as a limb symmetry index (LSI), which was calculated as the involved limb maximum torque divided by the uninvolved limb maximum torque expressed as a percentage.

Table 1. Participant demographic characteristics*

Characteristic	Total sample (n = 204)	Early OA (n = 41)†	Without early OA (n = 163)
Time since screening, years	5.5 ± 0.5	5.7 ± 0.6	5.5 ± 0.5
Age, years	32.7 ± 10.0	32.4 ± 9.8	32.6 ± 10.1
Female, no. (%)	96 (47)	21 (51)	76 (47)
BMI, kg/m^2‡	25.4 ± 4.0	25.5 ± 4.6	25.5 ± 3.8
Graft type, no. (%)§			
Nonoperative	55 (27)	8 (19)	47 (29)
Allograft	41 (20)	11 (27)	30 (19)
BPTB	30 (15)	6 (15)	24 (15)
Hamstrings	76 (38)	16 (39)	60 (37)
Operative status, no. (%)			
ACLR	149 (73)	33 (80)	116 (71)
Nonoperative rehabilitation	55 (27)	8 (20)	47 (29)
Quadriceps strength LSI, %	97.3 ± 15.1	91.9 ± 15.1	97.8 ± 14.8
KOOS			
Pain	93.8 ± 9.3	81.5 ± 11.9	97.06 ± 4.8
Symptoms	89.8 ± 12.6	70.8 ± 14.1	94.5 ± 5.9
ADL	97.6 ± 6.2	91.6 ± 11.4	99.1 ± 1.9
Sports/recreation	88.4 ± 16.3	69.4 ± 22.7	93.3 ± 9.3
QoL	79.4 ± 19.3	57.5 ± 20.1	84.9 ± 14.6

* Values are the mean ± SD unless indicated otherwise. ACLR = anterior cruciate ligament reconstruction; ADL = activities of daily living; BMI = body mass index; BPTB = bone–patellar tendon–bone; KOOS = Knee Injury and Osteoarthritis Outcome Score; LSI = limb symmetry index; OA = osteoarthritis; QoL = quality of life.
† Early OA as defined by KOOS criteria in Luyten et al (19).
‡ BMI was carried forward from the 2-year follow-up for 2 participants.
§ Two unknown graft types in the group without early OA.

KOOS. The KOOS, a patient-reported outcome of knee pain and function, was administered to all participants. The KOOS is comprised of 5 subscales (pain, symptoms, activities of daily living, sports and recreation function, and knee-related quality of life) (22). The KOOS subscales scores range from 0 to 100, with 100 indicating no impairment. We applied the patient-reported criteria of the proposed early clinical OA symptoms by Luyten et al (19). The criteria for early clinical knee OA was defined as ≤85% in 2 of 4 KOOS subscales (pain, symptoms, activities of daily living, knee-related quality of life) at the 5-year follow-up of the Delaware–Oslo ACL Cohort Study (19).

Statistical analyses. Statistical analyses were performed using SAS, version 9.4. A significance level of *P* less than 0.05 was set a priori. Descriptive statistics were calculated to describe the participant demographics. The exposures of interest were quadriceps strength LSI (continuous variable) and surgical status (ACLR or nonoperative management). For descriptive purposes, we also dichotomized quadriceps strength LSI into symmetric (≥90% LSI) and asymmetric (<90% LSI). The outcome was presence or absence of early clinical knee OA based on the Luyten et al (19) KOOS criteria (2 of 4 KOOS subscales scoring ≤85%). We calculated odds ratios (ORs) and 95% confidence intervals (95% CIs) using logistic regression to examine the association of quadriceps strength symmetry and surgical status with presence of early clinical knee OA 5 years after ACL rupture, adjusting for sex, age, body mass index, and meniscus injury at baseline.

RESULTS

Of 300 participants who enrolled in the Delaware–Oslo ACL Cohort Study at baseline, 204 (149 ACLR; 55 underwent rehabilitation alone) who returned for the 5-year follow-up completed the KOOS questionnaire and quadriceps strength testing (Table 1). Of 204 participants, 21% (41 of 200) were classified as having early clinical knee OA according to the modified Luyten et al criteria (19). Fifty-four percent of participants scored ≤85% on the KOOS knee-related quality of life subscale (111 of 204) (Table 2). Fewer participants had scores ≤85% on the symptoms (38 of 204) and pain (24 of 204) or activities of daily living (8 of 204) subscales.

Quadriceps asymmetry (<90% LSI) was observed in 54% of those with clinical knee OA and 35% of those without clinical knee OA. More symmetric quadriceps strength was associated with lower odds of clinical knee OA (adjusted OR [OR_{adj}] 0.96 [95% CI 0.93–0.99]). Specifically, for every 1% higher quadriceps strength LSI, there was 4% lower odds of clinical knee OA. Fourteen percent of those managed with rehabilitation alone and 22% of those after ACLR had clinical knee OA (9). Surgical status was not associated with clinical OA (OR_{adj} 0.58 [95% CI 0.23–1.50]).

Table 2. Total sample categorized by Luyten et al (19) criteria application*

	Value
KOOS ≤85	
Pain	24 (12)
Symptoms	38 (19)
ADL	8 (4)
QoL	111 (54)
No. of KOOS subscale scores ≤85	
0	89 (44)
1	74 (36)
2	24 (12)
3	9 (4)
4	8 (4)

* Values are the number (%). ADL = activities of daily living; KOOS = Knee Injury and Osteoarthritis Outcome Score; QoL = quality of life.

DISCUSSION

The results supported our primary hypothesis that higher quadriceps LSI was associated with lower odds of clinical knee OA. When considering clinical application and utility of our results, our analysis suggests that a modest 1% higher quadriceps LSI was associated with 4% lower odds of clinical knee OA. Five years after ACL rupture or reconstruction, 54% of those classified with clinical knee OA had a quadriceps LSI of <90%. The importance of maintaining quadriceps strength after ACL rupture that has been underscored in the literature is further supported by this analysis. Quadriceps weakness is associated with the development of symptomatic knee OA (13,23). In individuals with diagnosed knee OA, stronger quadriceps are correlated with increased physical activity (24). At the time of return to sport, patients with weaker quadriceps demonstrate poorer knee joint function and altered landing patterns (18,25). Our results, taken with previous literature, highlight the importance of maintaining symmetric quadriceps strength after traumatic knee injury, such as ACL rupture, for long-term joint health.

Treatment choice (i.e., ACLR plus rehabilitation versus rehabilitation alone) was not associated with early clinical knee OA in our cohort. Prior studies also reported that long-term outcomes, including development of OA, are not different between those who undergo ACLR versus rehabilitation alone after ACL injury (26,27). Maintaining long-term quadriceps strength symmetry is an important component of rehabilitation and long-term joint health regardless of how ACL ruptures are managed.

Our results support the use of the modified KOOS classification component of Luyten et al (19) to identify participants who may benefit from continued further clinical intervention (i.e., quadriceps strength training) beyond discharge from rehabilitation, and education on quadriceps weakness being a significant risk factor for the development and progression of knee OA (10). After ACL rupture, nearly all patients (98%) assume they have no or only slightly increased risk of developing posttraumatic

OA (28). Patients may benefit from discussions with health care practitioners regarding outcomes on the KOOS classification component of Luyten et al (19), including strategies for maintaining joint health such as maintaining quadriceps strength symmetry. Our results provide a direction for targeted secondary OA prevention after ACL rupture or reconstruction.

Practitioners and patients may be able to use the KOOS classification component of Luyten et al (19) to monitor clinical OA risk. Fifty-four percent of the study sample had deficits in knee-related quality of life (KOOS quality of life ≤85), and this included nearly all (40 of 41) participants classified with early clinical OA. While deficits in knee-related quality of life were common, participants were required to report deficits in at least 1 more subscale (pain, symptoms, or ADL) to be classified with clinical knee OA for the KOOS portion of the Luyten et al criteria (19). Worse scores on the KOOS pain subscale are associated with poorer patient-reported outcomes (e.g., worse function, greater fear of movement, worse knee confidence) after ACLR (29). Additionally, lower KOOS sports and recreation and pain subscales are associated with low self-reported knee stability in individuals with knee OA after ACLR (30).

When compared to the general population KOOS scores, those in our cohort without early clinical knee OA, as defined by the modified Luyten et al criteria, exceeded uninjured, age-matched scores on each subscale (31). Those in our sample with early clinical OA, however, had lower mean values on every subscale than uninjured, age-matched population scores (31). These data indicate that use of the KOOS subscales may assist in differentiating those with clinically relevant knee symptoms. Compared to individuals who report their knee status as acceptable after ACLR (32), our group with early clinical knee OA had worse KOOS pain, ADL, sports and recreation, and quality of life scores. Patient acceptable symptom state (PASS), a state in which individuals consider the current state of their knee to be satisfactory, was established in a cross-sectional analysis of individuals 1–5 years after ACLR (mean ± SD follow-up time 3.4 ± 1.3 years) (32). The PASS threshold was 88.9 for KOOS pain, 57.1 for KOOS symptoms, 100.0 for KOOS ADL, 75.0 for KOOS sports and recreation, and 62.5 for KOOS quality of life. In our sample classified with early OA, only the KOOS symptoms subscale exceeded the PASS threshold, while KOOS pain, KOOS ADL, KOOS sports and recreation, and KOOS quality of life did not. Conversely, in those without early OA, only the KOOS ADL subscale did not exceed the PASS threshold (99.1 versus 100). Our results, taken with these data, suggest that poor scores on at least 2 of 4 KOOS subscales (pain, symptoms, activities of daily living, knee-related quality of life) may be an important identifier of those who may be at risk of early OA development. If we can identify individuals who are at risk earlier, we can intervene with secondary prevention strategies, such as quadriceps strengthening, and potentially delay irreversible joint damage.

There are limitations of this study to consider when interpreting our results. Due to the cross-sectional study design and our statistical model, we cannot be sure that restoring quadriceps strength symmetry will change the clinical or structural trajectory for developing OA after ACL rupture, and we cannot assume a linear relationship. Additionally, we only applied the patient-reported component of the Luyten et al (19) criteria to our sample. The full Luyten et al (19) criteria have not yet been validated. Therefore, we are unable to conclude that quadriceps strength asymmetry was a direct contributor to clinical OA. Further analysis is needed to examine the association of quadriceps strength with early clinical knee OA when the full criteria are applied. Strength testing differed slightly between sites; however, we reported limb symmetry measures to ensure that strength data were comparable. Caution should be used when interpreting limb symmetry indexes, as using the contralateral limb as a benchmark may overestimate knee function; therefore, we cannot draw definite conclusions on absolute strength. Future analysis should consider the association of quadriceps strength with the development of radiographic OA.

In conclusion, more symmetric quadriceps strength, but not surgical status, 5 years after ACL rupture was associated with lower odds of clinical knee OA. These results indicate that maintaining stronger quadriceps after ACL rupture, with or without reconstruction, could reduce the odds of clinical knee OA.

ACKNOWLEDGMENTS

The authors acknowledge Håvard Moksnes, Ingrid Eitzen, Annika Storevold, Ida Svege, Espen Selboskar, Karin Rydevik, Marte Lund, Kristine Roberg Hytten, Andrew Lynch, Mathew Failla, Airelle Giordano, and Angela Smith for their assistance with data collection for this study. The authors acknowledge the Norwegian Sports Medicine Clinic, Nimi (www.nimi.no), for supporting the Norwegian Research Center for Active Rehabilitation with rehabilitation facilities and research staff. Additionally, the authors acknowledge the University of Delaware Physical Therapy Clinic rehabilitation resources and staff, and Martha Callahan and the Delaware Rehabilitation Institute ResCore staff.

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. Arhos 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. Arhos, Thoma, Risberg, Snyder-Mackler.

Acquisition of data. Grindem, Logerstedt, Risberg, Snyder-Mackler.

Analysis and interpretation of data. Arhos, Thoma, Grindem, Risberg, Snyder-Mackler.

REFERENCES

1. Meunier A, Odensten M, Good L. Long-term results after primary repair or non-surgical treatment of anterior cruciate ligament rupture: a randomized study with a 15-year follow-up. Scand J Med Sci Sports 2007;17:230–7.

2. Lohmander LS, Englund PM, Dahl LL. The long-term consequence of anterior cruciate ligament and meniscus injuries. Am J Sports Med 2007;35:1756–69.
3. Øiestad BE, Holm I, Risberg MA. Return to pivoting sport after ACL reconstruction: association with osteoarthritis and knee function at the 15-year follow-up. Br J Sports Med 2018;52:1199–204.
4. Felson DT, Hodgson R. Identifying and treating preclinical and early osteoarthritis. Rheum Dis Clin North Am 2014;40:699–710.
5. Rudolph KS, Eastlack ME, Axe MJ, Snyder-Mackler L. 1998 Basmajian Student Award Paper: movement patterns after anterior cruciate ligament injury: a comparison of patients who compensate well for the injury and those who require operative stabilization. J Electromyogr Kinesiol 1998;8:349–62.
6. Wellsandt E, Gardinier ES, Manal K, Axe MJ, Buchanan TS, Snyder-Mackler L. Decreased knee joint loading associated with early knee osteoarthritis after anterior cruciate ligament injury. Am J Sports Med 2016;44:143–51.
7. Wellsandt E, Axe MJ, Snyder-Mackler L. Poor performance on single-legged hop tests associated with development of posttraumatic knee osteoarthritis after anterior cruciate ligament injury. Orthop J Sport Med 2018;6:1–9.
8. Patterson B, Culvenor AG, Barton CJ, Guermazi A, Stefanik J, Morris HG, et al. Poor functional performance 1 year after ACL reconstruction increases the risk of early osteoarthritis progression. Br J Sports Med 2020;54:546–55.
9. Øiestad BE, Holm I, Aune AK, Gunderson R, Myklebust G, Engebretsen L, et al. Knee function and prevalence of knee osteoarthritis after anterior cruciate ligament reconstruction: a prospective study with 10 to 15 years of follow-up. Am J Sports Med 2010;38:2201–10.
10. Øiestad BE, Juhl CB, Eitzen I, Thorlund JB. Knee extensor muscle weakness is a risk factor for development of knee osteoarthritis: a systematic review and meta-analysis. Osteoarthritis Cartilage 2015;23:171–7.
11. Palmieri-Smith RM, Lepley LK. Quadriceps strength asymmetry after anterior cruciate ligament reconstruction alters knee joint biomechanics and functional performance at time of return to activity. Am J Sports Med 2015;43:1662–9.
12. Lewek M, Rudolph K, Axe M, Snyder-Mackler L. The effect of insufficient quadriceps strength on gait after anterior cruciate ligament reconstruction. Clin Biomech (Bristol, Avon) 2002;17:56–63.
13. Norte GE, Hertel J, Saliba SA, Diduch DR, Hart JM. Quadriceps function and patient-reported outcomes after anterior cruciate ligament reconstruction in patients with or without knee osteoarthritis. J Athl Train 2018;53:965–75.
14. Ericsson YB, Roos EM, Owman H, Dahlberg LE. Association between thigh muscle strength four years after partial meniscectomy and radiographic features of osteoarthritis 11 years later. BMC Musculoskelet Disord 2019;20:1–10.
15. Grindem H, Snyder-Mackler L, Moksnes H, Engebretsen L, Risberg MA. Simple decision rules can reduce reinjury risk by 84% after ACL reconstruction: the Delaware-Oslo ACL cohort study. Br J Sports Med 2016;50:804–8.
16. Kyritsis P, Bahr R, Landreau P, Miladi R, Witvrouw E. Likelihood of ACL graft rupture: not meeting six clinical discharge criteria before return to sport is associated with a four times greater risk of rupture. Br J Sports Med 2016;50:946–51.
17. Adams D, Logerstedt D, Hunter-Giordano A, Axe MJ, Snyder-Mackler L. Current concepts for anterior cruciate ligament reconstruction: a criterion-based rehabilitation progression. J Orthop Sport Phys Ther 2012;42:601–14.
18. Schmitt LC, Paterno MV, Hewett TE. The impact of quadriceps femoris strength asymmetry on functional performance at return to sport following anterior cruciate ligament reconstruction. J Orthop Sports Phys Ther 2012;42:750–9.
19. Luyten FP, Bierma-Zeinstra SB, Accio FD, Kraus VB, Nakata K, Sekiya I, et al. Toward classification criteria for early osteoarthritis of the knee. Semin Arthritis Rheum 2018;47:457–63.
20. Hefti E, Müller W, Jakob RP, Stäubli HU. Evaluation of knee ligament injuries with the IKDC form. Knee Surg Sports Traumatol Arthrosc 1993;1:226–34.
21. Eitzen I, Moksnes H, Snyder-Mackler L, Risberg MA. A progressive 5-week exercise therapy program leads to significant improvement in knee function early after anterior cruciate ligament injury. J Orthop Sports Phys Ther 2010;40:705–21.
22. Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynnon BD. Knee Injury and Osteoarthritis Outcome Score (KOOS): development of a self-administered outcome measure. J Orthop Sports Phys Ther 1998;28:88–96.
23. Segal NA, Torner JC, Felson D, Niu J, Sharma L, Lewis CE, et al. Effect of thigh strength on incident radiographic and symptomatic knee osteoarthritis in a longitudinal cohort. Arthritis Rheum 2009;61:1210–7.
24. Pietrosimone BG, Thomas AC, Saliba SA, Ingersoll CD. Relation between quadriceps strength and physical activity in people with knee osteoarthritis. Osteoarthritis Cartilage 2013;21:S99.
25. Schmitt LC, Paterno MV, Ford KR, Myer GD, Hewett TE. Strength asymmetry and landing mechanics at return to sport after anterior cruciate ligament reconstruction. Med Sci Sports Exerc 2015;47:1426–34.
26. Smith TO, Postle K, Penny F, McNamara I, Mann CJ. Is reconstruction the best management strategy for anterior cruciate ligament rupture? A systematic review and meta-analysis comparing anterior cruciate ligament reconstruction versus non-operative treatment. Knee 2014;21:462–70.
27. Chalmers PN, Mall NA, Moric M, Sherman SL, Paletta GP, Cole BJ, et al. Does ACL reconstruction alter natural history? A systematic literature review of long-term outcomes. J Bone Joint Surg Am 2014;96:292–300.
28. Feucht MJ, Cotic M, Saier T, Minzlaff P, Plath JE, Imhoff AB, et al. Patient expectations of primary and revision anterior cruciate ligament reconstruction. Knee Surg Sports Traumatol Arthrosc 2016;24:201–7.
29. Hart HF, Culvenor AG, Guermazi A, Crossley KM. Worse knee confidence, fear of movement, psychological readiness to return-to-sport and pain are associated with worse function after ACL reconstruction. Phys Ther Sport 2020;41:1–8.
30. Hart HF, Collins NJ, Ackland DC, Crossley KM. Is self-reported knee stability associated with symptoms, function, and quality of life in people with knee osteoarthritis after anterior cruciate ligament reconstruction? Clin J Sport Med 2020;30:e134–8.
31. Paradowski PT, Bergman S, Sundén-Lundius A, Lohmander LS, Roos EM. Knee complaints vary with age and gender in the adult population: population-based reference data for the Knee Injury and Osteoarthritis Outcome Score (KOOS). BMC Musculoskelet Disord 2006;7:1–8.
32. Muller B, Yabroudi MA, Lynch A, Lai CL, van Dijk CN, Fu FH, et al. Defining thresholds for the patient acceptable symptom state for the IKDC subjective knee form and KOOS for patients who underwent ACL reconstruction. Am J Sports Med 2016;44:2820–6.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 3, March 2022, pp 392–402
DOI 10.1002/acr.24470

Costs of Joint Replacement in Osteoarthritis: A Study Using the National Joint Registry and Clinical Practice Research Datalink Data Sets

Jose Leal,[1] Jacqueline Murphy,[1] Cesar Garriga,[1] Antonella Delmestri,[1] Amar Rangan,[2] Andrew Price,[1] Andrew Carr,[1] Daniel Prieto-Alhambra,[3] and Andrew Judge[4]

Objective. To estimate the costs of primary hip and knee replacement in individuals with osteoarthritis up to 2 years postsurgery, compare costs before and after the surgery, and identify predictors of hospital costs.

Methods. Patients age ≥18 years with primary planned hip or knee replacements and osteoarthritis in England between 2008 and 2016 were identified from the National Joint Registry and linked with Hospital Episode Statistics data containing inpatient episodes. Primary care data linked with hospital outpatient records were also used to identify patients age ≥18 years with primary hip or knee replacements between 2008 and 2016. All health care resource use was valued using 2016/2017 costs, and nonparametric censoring methods were used to estimate total 1-year and 2-year costs.

Results. We identified 854,866 individuals undergoing hip or knee replacement. The mean censor-adjusted 1-year hospitalization costs for hip and knee replacement were £7,827 (95% confidence interval [95% CI] 7,813, 7,842) and £7,805 (95% CI 7,790, 7,818), respectively. Complications and revisions were associated with up to a 3-fold increase in 1-year hospitalization costs. The censor-adjusted 2-year costs were £9,258 (95% CI 9,233, 9,280) and £9,452 (95% CI 9,430, 9,475) for hip and knee replacement, respectively. Adding primary and outpatient care, the mean total hip and knee replacement 2-year costs were £11,987 and £12,578, respectively.

Conclusion. There are significant costs following joint replacement. Revisions and complications accounted for considerable costs and there is a significant incentive to identify best approaches to reduce these.

INTRODUCTION

Knee and hip replacement significantly improve the quality of life of individuals with osteoarthritis and have been shown to be very cost-effective compared to no surgery (1,2). In the UK, there were 96,117 primary hip procedures and 106,334 primary knee procedures in 2017, of which 90% and 99% of hip and knee replacements, respectively, had osteoarthritis as the indication for surgery (3).

There is limited evidence about the primary care and hospital costs of primary planned joint replacement in the subsequent years after surgery. It is important to have up-to-date and robust data of the costs of joint replacement and its drivers to inform decisions about changes in health service delivery and to produce good practice guidelines (4). Investment and disinvestment decisions regarding novel interventions in this area are driven by cost-effectiveness evidence (5,6), where resource use and costs are a key input.

The views expressed in this publication are those of the authors and do not necessarily reflect those of the NHS, the National Institute for Health Research (NIHR) or the Department of Health and Social Care, the National Joint Registry Steering Committee, the Healthcare Quality Improvement Partnership, or the Medicines and Healthcare Products Regulatory Agency.

Supported by the NIHR Health Services and Delivery Research program (#14/46/02). The work of Mr. Rangan and Mr. Carr was supported by the NIHR Oxford Biomedical Research Centre. Dr. Prieto-Alhambra's work was supported by the NIHR in the form of a Senior Research Fellowship (SRF-2018-11-ST2-004). Dr. Judge's work was supported by the NIHR Biomedical Research Centre at the University Hospitals Bristol NHS Foundation Trust and the University of Bristol.

[1]Jose Leal, DPhil, Jacqueline Murphy, MSc, Cesar Garriga, PhD, Antonella Delmestri, PhD, Andrew Price, BA, MBB, DPhil, Andrew Carr, ChM, FMedSci: University of Oxford, Oxford, UK; [2]Amar Rangan, FRCS: University of Oxford, Oxford, and National Joint Registry for England, Wales, Northern Ireland and the Isle of Man, Hemel Hempstead, UK; [3]Daniel Prieto-Alhambra, MD, PhD: University of Oxford and NIHR Oxford Biomedical Research Centre, John Radcliffe Hospital Oxford, Oxford, UK; [4]Andrew Judge, PhD: University of Oxford, Oxford, and University of Bristol, Bristol, UK.

Dr. Leal and Ms. Murphy contributed equally to this work.

Mr. Rangan has received grants from DePuy Ltd. Dr. Price has received consulting fees, speaking fees, and/or honoraria from Zimmer Biomet (more than $10,000). Dr. Prieto-Alhambra has received grants from Amgen, UCB Biopharma, Servier, Astellas, and Novartis. Dr. Judge has received consulting fees, speaking fees, and/or honoraria from Freshfields Bruckhaus Deringer and Anthera Pharmaceuticals (more than $10,000 each). No other disclosures relevant to this article were reported.

Address correspondence to Jose Leal, DPhil, Nuffield Department of Population Health, University of Oxford, Old Road Campus, OX3 7LF, UK. Email: jose.leal@dph.ox.ac.uk.

Submitted for publication February 4, 2020; accepted in revised form September 24, 2020.

SIGNIFICANCE & INNOVATIONS

- Joint replacement in osteoarthritis is associated with considerable health care costs and variation across surgery procedures.
- Revisions and complications were associated with up to a 3-fold increase in 1-year hospitalization costs.
- Costs in the second year after joint replacement were higher compared to costs in the year prior to surgery.

The primary aim of this study was to estimate the primary care and hospital costs of primary joint replacement up to 2 years postsurgery. We used data from the UK National Joint Registry linked with hospital data records in England and data from a large patient-level primary care data set representative of the English population. Secondly, we contrasted the resource use and costs by operation types. Finally, we reported the main predictors of hospital costs following joint replacement.

PATIENTS AND METHODS

Setting and data sources. We adopted an incidence-based approach (7) to estimate the primary and hospital care costs associated with hip and knee replacement. This approach estimated the costs of individuals from joint replacement backwards and forwards to the earliest and latest observed follow-up point, respectively.

Data from the UK National Joint Registry (NJR) were linked with Hospital Episode Statistics (HES), which contains records of all admitted patient care episodes undertaken in NHS trusts in England. NJR contains data on hip replacement surgeries from all hospitals in England and includes 2 million patients since 2003, currently covering 95% and 96% of primary hip and knee replacements, respectively (8).

Before personal data and sensitive personal data are recorded in NJR, express written patient consent is provided. With support under Section 251 of the NHS Act 2006, the ethics and confidentiality committee allows the NJR to collect patient data where consent is indicated as not recorded (Confidentiality Advisory Group [CAG] reference: PIAG 2-05(j)/2006). This study did not require ethical approval because it analyzed information previously collected in the course of normal care, and patients or service users were not identifiable to the research team carrying out the analysis (Medical Sciences Interdivisional Research Ethics Committee, University of Oxford; CAG reference: 16/CAG/0111). Planned hip and knee replacements in the HES data set were linked to patient-reported outcome measures, i.e., Oxford Hip Score/Oxford Knee Score (OHS/OKS) (9,10) and EuroQol 5-domain questionnaire in 3 levels (EQ-5D-3L) (11) before surgery and 6 months after surgery.

The Clinical Practice Research Datalink (CPRD) GOLD data set (obtained under license from the UK Medicines and Healthcare Products Regulatory Agency [MHRA]) contained data on patient consultations entered by the general practitioner, medical history, referrals data, test results, and all pharmaceutical prescriptions from general practitioner electronic health records. Hip and knee replacement were identified using predefined Read codes (see Supplementary Appendix A, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24470/abstract). The CPRD GOLD data set was linked to hospital outpatient records in HES and to Office for National Statistics mortality data. The study was approved by Independent Scientific Advisory Committee for MHRA Database Research (protocol number 11_050AMnA2RA2).

Study participants. To estimate hospitalization costs, we only included individuals identified in the NJR-HES–linked data set with a planned surgery for joint replacement between April 2008 and January 2017. Patients without a concordant date of replacement between NJR and HES databases were excluded from the analysis. To estimate outpatient and primary care costs, we only included patients in the CPRD GOLD data set with a first ever clinical or referral record of planned joint replacement occurring from April 1, 2008 until December 31, 2016.

Ascertainment of change in patient-reported outcomes at 6 months. We estimated the absolute change in OHS/OKS and EQ-5D-3L index scores (6 months–baseline score) to obtain a measure of change associated with the surgery. The scores from the 12 questions in the OHS/OKS were summed to obtain the total score spanning from 0 (worst possible) to 48 (best possible). The EQ-5D-3L responses were converted into utility scores using the UK value set (12). Higher positive values for OHS/OKS and EQ-5D-3L score changes between time points represented greater reduction in pain, improvement in function, and quality of life self-reported by the patient.

Ascertainment of death, complications, and revisions at 1 year. All-cause mortality was estimated at 1 year from the day of planned admission due to joint replacement and using the date of death from the Office for National Statistics mortality database. We defined postoperative complications as 1 or more events happening up to 1 year after joint replacement: stroke (excluding transient ischemic attack), respiratory infection, acute myocardial infarction, pulmonary embolism/deep vein thrombosis, urinary tract infection, wound disruption, surgical site infection, fracture after implantation, complication of prosthesis, neurovascular injury, acute renal failure, and blood transfusion. A group of 4 orthopedic surgeons independently went through all the relevant International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10) diagnoses and Office of Population Censuses and Surveys Classification of

Interventions and Procedures version 4 operation codes, and came to a consensus on the final list of codes for complications relevant to this study. This list of codes was further checked by a senior data manager (AD) who conducted additional searches based on the list of codes identified to ensure that no potential relevant codes had been missed (see Supplementary Appendix B, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24470/abstract). We also identified revisions occurring up to 1 year following joint replacement from revisions declared to the NJR registry by the surgeons (13) and revisions reported to HES using codes from Supplementary Appendix C, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24470/abstract.

Costs. Each finished consultant episode in a hospital admission was assigned into a Healthcare Resource Group (HRG) via the 2016/2017 Casemix Grouper Software (HRG4+) (14). HRGs are standard groups of clinically similar treatments that consume a common set of health care resources. HRGs for each finished consultant episode were valued using NHS reference costs from 2016/2017 (15) and appropriate methodology (16) and were summed to produce the total cost per hospital admission.

Primary care contacts and tests were costed using 2016/2017 unit costs from national cost databases (17) (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24470/abstract for full details of methodology). Pharmaceuticals were costed by matching each prescribed medication to a British National Formulary code and valuing these using 2016/2017 cost data from NHS digital prescription cost analysis (18) Total costs per patient were aggregated into monthly and annual amounts for the purposes of the analysis

Statistical analysis. The NJR-HES database was censored on January 20, 2017, and complete follow-up was not available for all cases. Hence, we report total hospital inpatient costs for those patients with complete follow-up data at years 1 and 2 following joint replacement and for the whole sample after adjusting for censoring using the methodology developed by Lin et al (19). Costs are reported as means together with their 95% confidence intervals (95% CIs), obtained from 1,000 bootstrap estimates.

Predictors of hospitalization costs of joint replacement were estimated using a generalized linear model (GLM). Based on our review of the literature, we examined the following predictors of costs in the year of the joint replacement: age, sex, EQ-5D-3L/OHS/OKS before surgery and change at 6 months, complications and revision up to 1 year after surgery, multiple deprivation index, Charlson comorbidity score up to surgery (see Supplementary Appendix D, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24470/abstract), body mass index (BMI) prior to surgery, type of joint replacement (partial, total, and patellofemoral for knee; resurfacing or total for hip), surgical variables, American Society of Anesthesiologists (ASA) grade before surgery, thrombolysis agents used (low molecular weight heparin, none, aspirin, and other), type of anesthesia (general, epidural, spinal, and nerve block), death, and year of surgery.

We used *t*-tests and Pearson's chi-square tests to evaluate the missingness for the potential predictors of costs (e.g., BMI, EQ-5D-3L/OHS/OKS scores before surgery and change at 6 months) in terms of age, sex, hospitalization costs, length of stay, Charlson comorbidity score, and type of joint replacement. We also performed multiple imputation of the missing data using a chained model with 20 iterations regressed on nonmissing variables to inform the prediction models (see Supplementary Appendix E, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24470/abstract for more details) (20).

The choice of the GLM model family and link functions was informed by the modified Park test and the Box-Cox test, respectively (21). We applied stepwise backward selection (at P value less than 0.05) per 300 bootstrap samples to identify variables that were consistently selected for at least 50% of the analyses and to inform the final models. A 2-tailed *t*-test with $\alpha = 0.01$ (to account for the large sample size) was used to determine whether each coefficient was statistically significantly different from zero, and their selection as predictors of costs was informed using Akaike's information criterion, mean square error, and likelihood test. All analyses were performed using Stata software, version 15.

RESULTS

Patient sample. Between April 1, 2008 and January 30, 2017, we identified 397,119 and 457,747 patients with osteoarthritis as having had a primary hip or knee replacement, respectively, in the NJR-HES–linked data set. Table 1 reports the baseline characteristics of the 2 cohorts. There were more women in the knee replacement cohort than in the hip replacement cohort (57.0% versus 40.4%). Individuals undergoing hip replacement were slightly younger (69.1 years versus 69.5 years) and with lower BMI (28.8 kg/m^2 versus 30.7 kg/m^2, i.e., overweight versus obese) compared to individuals undergoing knee replacement. Furthermore, the absolute change for Oxford and EQ-5D-3L scores was slightly lower in the hip replacement cohort (17.4 points versus 18.2 points for OHS/OKS and 0.33 versus 0.37 for EQ-5D-3L utilities). Osteoarthritis was the most common indication for joint replacement, with only 3.2% (hips) and 1.2% (knees) of cases having an indication other than osteoarthritis alone.

Patient outcomes and hospitalization costs. The mean ± SD duration of follow-up for the hip and knee replacement

Table 1. Patient characteristics of study cohorts at primary joint replacement*

Characteristic	Hip replacement (n = 397,119)	Knee replacement (n = 457,747)
Age, mean ± SD years	69.1 ± 10.8	69.5 ± 9.5
Female sex	40.4	57.0
White ethnicity†	86.1	82.4
Index of multiple deprivation, mean ± SD‡	18.0 ± 13.2	19.4 ± 14.0
Body mass index, mean ± SD§	28.8 ± 5.2	30.7 ± 5.4
Underweight (<18.5)	0.7	0.2
Normal (18.5–24.9)	19.4	9.7
Overweight (25–29.9)	39.8	34.4
Class I obese (30–34.9)	26.6	32.8
Class II obese (35–39.9)	10.1	16.4
Class III obese (≥40)	3.4	6.7
Oxford Hip/Knee Score before surgery, mean ± SD¶	17.4 ± 8.2	18.2 ± 7.8
EQ-5D-3L score before surgery, mean ± SD#	0.33 ± 0.32	0.37 ± 0.32
Location**		
Urban	71.3	74.7
Town and fringe	12.8	11.8
Village/isolated	15.9	13.5
Charlson comorbidity index, mean ± SD	0.37 ± 0.75	0.4 ± 0.8
Median (interquartile range)	0 (0–1)	0 (0–1)
ASA grade		
Fit and healthy (I)	13.8	10.2
Mild disease not incapacitating (II)	70.3	73.7
Incapacitating systemic disease (III)	15.5	15.9
Life-threatening disease or expected to die within 24 hours (IV and V)	0.4	0.3
Indication		
Osteoarthritis	96.8	98.8
Osteoarthritis and other	3.2	1.2
Operation type, no. (%)††		
Total joint replacement	381,145 (98.1)	418,510 (92.4)
Partial joint replacement	–	34,299 (7.8)
Patellofemoral joint replacement	–	4,939 (1.1)
Metal-on-metal resurfacing	7,271 (1.9)	–
Implant type‡‡		
Bicondylar	–	92.0
Metal-on-metal	4.6	–
Nonmetal-on-metal	95.4	–
Anesthesia§§		
General	38.9	35.4
Epidural	4.6	4.6
Nerve block	8.0	15.3
Spinal	71.0	68.5
Thromboprophylaxis for joint replacement		
None	3.1	3.7
Aspirin only	5.1	5.6
LMWH (with or without other)	66.0	72.3
Other (no LMWH)	25.8	18.4

* Values are the percentage unless indicated otherwise. ASA = American Society of Anesthesiologists; EQ-5D-3L = EuroQol 5-domain questionnaire in 3 levels; LMWH = low molecular weight heparin.
† 1.5% and 1.3% missing in the hip and knee replacement cohorts, respectively.
‡ 1.1% and 1.0% missing in the hip and knee replacement cohorts, respectively.
§ 29.0% and 29.1% missing in the hip and knee replacement cohorts, respectively.
¶ 41.2% and 45.3% missing in the hip and knee replacement cohorts, respectively.
# 41.9% and 45.9% missing in the hip and knee replacement cohorts, respectively.
** 0.3% missing in each cohort.
†† 2.2% missing in each cohort.
‡‡ 1.3% missing in each cohort.
§§ 0.5% missing in each cohort.

cohorts was 3.9 ± 2.5 years (Table 2). The mean ± SD difference between 6 months and preoperative OHS/OKS was 20.1 ± 10.2 points (n = 202,761) for hip replacement and 15.3 ± 10.0 points (n = 216,322) for knee replacement.

The mean hospitalization costs associated with index admission for hip replacement were £6,208 (median 5,824; SD 969) compared to £6,122 (median 5,692; SD 967) for knee replacement. Mean length of stay in the index admission was 4.8 days

Table 2. Patient outcomes and hospitalization costs*

	Hip replacement	Knee replacement
Follow-up time, years	3.9 ± 2.5	3.9 ± 2.5
Mortality within 1 year, no. (%)†	4,071 (1.2)	2,965 (0.8)
Initial hospitalization (index admission to discharge)‡		
Hospital length of stay, days	4.8 ± 3.8	4.8 ± 3.5
Hospitalization costs, £	6,208 ± 969	6,122 ± 967
Oxford Hip/Knee Score change at 6 months§	20.1 ± 10.2	15.3 ± 10.0
EQ-5D-3L score change at 6 months¶	0.40 ± 0.34	0.29 ± 0.33
Hospitalization costs within 1 year of replacement, £†		
Index hospitalization	6,207 ± 990	6,110 ± 979
Emergency hospitalizations after discharge	648 ± 2,880	606 ± 2,730
Planned hospitalizations after discharge	963 ± 2,825	1,067 ± 2,850
Total	7,817 ± 4,618	7,784 ± 4,520
Total length of hospital stay within 1 year of replacement, days†		
Index hospitalization	4.9 ± 3.8	4.8 ± 3.5
Emergency hospitalizations after discharge	1.4 ± 7.4	1.4 ± 7.2
Planned hospitalizations after discharge	0.9 ± 5.3	1.0 ± 5.5
Total	7.3 ± 11.2	7.2 ± 11.2
Hospitalization costs within year 2 after joint replacement, £#		
Emergency hospitalizations	524 ± 2,598	549 ± 2,692
Planned hospitalizations	908 ± 2841	1,090 ± 3,020
Total costs	1,432 ± 4,169	1,639 ± 4,353
Total length of hospital stay within year 2 after joint replacement#	1.9 ± 9.1	2.1 ± 9.5

* Values are the mean ± SD unless indicated otherwise. EQ-5D-3L = EuroQol 5-domain questionnaire in 3 levels.
† 344,721 and 394,118 individuals with complete follow-up, including those who died in that year, in the hip and knee cohorts, respectively.
‡ 397,119 and 457,747 individuals in the hip and knee replacement cohorts, respectively.
§ 202,761 and 216,322 individuals with presurgery and 6 months Oxford Hip Score/Oxford Knee Score in the hip and knee replacement cohorts, respectively.
¶ 187,636 and 201,077 individuals with presurgery and 6 months EQ-5D-3L scores in the hip and knee replacement cohorts, respectively.
# 293,618 and 333,123 individuals with complete follow-up, including those who died within 2 years of hip and knee replacement, respectively.

(median 4; SD 3.8; interquartile range [IQR] 3–6) and 4.8 days (median 4; SD 3.5; IQR 3–5) for hip and knee replacement, respectively.

Within 1 year of joint replacement, the mean hospitalization costs were estimated at £7,817 (median 6,258; SD 4,618) and £7,784 (median 6,226; SD 4,520) for hip and knee replacement, respectively, of which the index admission accounted for 79.4% and 78.5% of the total. Hospitalization costs and length of stay within 1 year were highly correlated for both types of joint replacement (Spearman's correlation coefficient 0.84, $P < 0.001$).

The 3 most common reasons for hospital readmission within the first year of joint replacement were similar in both cohorts: musculoskeletal (ICD-10 chapter 13: 32–35% of readmission costs), injury (ICD-10 chapter 19: 21%), and circulatory system (ICD-10 chapter 9: 8–9%) (see Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24470/abstract). For hip replacement, 2,404 patients (0.7%) with complete 1-year follow-up had a hip revision in the following year according to the NJR registry. We found 610 more 1-year revisions using HES, giving a total of 3,014 (0.9%). For knee replacement, 1,769 patients (0.5%) with complete 1-year follow-up had a knee revision in the following year according to the NJR registry. We found 178 more 1-year revisions using HES, giving a total of 1,947 (0.5%).

For hip replacement, individuals undergoing metal-on-metal resurfacing had on average lower 1-year and 2-year costs and length of stay (at 2 years: mean ± SD £7,374 ± 4,246 and 5.6 ± 7.8 days [n = 6,643]) compared to individuals undergoing total hip replacement (at 2 years: £9,321 ± 6,971 and 9.5 ± 16.5 days [n = 286,975]) (see Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24470/abstract for full details of methodology).

For knee replacement, individuals undergoing unicondylar joint replacement had on average lower 1-year and 2-year costs and length of stay (at 2 years: mean ± SD £8,198 ± 5,145 and 5.6 ± 9.5 days [n = 24,203]) compared to individuals undergoing patellofemoral joint replacement (at 2 years: £9,209 ± 6,252 and 7.4 ± 12.7 days [n = 3,726]) and total knee replacement (at 2 years: £9,548 ± 7,088 and 9.8 ± 17.0 days [n = 305,194]) (see Supplementary Table 4, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24470/abstract).

Adjusting for censoring, the mean 1-year costs were similar to the complete follow-up analysis (including individuals who died in that year) at £7,827 (95% CI 7,813, 7,842) and £7,805 (95% CI 7,790, 7,818) for hip and knee replacement, respectively. For hip replacement, the mean costs in the first 2 years following joint

replacement (2-year) adjusted for censoring were £9,258 (95% CI 9,233, 9,280) compared to £9,277 using only individuals with complete follow-up (including those who died in that year [n = 293,618]). For knee replacement, the costs in the first 2 years following joint replacement (2-year) adjusted for censoring were £9,452 (95% CI 9,430, 9,475) and similar to £9,446 using only individuals with complete follow-up (n = 333,123). Supplementary Table 5, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24470/abstract, reports hospital admissions, length of stay, and costs during the first 2 years following joint replacement.

Predictors of hospitalization costs in the first year following joint replacement. Approximately 50% and 70% of patients had missing data for Oxford and EQ-5D-3L scores (before surgery and at 6 months), BMI, or other variables to inform the prediction of hospitalization costs for hip and knee replacement, respectively (see Supplementary Table 6, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24470/abstract). Following multiple imputation, the predictors of hospitalization costs for hip and knee replacement are shown in Tables 3 and 4, respectively. A GLM model with gamma family and identity link function had the best fit.

Adjusting for all covariates, conventional total hip replacement was more expensive on average than metal-on-metal resurfacing (£451; $P < 0.001$). Women had higher mean hospitalization costs than men (£167; $P < 0.001$), and older and more deprived individuals were associated with higher costs. Individuals with higher quality of life values (EQ-5D-3L and OHS) prior to surgery and reporting improvements at 6 months were associated with lower hospitalization costs. There was strong evidence ($P < 0.01$) that ceramic on ceramic, ceramic on metal, and metal on ceramic bearing were

Table 3. Predictors of 1-year hospitalization costs following hip replacement (n = 330,765)*

	Value	Mean additional cost, £ (95% CI)	$P > z$
Type of joint replacement			
Total hip replacement	97.9	Ref.	
Metal-on-metal resurfacing	2.1	–451 (–556, –347)	<0.001
Age at replacement (centered at 69 years)	69.1†	28 (27, 30)	<0.001
Age at replacement squared, years	–	0.9 (0.8, 0.9)	<0.001
Sex			
Male	59.5	Ref.	
Female	40.5	167 (147, 188)	<0.001
Charlson comorbidity score	0.4†	380 (362, 399)	<0.001
Body mass index at hip replacement	28.8†	–4 (–6, –1)	0.002
EQ-5D-3L score at baseline (per 0.10 increase)	0.3†	–105 (–113, –96)	<0.001
EQ-5D-3L score change at 6 months (per 0.10 increase)	0.4†	–97 (–104, –9)	<0.001
Hip score at baseline	17.4†	–30 (–32, –27)	<0.001
Hip score change at 6 months	20.1†	–17 (–19, –14)	<0.001
Calendar year of replacement (centered at 2012)		–31 (–35, –26)	<0.001
ASA grade			
Fit and healthy (I)	13.9	–150 (–174, –126)	<0.001
Mild disease not incapacitating (II)	70.4	Ref.	
Incapacitating systemic disease (III)	15.3	637 (600, 675)	<0.001
Life-threatening disease or expected to die within 24 hours (IV and V)	0.4	2,112 (1,772, 2,452)	<0.001
Head size, mm			
≤28	42.2	Ref.	
29–35	31.4	45 (22, 69)	<0.001
36–42	23.4	56 (27, 85)	<0.001
43–48	1.4	29 (–72, 129)	0.579
49–52	1.2	66 (–59, 191)	0.300
≥53	0.4	226 (60, 392)	0.008
Bearing surfaces			
Metal on polyethylene	61.9	Ref.	
Metal on metal	4.3	–29 (–105, 47)	0.450
Ceramic on ceramic	16.9	–40 (–69, –10)	0.009
Ceramic on polyethylene	16.6	–24 (–51, 4)	0.094
Other (ceramic on metal or metal on ceramic)	0.4	–194 (–324, –64)	0.003
Surgeon volume of hip procedures (per 100 additional procedures)	97.4†	–16 (–28, –4)	0.007
Complications within 1 year	6.0	6,601 (6,472, 6,731)	<0.001
Revision within 1 year	0.9	11,255 (10,800, 11,709)	<0.001
Death	1.0	4,682 (4,374, 4,991)	<0.001
Constant	–	8,600 (8,500, 8,700)	<0.001

* Values are the percentage unless indicated otherwise. 95% CI = 95% confidence interval; ASA = American Society of Anesthesiologists; EQ-5D-3L = EuroQol 5-domain questionnaire in 3 levels; Ref. = reference.
† Mean value of variable untransformed.

associated with lower mean 1-year hospitalization costs than metal on polyethylene bearings (the most common bearing type in the cohort). Costs were also lower in recent years (–£31 per year; $P < 0.001$), holding all else constant. Complications and revisions within the year were significantly associated with higher mean costs, with an additional £6,601 (1.9-fold increase; $P < 0.001$) and £11,255 (2.5-fold increase; $P < 0.001$), respectively, and £17,857 together (3.4-fold increase). Holding all else constant, the complications associated with the highest increase in 1-year costs were blood transfusion (an additional £7,782), surgical site infection (£6,799), stroke (£6,791), fracture after implant (£6,585), and wound disruption (£6,209) (see Supplementary Table 7, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24470/abstract).

Adjusting for all covariates, total knee replacement was significantly associated with higher 1-year hospitalization costs than

Table 4. Predictors of 1-year hospitalization costs following knee replacement (n = 391,691)*

	Value	Mean additional cost, £ (95% CI)	*P* > z
Type of joint replacement			
Total knee replacement	92.5	Ref.	<0.001
Partial knee replacement	7.4	–404 (–443, –366)	<0.001
Patellofemoral replacement	0.1	–137 (–237, –38)	0.007
Age at replacement (centered at 69 years)	69.5†	31 (30, 32)	<0.001
Age at replacement squared	–	1.1 (1.0, 1.2)	<0.001
Sex			
Male	56.8	Ref.	
Female	43.2	255 (234, 277)	<0.001
Charlson comorbidity score	0.4	364 (348, 380)	<0.001
Year of surgery (centered in 2012)		–14 (–19, –10)	<0.001
IMD score (divided by 100)	19.4†	–276 (–350, –203)	<0.001
EQ-5D-3L score at baseline (per 0.10 increase)	0.4†	–97 (–105, –89)	<0.001
EQ-5D-3L score change at 6 months (per 0.10 increase)	0.3†	–94 (–101, –87)	<0.001
Knee score at baseline	18.2†	–28 (–30, –25)	<0.001
Knee score change at 6 months	15.2†	–10 (–12, –8)	<0.001
ASA grade			
Mild disease not incapacitating (II)	73.7	Ref.	
Fit and healthy (I)	10.2	–153 (–184, –121)	<0.001
Incapacitating systemic disease (III)	15.7	617 (585, 549)	<0.001
Life-threatening disease or expected to die within 24 hours (IV and V)	0.3	1,605 (1,346, 1,863)	<0.001
Deformity degrees			
<10	65.2	Ref.	
10–30	33.7	61 (38, 85)	<0.001
>30	1.1	507 (396, 618)	<0.001
Range of flexion, degrees			
91–110	45.3	Ref.	
<70	2.1	93 (15, 171)	0.027
70–90	19.7	56 (26, 86)	<0.001
>110	32.9	–15 (–41, 11)	0.238
Type of surgeon			
Consultant	78.5	Ref.	
Other	21.5	54 (29, 78)	<0.001
Approach			
Medial parapatellar	93.0	Ref.	
Lateral parapatellar	1.0	175 (72, 279)	0.001
Midvastus	3.1	30 (–26, 87)	0.295
Subvastus	1.2	138 (43, 232)	0.004
Other	1.7	–20 (–96, 56)	0.603
Type of fixation			
Cemented	95.0	Ref.	
Uncemented	4.2	–71 (–119, –22)	0.004
Hybrid	0.7	54 (–67, 175)	0.382
General anesthesia	36.5	77 (56, 87)	<0.001
Complications within 1 year	6.0	6,220 (6,139, 6,301)	<0.001
Revision within 1 year	0.5	10,406 (10,012, 10,799)	<0.001
Death	0.8	4,622 (4,390, 4,854)	<0.001
Constant	–	8,152 (8,094, 8,210)	<0.001

* Values are the percentage unless indicated otherwise. 95% CI = 95% confidence interval; ASA = American Society of Anesthesiologists; EQ-5D-3L = EuroQol 5-domain questionnaire in 3 levels; IMD = Index of Multiple Deprivation; Ref. = reference.
† Mean value of variable untransformed.

unicondylar knee replacement (£404; $P < 0.001$). Women had higher mean hospitalization costs than men (£255; $P < 0.001$), and costs increased with age (£31 per additional year; $P < 0.001$) and higher deprivation. Individuals with higher quality of life values (EQ-5D-3L and OKS) at baseline and those reporting improvements at 6 months had lower hospitalization costs. Higher deformity and lower range of flexion were also significantly associated with higher costs. Costs were also lower in recent years (–£14 per year; $P < 0.001$), holding all else constant. Complications and revisions within the first year were significantly associated with higher costs, with an additional £6,220 (1.8-fold increase; $P < 0.001$) and £10,406 (2.3-fold increase; $P < 0.001$), respectively, and £16,626 together (3.0-fold increase). Holding all else constant, the knee surgery complications associated with the highest increase in 1-year hospitalization costs were fracture after implant (an additional £9,875), blood transfusion (£7,691), stroke (£6,749), wound disruption (£6,889), and urinary tract infection (£6,529) (see Supplementary Table 7, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24470/abstract).

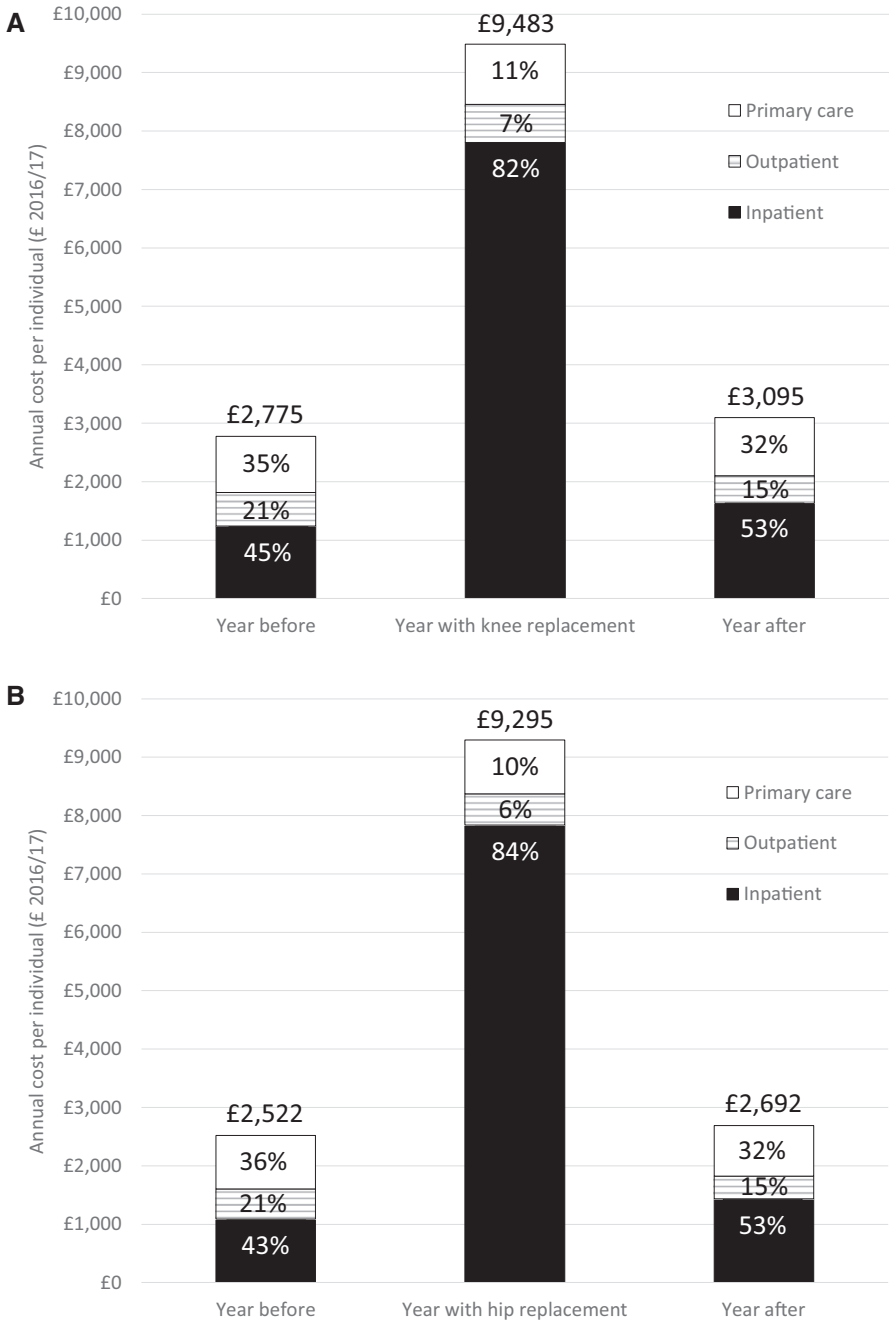


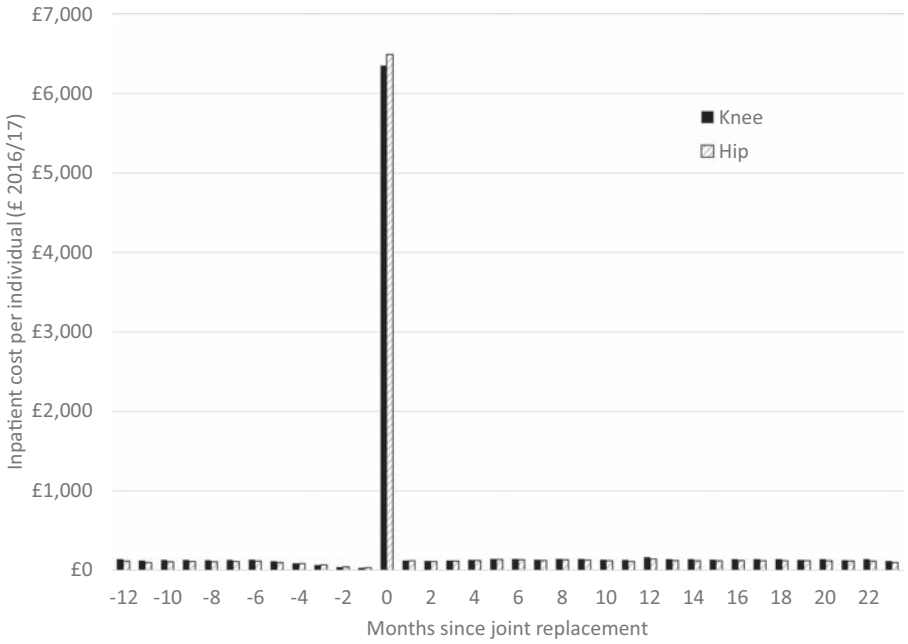
Figure 1. Costs in the months before and after knee replacement (**A**) and hip replacement (**B**), with complete cases, including those who died in that year.

For completeness, Supplementary Tables 8 and 9, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24470/abstract, report the predictors of 1-year hospitalization costs for hip and knee replacement, respectively, using only the subgroups of individuals with no missing data (complete cases). The results were similar in terms of direction and magnitude of the associations between hospitalization costs and covariates. The cohorts with complete data had lower mortality rates at 1 year and lower hospitalization costs compared to cases with missing data (see Supplementary Tables 10 and 11, available on the Arthritis Care & Research website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24470/abstract).

Costs before and after joint replacement. Adding primary, outpatient, and inpatient hospitalization costs, the mean costs associated with hip replacement amounted to £9,295 in the year of surgery compared to £9,483 following knee replacement (Figure 1). Hospitalization costs accounted for the highest proportion of the total 1-year cost for both hip and knee replacement (82–84%).

Using the annual number of UK joint replacements in 2017, the NHS primary and hospital costs were estimated at £899 million (n = 96,717) and £1,008 million (n = 106,334) in the year of the hip and knee replacement, respectively. In the second year after joint replacement, total costs were £2,692 for hip and £3,095 for knee replacement cohorts, with inpatient costs being the largest component (53% for both knee and hip).

Figure 2 reports the hospitalization costs in the months before and after joint replacement. The annual hospitalization costs in the year of joint replacement were £6,753 (95% CI 6,732, 6,774) and £6,563 (95% CI 6,544, 6,583) higher for hip and knee replacement, respectively, compared to that of the previous year. However, there was a decrease in hospitalization costs in the 5 months prior to surgery, reflecting lower hospital admissions leading up to the planned admission. Costs in the second year after joint replacement were £389 (95% CI 370, 407) and £349 (95% CI 329, 368) higher compared to costs



	Year before	Year 1	Year 2
Knee*	£1,240	£7,803	£1,628
Difference vs. year before	-	£6,563 (6,544 to 6,583)	£389 (370 to 407)
	N=297,794	N=297,794	N=297,794
Hip*	£1,085	£7,838	£1,434
Difference vs. year before	-	£6,753 (6,732 to 6,774)	£349 (329 to 368)
	N=257,563	N=257,563	N=257,563

Figure 2. Hospitalization costs in the months before and after joint replacement. Columns for Year 1 and Year 2 show the value (95% confidence interval). * = complete cases, including those who died in that year.

in the year prior to surgery for knee and hip replacement, respectively.

A similar pattern was observed with primary care and outpatient costs (see Supplementary Figures 1 and 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24470/abstract). However, outpatient costs in the second year after surgery were significantly lower than in the year preceding the surgery for both types of joint replacement (–£105 [95% CI –78, –133] and –£126 [95% CI –109, –143] for knee and hip, respectively). In contrast, primary care costs were lower in the second year after surgery for hip replacement (by –£53) but higher for knee replacement (by £37) compared to the year preceding surgery (see Supplementary Figures 1 and 2).

DISCUSSION

In this study, we estimated the immediate- and medium-term (up to 2 years) hospital and primary care costs of joint replacement compared with costs in the year prior to surgery in a large representative sample of patients in England, and explored the main variables influencing these costs. We also identified revisions and complications within the first year of joint replacement as major drivers of hospitalization costs, accounting for up to a 3-fold increase in costs.

Previous studies have examined the costs of joint replacement but consisted of smaller samples and without linkage to NJR data (22,23). We were able to examine the hospitalization costs of different types of joint replacement through their identification in the NJR data set and linkage to hospital records. We found unicondylar knee replacement to have lower 1-year costs than total knee replacement, and metal-on-metal resurfacing also had lower costs than conventional total hip replacement, even after adjusting for potential confounders. However, these cost differences could be offset with longer follow-up than 2 years if revision rates are observed to be relatively higher with unicondylar and metal-on-metal resurfacing.

We assessed hospitalization costs by month before and after surgery and identified a reduction in hospitalization costs in the 5 months prior to surgery for both types of joint replacement, reflecting fewer hospitalizations leading up to the planned admission. Furthermore, primary care costs were slightly lower in the second year after surgery for hip replacement but slightly higher for knee replacement compared to the year preceding surgery, possibly reflecting differences in recovery times between the 2 procedures.

Overall, we also found the predictors of costs to be similar for hip and knee replacement. Consistent with previous work (22), we found preoperative quality of life, as measured using Oxford and EQ-5D-3L scores, to be associated with hospitalization costs; 1-year costs were higher for individuals with worse preoperative quality of life even after adjusting for other covariates. Also, 1-year costs were lower for individuals reporting larger improvements in quality of life at 6 months.

Knee and hip replacement costs are significant, but these are very cost-effective procedures compared to no joint surgery in individuals with osteoarthritis (1,2,24–26). There is then an economic incentive to fund research aimed at identifying cost-effective ways of further improving the quality of life of patients with osteoarthritis following joint replacement and reducing the risk of revisions and complications.

A key advantage of this study was the use of the NJR data set, which is the largest arthroplasty data set in the world, linked to hospital care data and supplemented with a large primary care data set. Hence, our data are representative of the range of individuals with osteoarthritis undergoing joint replacement in England and are generalizable for use in other similar health care systems. However, our study had some limitations. NJR data were obtained for individuals undergoing joint replacement with osteoarthritis as an indication for surgery. Hence, individuals without osteoarthritis as one of the indications were not available for analysis, e.g., rheumatoid arthritis or fractures. Furthermore, private joint replacements were not available in the hospital care data set and we were not able to relate hospital readmissions to joint replacement. However, the majority of costs following joint replacement were associated with readmissions due to musculoskeletal and injury reasons (53–56% of all readmission costs). The study also lacked a control group, and additional costs associated with joint replacement were estimated by comparing the costs in the year of replacement with those in the previous year. Another potential limitation of the analysis is the use of HRGs and reference costs as opposed to detailed microcosting approaches to estimate hospitalization costs. HRGs and reference costs are nationally representative but may lack the precision to capture changes in resource use across individuals within the same HRG. To mitigate these issues, we followed best practice to ensure that all hospital contacts were captured and costed appropriately (16,27).

Finally, a large proportion of the cohort had missing data for 1 or more key covariates of the hospitalization costs, in particular EQ-5D-3L/OHS responses and BMI, which necessitated the use of missing data methods, specifically multiple imputation. A key assumption using multiple imputation was that the missing data were missing at random; that is, the missingness can be adjusted for (i.e., explained) using the observed data. This assumption is always untestable, but due to the large number of relevant covariates in our linked data, we judged it to be reasonable in this case. For completeness, we also present the results of the analysis using complete cases in the supplementary information, which we found to be similar to the findings following multiple imputation.

In conclusion, our results show the impact of hip and knee replacement on primary and hospital care and its predictors in England. We highlight the differences in costs between the types

of replacement and the significant impact of revisions and complications in individuals with osteoarthritis. Our results can be used as inputs in future work assessing the cost and cost-effectiveness of hip and knee replacement, and in particular to explore heterogeneity between patient subgroups. Our findings will be useful to commissioners, providers, and researchers interested in the prevention and management of osteoarthritis.

ACKNOWLEDGMENTS

We thank the patients and staff of all the hospitals in England, Wales, and Northern Ireland who have contributed data to the National Joint Registry. We are grateful to the Healthcare Quality Improvement Partnership and the NJR Research Committee and staff at the NJR Centre for facilitating this work. The authors have conformed to the NJR's standard protocol for data access and publication.

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. Leal 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. Leal, Price, Carr, Prieto-Alhambra, Judge.

Acquisition of data. Leal, Delmestri, Rangan, Judge.

Analysis and interpretation of data. Leal, Murphy, Garriga, Judge.

REFERENCES

1. Dakin H, Gray A, Fitzpatrick R, MacLennan G, Murray D. Rationing of total knee replacement: a cost-effectiveness analysis on a large trial data set. BMJ Open 2012;2:e000332.
2. Fordham R, Skinner J, Wang X, Nolan J. The economic benefit of hip replacement: a 5-year follow-up of costs and outcomes in the Exeter Primary Outcomes Study. BMJ Open 2012;2:e000752.
3. National Joint Registry. 15th annual report 2018: surgical data to 31 December 2017. URL: https://reports.njrcentre.org.uk/.
4. Luchinskaya D, Simpson P, Stoye G. UK health and social care spending. URL: https://www.ifs.org.uk/uploads/publications/budgets/gb2017/gb2017ch5.pdf.
5. National Institute for Health and Care Excellence. Total hip replacement and resurfacing arthroplasty for end-stage arthritis of the hip. URL: https://www.nice.org.uk/guidance/ta304.
6. National Institute for Health and Care Excellence. Guideline scope: hip, knee and shoulder joint replacement. URL: https://www.nice.org.uk/guidance/gid-ng10084/documents/draft-scope.
7. Lipscomb J, Yabroff KR, Brown ML, Lawrence W, Barnett PG. Health care costing: data, methods, current applications. Med Care 2009;47:S1–6.
8. National Joint Registry. 14th annual report: surgical data to 31 December 2016. URL: https://reports.njrcentre.org.uk/.
9. Dawson J, Fitzpatrick R, Carr A, Murray D. Questionnaire on the perceptions of patients about total hip replacement. J Bone Joint Surg Br 1996;78:185–90.
10. Dawson J, Fitzpatrick R, Murray D, Carr A. Questionnaire on the perceptions of patients about total knee replacement. J Bone Joint Surg Br 1998;80:63–9.
11. EuroQol Group. EuroQol: a new facility for the measurement of health-related quality of life. Health Policy 1990;16:199–208.
12. Dolan P. Modeling valuations for EuroQol health states. Med Care 1997;35:1095–108.
13. National Joint Registry. Data collection forms: revision procedure for hip. 2014. URL: http://www.njrcentre.org.uk/njrcentre/Portals/0/Documents/England/Data%20collection%20forms/MDSv6.0_H2_V004.pdf.
14. NHS Digital. HRG4+ 2016/17 reference costs grouper. URL: https://digital.nhs.uk/services/national-casemix-office/downloads-groupers-and-tools/grouper-and-tools-archive/costing-hrg4-2016-17-reference-costs-grouper.
15. NHS Improvement. Reference costs 2016/17: national schedule of reference costs 1. URL: https://www.england.nhs.uk/national-cost-collection/.
16. Leal J, Manetti S, Buchanan J. The impact of hospital costing methods on cost-effectiveness analysis: a case study. Pharmacoeconomics 2018;36:1263–72.
17. Curtis L, Burns A. Unit costs of health and social care 2017. URL: https://doi.org/10.22024/UniKent/01.02/65559.
18. NHS Digital. Prescription cost analysis: England 2017. URL: https://digital.nhs.uk/data-and-information/publications/statistical/prescription-cost-analysis/prescription-cost-analysis-england-2017.
19. Lin DY, Feuer EJ, Etzioni R, Wax Y. Estimating medical costs from incomplete follow-up data. Biometrics 1997;53:419–34.
20. Faria R, Gomes M, Epstein D, White IR. A guide to handling missing data in cost-effectiveness analysis conducted within randomised controlled trials. Pharmacoeconomics 2014;32:1157–70.
21. Deb P, Norton EC. Modeling health care expenditures and use. Annu Rev Public Health 2018;39:489–505.
22. Eibich P, Dakin HA, Price AJ, Beard D, Arden NK, Gray AM. Associations between preoperative Oxford hip and knee scores and costs and quality of life of patients undergoing primary total joint replacement in the NHS England: an observational study. BMJ Open 2018;8:e019477.
23. Burn E, Edwards CJ, Murray DW, Silman A, Cooper C, Arden NK, et al. Trends and determinants of length of stay and hospital reimbursement following knee and hip replacement: evidence from linked primary care and NHS hospital records from 1997 to 2014. BMJ Open 2018;8:e019146.
24. Jenkins PJ, Clement ND, Hamilton DF, Gaston P, Patton JT, Howie CR. Predicting the cost-effectiveness of total hip and knee replacement: a health economic analysis. Bone Joint J 2013;95-B:115–21.
25. Ferket BS, Feldman Z, Zhou J, Oei EH, Bierma-Zeinstra SM, Mazumdar M. Impact of total knee replacement practice: cost effectiveness analysis of data from the Osteoarthritis Initiative. BMJ 2017;356:j1131.
26. Rasanen P, Paavolainen P, Sintonen H, Koivisto AM, Blom M, Ryynanen OP, et al. Effectiveness of hip or knee replacement surgery in terms of quality-adjusted life years and costs. Acta Orthop 2007;78:108–15.
27. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013. URL: https://www.nice.org.uk/process/pmg9/chapter/foreword.

Arthritis Care & Research
Vol. 74, No. 3, March 2022, pp 403–409
DOI 10.1002/acr.24484

American College of Rheumatology
Empowering Rheumatology Professionals

Association Between the Severity of Periodontitis and Osteoarthritis in Middle-Aged and Older Patients With Type 2 Diabetes Mellitus: A Nationwide Population-Based Study

Hyoung-Sik Kim,[1] Hye-Min Park,[2] Haeyoung Kim,[3] Hye Sun Lee,[1] Da-Hye Son,[1] and Yong-Jae Lee[1]

Objective. Periodontitis and osteoarthritis are major public health concerns that result in decreased quality of life among middle-aged and older adults. We sought to examine whether the severity of periodontitis is related to osteoarthritis according to the presence of type 2 diabetes mellitus.

Methods. This study included 3,527 participants age ≥50 years from the Korean National Health and Nutrition Examination Survey. Periodontitis was assessed using the Community Periodontal Index; severe periodontitis was defined as periodontal tissue forming deep periodontal pockets ≥6-mm depth. Osteoarthritis was defined as Kellgren/Lawrence grade ≥2 on radiographic images of the knee or hip area with joint pain. The odds ratios (ORs) and 95% confidence intervals (95% CIs) for osteoarthritis according to the severity of periodontitis, stratified by type 2 diabetes mellitus, were calculated using multiple logistic regression analyses.

Results. Participants with type 2 diabetes mellitus were more likely to have osteoarthritis as the severity of periodontitis increased (nonsevere periodontitis OR 1.23 [95% CI 0.67–2.32]; severe periodontitis OR 3.01 [95% CI 1.51–5.84]) after adjusting for age, sex, body mass index, smoking status, alcohol consumption, regular exercise, education level, household income, hypertension, and frequent tooth-brushing. However, this positive association was not found in individuals without type 2 diabetes mellitus after adjusting for the same covariables.

Conclusion. Severe periodontitis was positively and significantly associated with osteoarthritis in middle-aged and older individuals with type 2 diabetes mellitus. Our findings suggest that the oral inflammation manifesting in periodontitis may be at least partly involved in the pathogenesis of osteoarthritis, particularly in patients with type 2 diabetes mellitus.

INTRODUCTION

Periodontitis is characterized by gingival inflammation accompanied by the loss of periodontal connective tissue. Periodontitis is a major oral disease worldwide and is known to decrease chewing function and quality of life, leading to eventual tooth loss with increased health care utilization (1). Emerging studies have suggested that periodontitis is not solely the result of local oral inflammation but is also due to systemic inflammation related to atherosclerotic cardiovascular disease and type 2 diabetes mellitus (2). In South Korea, the prevalence of periodontal disease was 23.9% among individuals age ≥19 years in 2010 according to the Korea Centers for Disease Control and Prevention (3). Regarding the severity of periodontitis, a large-sample study concluded that approximately 8.9% of adults in the US have severe periodontitis (4), which is more prevalent in individuals with type 2 diabetes mellitus than those without (5).

Osteoarthritis is characterized by the gradual loss of articular cartilage and secondary subchondral bone changes, particularly in the weight-bearing joints, owing to multifactorial pathophysiology (6). According to research from the Centers for Disease Control and Prevention, approximately 240 million people worldwide are affected by osteoarthritis, including >30 million in the US (7). Data from the national surveys in South Korea reported that the prevalence of knee osteoarthritis in people age ≥50 years according to sex was 4.4% in men and 19.2% in women,

[1]Hyoung-Sik Kim, MD, PhD, Hye Sun Lee, PhD, Da-Hye Son, MD, Yong-Jae Lee, MD, MPH, PhD: Yonsei University College of Medicine, Seoul, Republic of Korea; [2]Hye-Min Park, MD: Cha University and Yonsei University, Seoul, Republic of Korea; [3]Haeyoung Kim, MD: Yonsei University College of Medicine and ChunLab, Inc., Seoul, Republic of Korea.

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

Address correspondence to Yong-Jae Lee, MD, MPH, PhD, Professor, Department of Family Medicine, Yonsei University College of Medicine, Gangnam Severance Hospital, 211 Eonju-ro, Gangnam-gu, Seoul, Republic of Korea, 06273. Email: ukyjhome@yuhs.ac.

Submitted for publication February 19, 2020; accepted in revised form October 6, 2020.

SIGNIFICANCE & INNOVATIONS

- The bidirectional associations of type 2 diabetes mellitus with periodontitis and osteoarthritis have been recognized in previous studies.
- Severe periodontitis was positively associated with osteoarthritis in individuals with type 2 diabetes mellitus.
- The oral inflammation manifesting in periodontitis may be at least partly involved in the pathogenesis of osteoarthritis, particularly in patients with type 2 diabetes mellitus.

respectively (8). Osteoarthritis is one of the leading causes of physical disability and impairment in older adults and is a major public health problem worldwide (9). The incidence of osteoarthritis increases with age and is higher in individuals with type 2 diabetes mellitus (10). While degeneration of cartilage is the main pathologic finding in osteoarthritis, emerging evidence suggests that systemic low-grade inflammation is closely related to the initiation and progression of osteoarthritis (11). For example, as a source of metabolic inflammation, obesity may increase cartilage loss and joint pain by releasing adipokines and proinflammatory cytokines from adipose tissue (12).

Since the bidirectional associations of type 2 diabetes mellitus with periodontitis and osteoarthritis have been recognized in previous studies (13,14), we can infer that there is a relationship between periodontitis and osteoarthritis in diabetic patients. Thus, we conducted a cross-sectional study to examine how the severity of periodontitis is related to osteoarthritis in accordance with the presence of type 2 diabetes mellitus.

MATERIALS AND METHODS

Study population. This study was based on data obtained from the 2008 to 2015 Korean National Health and Nutrition Examination Survey (KNHANES), a cross-sectional and nationally representative survey conducted by the Korean Ministry of Health and Welfare (15). Household samples were randomly selected using a stratified, multistage design based on age, sex, and geographic area. Sampling weights were assigned to each individual to obtain results representing the entire Korean population. Participants responded to a 4-part questionnaire that included a health interview, health behavior survey, health examination, and nutrition survey. Participants were informed that they were randomly selected as a household to voluntarily participate in a nationally representative survey conducted by the Korean Ministry of Health and Welfare and that they had the right to refuse to participate according to the National Health Enhancement Act based on the National Statistics Law of Korea for the 2008 to 2015 KNHANES. All study participants provided their written informed consent. The KNHANES was managed by the institutional review board of the Korea Centers for Disease Control and Prevention (2010-02CON-21-C, 2011-02CON-06-C, 2012-01EXP-01-2C, 2013–07CON-03-4C, 2013-12EXP-03-5C, and 2015-01-02-6C) and conducted according to the ethical principles of the Declaration of Helsinki. We extracted 5,598 participants age ≥50 years whose data on periodontal examinations and radiographic images were available from the 2008 to 2015 KNHANES data sets. Among them, we excluded participants who did not complete the health examination survey and participants with other missing data (n = 2,071). Finally, the sample size of our study included 3,527 respondents of both sexes.

Data collection. According to standardized protocols, anthropometric measurements were taken by trained medical staff and all equipment was calibrated regularly, using the Asia-Pacific regional guidelines of the World Health Organization (WHO) and the International Obesity Task Force. We categorized body mass index as follows: underweight (<18.5 kg/m^2), normal weight (18.5–22.9 kg/m^2), overweight (23.0–24.9 kg/m^2), and obese (≥25 kg/m^2) (16). In the health interviews, we collected data on participant age, sex, health-related behaviors (e.g., cigarette smoking, alcohol intake, and exercise), residential area, and occupation using a self-reported questionnaire. Smoking status was classified as never smoker, exsmoker, or current smoker. Alcohol consumption was categorized according to frequency, regardless of the type of alcohol consumed. Current drinkers were defined as those who consumed a mean of ≥7 units of alcohol for men or ≥5 units for women at least 1 day per month over the past year. Nondrinkers included those who never drink or who consumed 1 or fewer glasses of alcohol per month over the past year. Regular aerobic exercise was regarded as ≥20 minutes of vigorous-intensity physical activity ≥3 days per week or ≥30 minutes of light or moderate-intensity physical activity ≥5 days per week.

Household income was classified into the following 4 groups: the lowest quartile, 2nd to 3rd quartile, 3rd to 4th quartile, and the highest quartile. Education level was assessed according to the number of years of schooling and classified into 4 categories as follows: <7 years (elementary school), 7–9 years (middle school), 10–12 years (high school), and ≥13 years (college or higher). Participants with type 2 diabetes mellitus were identified according to the use of insulin or other diabetes medications or were diagnosed with a fasting blood glucose of ≥126 mg/dl. Hypertension was defined as a systolic blood pressure of ≥140 mm Hg, a diastolic blood pressure of ≥90 mm Hg, or the current use of antihypertensive medication. Dyslipidemia was defined as a triglyceride level of ≥150 mg/dl, a high-density lipoprotein cholesterol concentration of ≤40 mg/dl for men and of ≤50 mg/dl for women, or current use of antidyslipidemic medications. Oral health-related behaviors included the frequency of tooth-brushing, the use of floss, and the use of an

Table 1. Community Periodontal Index (CPI) score for periodontitis and Kellgren/Lawrence (K/L) grade for osteoarthritis based on radiologic findings

CPI score	
0	Healthy periodontal tissue
1	Bleeding periodontal tissue when observed with the eye and a dental mirror
2	Periodontal tissue with plaque
3	Periodontal tissue forming a shallow periodontal pocket with depth of 4–5 mm
4	Periodontal tissue forming deep periodontal pockets with depth of ≥6 mm
K/L grade	
0	No radiologic findings of osteoarthritis
1	Doubtful: questionable joint space narrowing or possible osteophytic lipping
2	Minimal: definitive small osteophytes, minimal/mild joint space narrowing
3	Moderate: definitive moderate multiple osteophytes, joint space narrowing of at least 50%
4	Severe: large osteophytes, severely impaired joint space, subchondral bone cysts, and sclerosis

interdental brush. Frequent tooth-brushing was defined as brushing occurring ≥2 times daily.

Definition of periodontitis. A periodontal examination was performed for all study participants using the Community Periodontal Index (CPI) scoring system of the WHO (17). Approximately 20 grams of CPI probing force was applied to the periodontal tissue pocket, according to the WHO guidelines (18). To evaluate the condition of the periodontium, we used the CPI for Treatment Needs, which is widely applied in large population groups. Teeth were divided into sextants by number as follows: 1–5 (maxillary right posterior), 6–11 (maxillary anterior), 12–16 (maxillary left posterior), 17–21 (mandibular left posterior), 22–27 (mandibular anterior), and 28–32 (mandibular right posterior).

Table 2. Demographic and clinical characteristics of study participants in relationship to the presence of osteoarthritis and type 2 diabetes mellitus*

	With type 2 diabetes mellitus			Without type 2 diabetes mellitus		
Characteristic	OA neg.	OA pos.	*P*	OA neg.	OA pos.	*P*
Unweighted, no. (%)	1,451 (83.9)	324 (16.1)	–	1,546 (89.2)	206 (10.8)	–
Age, years	63.5 ± 0.3	69.9 ± 0.5	<0.001	59.6 ± 0.2	64.9 ± 0.7	<0.001
Female, % (SE)	42.3 ± 1.4	84.5 ± 2.3	<0.001	51.8 ± 1.1	70.7 ± 4.0	<0.001
Body mass index, kg/m^2, % (SE)			<0.001			0.047
Underweight (<18.5)	1.4 ± 0.4	1.8 ± 0.8	–	2.1 ± 0.4	1.5 ± 0.9	–
Normal (18.5–22.9)	30.1 ± 1.5	18.6 ± 2.7	–	36.6 ± 1.4	29.7 ± 3.9	–
Overweight (23.0–24.9)	24.2 ± 1.3	17.6 ± 2.4	–	28.4 ± 1.4	24.9 ± 3.1	–
Obese (≥25.0)	44.3 ± 1.6	61.3 ± 3.3	–	32.9 ± 1.3	43.9 ± 4.1	–
Current smoker, % (SE)	22.3 ± 1.4	10.3 ± 2.3	<0.001	17.4 ± 1.1	12.8 ± 3.8	0.031
Current drinker, % (SE)	29.6 ± 1.8	16.1 ± 3.1	<0.001	26.0 ± 1.5	17.9 ± 3.7	0.065
Regular exercise	36.1 ± 1.4	32.6 ± 3.3	0.330	33.3 ± 1.4	31.4 ± 4.0	0.668
Education			<0.001			<0.001
Elementary school	43.9 ± 1.6	82.9 ± 2.8	–	33.3 ± 1.8	59.4 ± 3.8	–
Middle school	19.2 ± 1.3	9.3 ± 2.2	–	17.9 ± 1.2	19.0 ± 3.6	–
High school	25.5 ± 1.4	6.6 ± 1.8	–	32.4 ± 1.7	18.9 ± 3.4	–
≥College	11.4 ± 1.0	1.3 ± 0.6	–	16.4 ± 1.4	2.7 ± 1.1	–
Household income			<0.001			0.001
Q1 (lowest)	31.9 ± 1.6	56.0 ± 3.7	–	21.8 ± 1.5	36.7 ± 3.8	–
Q2	28.1 ± 1.6	20.5 ± 2.7	–	24.0 ± 1.5	24.7 ± 3.3	–
Q3	22.2 ± 1.3	13.8 ± 2.9	–	25.0 ± 1.4	19.0 ± 3.1	–
Q4 (highest)	17.8 ± 1.3	9.7 ± 1.9	–	29.2 ± 1.9	19.6 ± 3.8	–
Hypertension, % (SE)	74.5 ± 1.6	79.8 ± 3.4	0.167	26.3 ± 1.4	47.1 ± 4.3	<0.001
Dyslipidemia, % (SE)	53.1 ± 2.2	60.6 ± 4.4	0.119	16.3 ± 1.1	25.3 ± 3.6	0.007
Cardiovascular diseases, % (SE)	17.8 ± 1.7	16.6 ± 3.3	0.746	3.0 ± 0.5	3.0 ± 1.0	0.974
Periodontitis assessed by CPI, % (SE)			0.035			0.466
0 (normal)	50.9 ± 1.8	44.4 ± 3.6	–	57.7 ± 2.0	61.2 ± 4.8	–
1–3 (nonsevere periodontitis)	33.7 ± 1.6	32.3 ± 3.5	–	28.4 ± 1.8	29.3 ± 3.9	–
4 (severe periodontitis)	15.4 ± 1.2	23.3 ± 3.4	–	13.9 ± 1.4	9.5 ± 3.2	–
Oral health-related behaviors, % (SE)						
Frequent tooth-brushing	38.1 ± 1.6	32.5 ± 3.4	0.127	44.6 ± 1.7	39.5 ± 3.6	0.204
Use of floss	9.7 ± 1.1	5.6 ± 1.7	0.081	12.6 ± 0.9	6.3 ± 1.9	0.018
Use of interdental brush	20.3 ± 1.5	9.1 ± 1.9	<0.001	24.8 ± 1.4	15.2 ± 3.0	0.012

* Values are the mean ± SE of the mean, unless indicated otherwise. CPI = Community Periodontal Index; neg. = negative; OA = osteoarthritis; pos. = positive; Q = quartile.

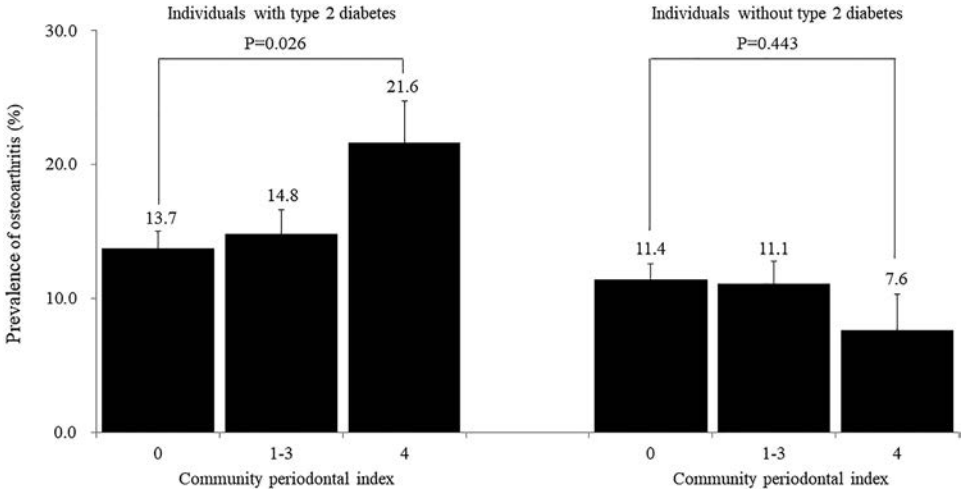


Figure 1. Proportion of osteoarthritis according to the severity of periodontitis using the Community Periodontal Index.

The depth of periodontal pockets was measured from selected index teeth (2, 3, 8, 14, 15, 18, 19, 24, 30, and 31, in that order). Trained dentists performed the examinations and assigned CPI scores from 0 to 4 (Table 1) (19). In our study, nonsevere periodontitis was indicated with a CPI score of 1 to 3 and severe periodontitis with a CPI score of 4 in at least 1 site. The mean kappa index for interrater reliability of periodontitis was $\kappa = 0.871$ (20).

Definition of osteoarthritis. Radiographic examinations of the knee and hip were performed using an SD 3000 Synchro Stand instrument (Accele Ray, Shinyoung). Weight-bearing anteroposterior, bilateral anteroposterior, and lateral plain radiographs were taken to assess the knees, while anteroposterior and bilateral plain radiographs were obtained to evaluate the hip. All radiographic images were graded from 0 to 4 points by 2 specialized musculoskeletal radiologists using the Kellgren/Lawrence (K/L) grading system (Table 1) (21). If there was a discrepancy of 1 grade between the 2 radiologists, the higher K/L grade was accepted. If the difference was >1 grade, the case was referred to a third radiologist and the grade of the first 2 most consistent with the third grade suggested was accepted. Knee/hip joint pain was self-reported using the question, "Have you experienced knee/hip pain for 30 or more days over the past 3 months?" In our study, osteoarthritis was defined as the presence of K/L grade on radiographic images, in either the knee or hip area, with pain for >30 days over the past 3 months.

Statistical analysis. In this survey, the sampling units were households selected through a stratified, multistage, probability-sampling design based on geographic area, sex, and age group. The sample weights were constructed for sample participants to represent the Korean population by accounting for the complex survey design, survey nonresponse, and poststratification. Therefore, in statistical analysis, we applied sampling weights to account for complex sampling. Because of the survey's characteristics, we presented the results as mean ± SE of the mean or proportion ± SE of the proportion. Differences in clinical characteristics were compared using a weighted *t*-test for continuous variables and the weighted Rao-Scott chi-square test for categorical variables. The odds ratios (ORs) and 95% confidence intervals (95% CIs) for osteoarthritis according to the severity of periodontitis, stratified by type 2 diabetes mellitus, were calculated using multiple logistic regression analyses. For this analysis, variables with a *P* value less than 0.2 in the univariate analysis and clinically important variables were entered. All analyses were conducted using SAS statistical software, version 9.4. All statistical tests were 2-sided and statistical significance was set at *P* value less than 0.05.

RESULTS

Table 1 illustrates the CPI score for periodontitis and the K/L grade for osteoarthritis based on the radiologic findings. Table 2 shows the demographic and clinical characteristics of study participants in relation to the presence of osteoarthritis and type 2 diabetes mellitus. The mean age and the proportions of female sex and obesity were higher, whereas the proportion of current smokers, the socioeconomic status manifested by household income, and the educational level were lower in the osteoarthritis group, regardless of type 2 diabetes mellitus. The frequency of both hypertension and dyslipidemia was higher in the osteoarthritis group without type 2 diabetes mellitus. Regarding oral health-related behaviors, individuals without osteoarthritis tended more often to use an interdental brush and dental floss, regardless of type 2 diabetes mellitus. Concerning the severity of periodontitis assessed using the CPI, the proportion with severe periodontitis

Table 3. Odds ratios for osteoarthritis according to the severity of periodontitis*

Periodontitis assessed by CPI	Type 2 diabetes mellitus	No type 2 diabetes mellitus
0 (normal)	1.00 (ref.)	1.00 (ref.)
1–3 (nonsevere periodontitis)	1.23 (0.67–2.32)	0.99 (0.61–1.61)
4 (severe periodontitis)	3.01 (1.51–5.84)	0.77 (0.37–1.62)

* Multiple logistic regression analysis included age, sex, body mass index (underweight, normal weight, overweight, and obese), smoking status (nonsmoker, exsmoker, and current smoker), alcohol consumption, regular exercise, education level, household income, hypertension, and frequent tooth-brushing. CPI = Community Periodontal Index; ref. = reference.

was significantly higher in the osteoarthritis group with type 2 diabetes mellitus.

Figure 1 shows the proportion of osteoarthritis according to the severity of periodontitis using the CPI. We determined that the proportion of osteoarthritis increased significantly according to the severity of periodontitis in individuals with type 2 diabetes mellitus.

Table 3 shows the results of multiple logistic regression analyses, which were conducted to assess the odds of predicting the presence of osteoarthritis according to the severity of periodontitis. Compared with individuals with type 2 diabetes mellitus without periodontitis, the OR for osteoarthritis in individuals with severe periodontitis and type 2 diabetes mellitus was 3.01 (95% CI 1.51–5.84) after adjusting for age, sex, body mass index, smoking status, alcohol intake, physical exercise, employment, household income, hypertension, and frequent tooth-brushing

DISCUSSION

In this cross-sectional study performed using a nationally representative sample of Korean adults, periodontitis was significantly associated with osteoarthritis in patients with type 2 diabetes mellitus, after adjusting for potential confounding variables. Our results are consistent with those of previous studies showing that individuals with periodontitis are more likely to have a higher prevalence of knee osteoarthritis than that in healthy controls (20). Kim et al reported that the adjusted OR for knee osteoarthritis was 1.47 (95% CI 1.17–1.85) in participants with severe periodontitis (defined as a CPI score of 4) among 7,969 Korean adults age ≥50 years. However, in their study, type 2 diabetes mellitus was not fully considered by presenting separate data on type 2 diabetes mellitus in multivariate analysis. As shown in the current study, there was a significant difference apparent in the prevalence of osteoarthritis according to the presence of type 2 diabetes mellitus. Moreover, our study revealed that positive associations between periodontitis and osteoarthritis can be especially applicable in patients with diabetes mellitus.

The exact mechanism by which periodontitis is associated with osteoarthritis is not well-known, but several possible explanations for the association deserve consideration. Inflammation of the oral cavity due to periodontitis could influence the development and progression of osteoarthritis via several virulence factors. Recent evidence has suggested that, in addition to the mechanical burden on the joint, reactive oxidative stress and low-grade chondrocyte inflammation are involved in the progression of osteoarthritis (22,23). Periodontitis originates from the formation of a dysbiotic biofilm that triggers host inflammatory responses through the expression of inflammatory cytokines (24). The oral and joint cavities are distant organs but could be linked via systemic inflammatory responses. Periodontal pathogens such as *Porphyromonas gingivalis* might spread to joints through the bloodstream, and the same bacterial DNA has been found in the periodontal tissue and synovial fluid of patients with osteoarthritis (25).

Hematogenous inoculation by periodontal bacteria may contribute to joint inflammation and damage. Both periodontitis and osteoarthritis are mediated by proinflammatory cytokines such as interleukin (IL)-6 and tumor necrosis factor (TNF) (6,26). Emerging evidence supports the idea that low-grade inflammation is closely related to the initiation and progression of periodontitis and osteoarthritis. Kwon et al reported that leukocyte count, a nonspecific marker of systemic inflammation, was elevated in the presence of periodontitis in 9,391 Korean adults (27). Levels of proinflammatory cytokines such as IL-6 and TNF are elevated in the synovial fluid and cartilage of patients with osteoarthritis (6), while chronic cytokine elevation is associated with an increased risk of osteoarthritis, according to a prospective study (11). These proinflammatory cytokines upregulate inflammatory responses and inhibit the synthesis of proteoglycan and type II collagen in chondrocytes (28). Thus, chronic systemic inflammation may be a process that can elucidate the systemic effects of periodontitis on the pathogenesis of osteoarthritis through a chain cascade.

Another noteworthy finding in our study was that there were differences in the association between periodontitis and osteoarthritis according to the presence of type 2 diabetes mellitus. Type 2 diabetes mellitus is an important risk factor associated with both the development of periodontitis and osteoarthritis (5,29). Whereas many factors are known to contribute to the development of diabetes mellitus, emerging evidence has suggested that diabetes mellitus is increasingly viewed as an inflammatory disease (30). In this regard, the inflammatory cascades from the oral cavity to joint tissue might be augmented predominantly through systemic inflammation in patients with type 2 diabetes mellitus. In line with the view of insulin resistance and low-grade

inflammation, a recent systematic review involving 7 eligible studies on osteoarthritis failed to demonstrate significant associations with metabolic syndrome, but most of the included studies adopted cross-sectional designs, with only partial adjustments made for potential confounding variables (31).

Some limitations should be acknowledged regarding the interpretation of the current study findings. First, this was a cross-sectional study, so we cannot infer a causal relationship between periodontitis and osteoarthritis. Further prospective research is warranted to elucidate any positive correlations between periodontitis and osteoarthritis. Second, as we included only members of an East Asian population, the results may not be generalizable to other ethnic groups. In addition, due to the nature of the secondary KNHANES data set, some possible residual confounding factors regarding the characteristics of osteoarthritis pain, synovial inflammatory markers such as IL-6, and diet quality were not included in the multiple logistic regression analysis model.

Despite these potential limitations, our results can be generalized to the entire Korean population by applying complex sampling design analysis. In addition, we used an objective tool to detect osteoarthritis in the knee and hip joints on the basis of both radiologic findings and osteoarthritis-related symptoms.

In conclusion, we found that severe periodontitis was positively and significantly associated with osteoarthritis in individuals middle-aged and older with type 2 diabetes mellitus. Our findings suggest that oral inflammation manifesting in periodontitis may be at least partially involved in the pathogenesis of osteoarthritis, particularly in patients with type 2 diabetes mellitus.

ACKNOWLEDGMENTS

We thank all those who conducted the KNHANES, as well as the individuals who participated in this survey.

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. Lee 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. H. S. Kim, Y. J. Lee, Park.

Acquisition of data. Son, Park, H. Kim.

Analysis and interpretation of data. H. S. Kim, H. S. Lee, Son, Y. J. Lee, Park, H. Kim.

REFERENCES

1. Griffin SO, Jones JA, Brunson D, Griffin PM, Bailey WD. Burden of oral disease among older adults and implications for public health priorities. Am J Public Health 2012;102:411–8.
2. Scannapieco FA, Bush RB, Paju S. Associations between periodontal disease and risk for atherosclerosis, cardiovascular disease, and stroke: a systematic review. Ann Periodontol 2003;8:38–53.
3. Kelesidis N. A racial comparison of sociocultural factors and oral health perceptions. J Dent Hyg 2014;88:173–82.
4. Eke PI, Dye BA, Wei L, Slade GD, Thornton-Evans GO, Borgnakke WS, et al. Update on prevalence of periodontitis in adults in the United States: NHANES 2009 to 2012. J Periodontol 2015;86:611–22.
5. Grossi SG, Genco RJ. Periodontal disease and diabetes mellitus: a two-way relationship. Ann Periodontol 1998;3:51–61.
6. Kapoor M, Martel-Pelletier J, Lajeunesse D, Pelletier JP, Fahmi H. Role of proinflammatory cytokines in the pathophysiology of osteoarthritis. Nat Rev Rheumatol 2011;7:33–42.
7. Nelson AE. Osteoarthritis year in review 2017: clinical. Osteoarthritis Cartilage 2018;26:319–25.
8. Lee S, Kim SJ. Prevalence of knee osteoarthritis, risk factors, and quality of life: the fifth Korean National Health And Nutrition Examination Survey. Int J Rheum Dis 2017;20:809–17.
9. Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012;380:2163–96.
10. Piva SR, Susko AM, Khoja SS, Josbeno DA, Fitzgerald GK, Toledo FG. Links between osteoarthritis and diabetes: implications for management from a physical activity perspective. Clin Geriatr Med 2015;31:67–87.
11. Robinson WH, Lepus CM, Wang Q, Raghu H, Mao R, Lindstrom TM, et al. Low-grade inflammation as a key mediator of the pathogenesis of osteoarthritis. Nat Rev Rheumatol 2016;12:580–92.
12. Azamar-Llamas D, Hernández-Molina G, Ramos-Ávalos B, Furuzawa-Carballeda J. Adipokine contribution to the pathogenesis of osteoarthritis. Mediators Inflamm 2017;2017:5468023.
13. Winning L, Linden GJ. Periodontitis and systemic disease: association or causality? Curr Oral Health Rep 2017;4:1–7.
14. Velasquez MT, Katz JD. Osteoarthritis: another component of metabolic syndrome? Metab Syndr Relat Disord 2010;8:295–305.
15. Kweon S, Kim Y, Jang MJ, Kim Y, Kim K, Choi S, et al. Data resource profile: the Korea National Health and Nutrition Examination Survey (KNHANES). Int J Epidemiol 2014;43:69–77.
16. World Health Organization. The Asian-Pacific perspective: redefining obesity and its treatment. 2000. URL: https://apps.who.int/iris/handle/10665/206936.
17. Ainamo J, Barmes D, Beagrie G, Cutress T, Martin J, Sardo-Infirri J. Development of the World Health Organization (WHO) community periodontal index of treatment needs (CPITN). Int Dent J 1982;32:281–91.
18. World Health Organization. Oral health surveys: basic methods: 5th ed. World Health Organization; 2013. p. 48–56
19. World Health Organization. Periodontal country profiles: World Health Organization. 2005. URL: https://www5.dent.niigata-u.ac.jp/~prevent/perio/contents.html.
20. Kim JW, Chung MK, Lee J, Kwok SK, Kim WU, Park SH, et al. Association of periodontitis with radiographic knee osteoarthritis. J Periodontol 2020;91:369–76.
21. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis 1957;16:494.
22. Henrotin Y, Kurz B, Aigner T. Oxygen and reactive oxygen species in cartilage degradation: friends or foes? Osteoarthritis Cartilage 2005;13:643–54.
23. Scanzello CR. Role of low-grade inflammation in osteoarthritis. Curr Opinion Rheumatol 2017;29:79–85.
24. Jiao Y, Hasegawa M, Inohara N. The role of oral pathobionts in dysbiosis during periodontitis development. J Dent Res 2014;93:539–46.
25. Temoin S, Chakaki A, Askari A, El-Halaby A, Fitzgerald S, Marcus RE, et al. Identification of oral bacterial DNA in synovial fluid of patients with arthritis with native and failed prosthetic joints. J Clin Rheumatol 2012;18:117–21.

26. Di Benedetto A, Gigante I, Colucci S, Grano M. Periodontal disease: linking the primary inflammation to bone loss. Clin Dev Immunol 2013;2013:503754.
27. Kwon YJ, Jeon KJ, Chung TH, Lee YJ. Elevated leukocyte count is associated with periodontitis in Korean adults: the 2012-2014 KNHANES. Oral Dis 2017;23:241–6.
28. Maldonado M, Nam J. The role of changes in extracellular matrix of cartilage in the presence of inflammation on the pathology of osteoarthritis. Biomed Res Int 2013;2013:284873.
29. Dubey NK, Ningrum DN, Dubey R, Deng YH, Li YC, Wang PD, et al. Correlation between diabetes mellitus and knee osteoarthritis: a dry-to-wet lab approach. Int J Mol Sci 2018;19:3021.
30. Donath MY, Shoelson SE. Type 2 diabetes as an inflammatory disease. Nat Rev Immunol 2011;11:98–107.
31. Li S, Felson DT. What is the evidence to support the association between metabolic syndrome and osteoarthritis? A systematic review. Arthritis Care Res (Hoboken) 2019;71:875–84.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 3, March 2022, pp 410–419
DOI 10.1002/acr.24477

Patient Perspectives Surrounding Intraarticular Injections for Knee Osteoarthritis: A Qualitative Study

Nora K. Lenhard,[1] Emma E. Williams,[1] Emma C. Lape,[1] Lindsey A. MacFarlane,[2] Elena Losina,[3] and Jeffrey N. Katz[2]

Objective. Intraarticular (IA) injections are used frequently for knee osteoarthritis (OA), but little is known about patients' attitudes toward these therapies. We aimed to better understand patients' perceptions of the facilitators of and barriers to IA injections for knee OA.

Methods. We conducted a qualitative, descriptive, exploratory study and held focus groups and individual interviews with participants with knee OA, including some who had and some who had not received IA injections. We conducted a thematic analysis to identify themes describing the factors that participants found influential when deciding whether to try an IA injection.

Results. We held 3 focus groups with 12 participants and conducted 3 individual interviews (15 participants total). We identified the following 4 themes that shaped participants' decisions to receive a specific injection: 1) the impact of OA on participants' lives; 2) participants' attitudes and concerns, including desire to avoid surgery, willingness to accept uncertain outcomes, and concerns about side effects and dependence; 3) the way participants gathered and processed information from physicians, peers, and the internet; and 4) the availability of injectable products. Participants weighed the desire to regain function and delay surgery with concerns about side effects, uncertain efficacy, and costs.

Conclusion. Participants were concerned about the effectiveness, toxicity, availability, and cost of injectable products. They balanced disparate sources of information, uncertain outcomes, limited product availability, and other injection-related concerns with a desire to decrease pain. These findings can provide clinicians, investigators, and public health professionals with insights into challenges that patients face when making injection decisions.

INTRODUCTION

Knee osteoarthritis (OA) is a painful, disabling condition affecting >14 million Americans (1). Presently, there are no disease-modifying therapies approved for clinical use that slow the progression of knee OA, a chronic condition; consequently, treatment is aimed at managing symptoms. Patients typically progress through a variety of treatments, including nonsteroidal antiinflammatory drugs (NSAIDs), physical therapy, and intraarticular (IA) injections, before considering total knee replacement (TKR) (2). Many products can be delivered intraarticularly, including glucocorticoids, hyaluronic acid (HA) derivatives, and platelet-rich plasma (PRP). While professional societies recommend the use of glucocorticoid injections based on their efficacy (2), recommendations for HA and PRP injections conflict (2,3), and studies show variable efficacy (4,5). Glucocorticoid injections reduce inflammation; HA and PRP injections work by different mechanisms that are not fully understood, although HA is a component of healthy cartilage, and PRP contains potent growth factors (6). While there are scant data on utilization of IA injections, one study among patients who had undergone TKR found that 29% had received a glucocorticoid injection and 6% an HA injection in the year prior to surgery (7).

While the number of injectable treatment options for knee OA has increased, little is known about factors that patients consider when making decisions about IA injections. Qualitative research is

Supported by Flexion Therapeutics. Dr. Katz's work was supported by Samumed and Flexion Therapeutics.

[1]Nora K. Lenhard, BA, Emma E. Williams, BA, Emma C. Lape, BA: Brigham and Women's Hospital, Boston, Massachusetts; [2]Lindsey A. MacFarlane, MD, MPH, Jeffrey N. Katz, MD, MSc: Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts; [3]Elena Losina, PhD: Brigham and Women's Hospital, Harvard Medical School and Boston University School of Public Health, Boston, Massachusetts

Dr. MacFarlane has received consulting fees from Flexion Therapeutics (less than $10,000). Dr. Losina has received consulting fees from Pfizer (less than $10,000) and grant support from Samumed, Flexion Therapeutics, and Pfizer. No other disclosures relevant to this article were reported.

Address correspondence to Jeffrey N. Katz, MD, MSc, Orthopaedic and Arthritis Center for Outcomes Research, Brigham and Women's Hospital, 75 Francis Street, Building for Transformative Medicine, Suite 5016, Boston, MA 02115. Email: jnkatz@bwh.harvard.edu.

Submitted for publication April 25, 2020; accepted in revised form September 29, 2020.

SIGNIFICANCE & INNOVATIONS

- Many intraarticular (IA) injectable products are available for knee osteoarthritis (OA), but little is known about the factors patients consider when deciding whether to receive an injection or which particular injectable product to use.
- Participants with knee OA noted that a variety of factors (including the impact of OA on their lives, their attitudes and concerns, the way they gathered and processed information about injection options, and the availability of certain products) shaped the decisions they made surrounding injections.
- Participants raised various concerns about IA injections, including their efficacy, the specific content of the injection and potential damage or side effects that these substances might cause, and out-of-pocket costs. Our findings suggest that it may be useful for clinicians to help patients navigate these concerns during shared decision-making discussions.

well suited to examine patient preferences and experiences, as it allows patients to share their points of view and experiences. Prior qualitative studies focused on OA have examined the experience of living with OA (8,9), patients' considerations of OA treatments (10), and decision-making surrounding total joint replacement (9). We are not aware of prior qualitative studies focused on injection therapies. We conducted focus groups and individual interviews with knee OA patients with the aim of better understanding the factors that patients consider when choosing whether to undergo a specific injection.

PATIENTS AND METHODS

Design. We conducted a qualitative, descriptive, exploratory study to better understand participants' experiences with injection-related choices. We included participants who had and had not received prior injections in order to better understand facilitators and barriers from different perspectives. We conducted a thematic analysis, which involves identifying and interpreting patterns of meaning ("themes") in qualitative data (11). We utilized an inductive approach, in which we developed themes from the data (rather than a deductive approach, in which data are interpreted according to prespecified theory) (12). We view concordance of findings across focus groups led by different interviewers, and across focus groups and individual interviews, as evidence of the validity of the findings.

Participants. We searched the Brigham and Women's Hospital (BWH) electronic medical record for patients with a knee OA diagnosis who were seen by selected rheumatologists and sports medicine physicians for knee OA between January 2018 and January 2019. To meet inclusion criteria, subjects had to live within 60 miles of BWH (as individuals who live farther away are less likely to participate), have a knee radiograph from the past 5 years, have evidence of cartilage damage or osteophytes on imaging, and have reported knee pain within the past 4 weeks. We excluded persons who had inflammatory arthritis, were claiming workers' compensation, or did not speak English. We sent letters inviting potentially eligible subjects to participate and then called subjects to gauge interest and assess eligibility. We obtained verbal consent of those interested and eligible to participate.

To characterize the demographic and clinical status of participants, we obtained data on age, sex, radiographic score (Kellgren/Lawrence grade), symptom duration, history of injection use, and pain severity. To assess pain severity, we asked 3 questions from the Knee Injury and Osteoarthritis Outcome Score Pain Scale (pain with walking, going up or down stairs, and standing upright, each rated from none [0] to extreme [4]). We summed these ratings, yielding a scale with possible range from 0 to 12. The Partners HealthCare Human Subjects Committee approved the study protocol.

Data collection. We conducted focus groups and individual interviews with subjects who were unable to attend the focus groups. The focus groups were conducted by pairs of moderators (ECL and EL, NKL and ECL, and ECL and EL for the 3 groups, respectively), including a biostatistician (EL) and 2 research assistants (NKL and ECL). None of the moderators had previous contact with study participants. The individual interviews (conducted by NKL and ECL) added to the data's richness by including perspectives of participants who otherwise would not have been able to attend. The focus groups and interviews followed a moderators' guide that probed participants' perceptions of the facilitators of and barriers to injections (Table 1). The moderators' guide was developed collaboratively by a group that included physicians, a biostatistician, and research assistants, all of whom had clinical and/or research experience with knee OA. Participants were compensated with a $25 Amazon gift card and dinner. All focus groups and interviews were audiorecorded and transcribed verbatim by Landmark Associates (thelai.com).

Thematic analysis. We analyzed the transcript data according to the thematic analysis procedures demonstrated in a study by Braun and Clarke (12). Investigators reviewed and discussed transcript data throughout the data collection process. We concluded recruitment once we felt that the transcripts were sufficiently rich to address the study's goals (13). This point, at which no new themes are generated, is often referred to as thematic saturation.

Three investigators (NKL, EEW, and LAM) read the transcripts to familiarize themselves with the data. These investigators then developed a coding scheme by labeling 1 of the focus group transcripts with codes (words or short phrases) that described the most basic segments of data relevant to the guiding questions,

Table 1. Questions included in the moderators' guide

Topic	Questions
Strategies for managing pain	What do you do to relieve your pain?
Attitudes toward new treatments	Can you recall a time when your healthcare provider offered you a new treatment for your knee osteoarthritis? Think about your level of pain and ability to do your daily activities over the past month. Given that current state, would you consider trying a new treatment? Why or why not? What kind of things would you consider when thinking about a new treatment? What information would you like to have before trying it? What kind of benefits would you hope for? What risks would be acceptable or unacceptable to you?
Perceptions of intraarticular injections	Tell us what you know or have heard about injection treatments for knee osteoarthritis. What kind of benefits would you expect to see? What kind of risks or side effects could result? Would you expect an injection to provide more relief than a medication you take by mouth or topical medication? What comes to mind when you picture an injection? Does that mental image trigger any fear or anxiety for you? What concerns, if any, do you have about injection treatments?* Do you have any reservations about the experience of being injected (pain, using a needle, etc.)?* Are you concerned about the costs of the injection (copays or paying out of pocket)?* What kind of benefits would you expect to see after an injection?* How long would you expect the benefits to last? Did you have any concerns that the injection might cause any harms or side effects? If so, where did you hear about potential risks?* Have you heard about injection treatments from others (friends or family) who also have knee osteoarthritis?* If you have had an injection, please tell us about that experience. Did it go how you expected, or was it different?† What concerns, if any, did you have before your injection? What made you able to overcome these concerns?† What prompted you to seek additional treatment?† How long did you expect the benefits to last?† Did the experience make you more or less comfortable about the idea of having an injection in the future? Why?†
Considering the options	What questions would you ask your doctor if you were considering having an injection? Imagine that a friend is considering having an injection for her knee osteoarthritis. What would you advise her? Imagine that your doctor is telling you about how long the pain relief will last after an injection treatment. How would you react if the pain relief lasted 1 month? 3 months? 6 months? Would that change your decision whether or not to have the treatment? What characteristics would make something the ideal pain management strategy for you?

* For those who have not received injections.
† For injection recipients.

including "What factors facilitate patients' decisions to receive an injection?" and "What factors serve as barriers?" Once the 3 investigators came to a consensus about the coding scheme, one investigator (EW) coded the rest of the transcript data using Dedoose, a qualitative analysis software (Dedoose.com).

In the next phase of analysis, investigators reviewed the coding scheme and read different portions of the transcripts to generate themes. Themes express broader patterns in the data by grouping individual codes and describing the relationships between these codes, as well as their relationship to the guiding question (12). This analytic phase produced a thematic scheme, which consists of a list of themes, subthemes, explanatory statements describing each theme's connection to the guiding questions, and supporting quotations. We created a thematic map portraying the relationships between themes (Figure 1). All investigators reviewed the themes to ensure that they were distinct and accurate and approved the thematic scheme and map.

RESULTS

Participants. We sent invitation letters to 106 potentially eligible subjects. Fifteen individuals chose to participate, 43 declined, and the remainder were ineligible (3 individuals), could not be scheduled or did not attend their scheduled group (12 individuals), or were unable to be reached (33 individuals). We conducted 2 focus groups involving 12 subjects (with 5 participants in 2 of the groups and 2 participants in 1 group) and 3 individual interviews. Participant characteristics are described in Table 2. The focus groups lasted 70–80 minutes, and interviews lasted 21–30 minutes.

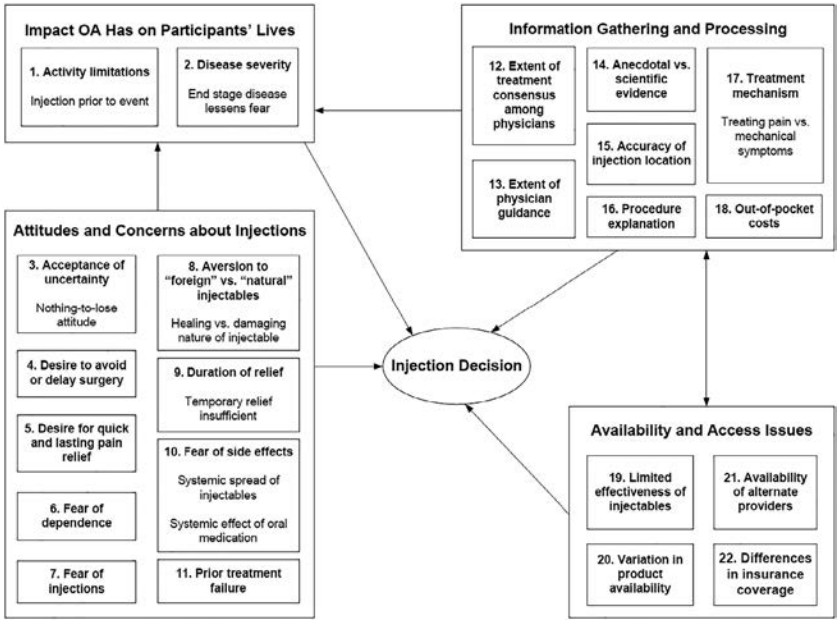


Figure 1. Thematic map visually depicting relationships between themes and subthemes identified through thematic analysis.

Thematic analysis. We identified 4 themes and 22 subthemes (Figure 1). In the following section, we present the themes, subthemes, and supporting data. Themes and selected quotes are summarized in Table 3; the complete thematic scheme is available (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24477/abstract). The following were the 4 themes: impact OA has on participants' lives; participants' attitudes and concerns about injections; ways in which participants gathered and processed information about injections; and availability of and access to different injectable products.

Table 2. Demographic and clinical characteristics of focus group and interview participants*

Characteristic	Value
Female sex	13 of 15
Age, median (range) years	61 (44–81)
Pain rating, median (12-point pain scale)	3 points
Distribution of pain ratings (12-point pain scale)	
1–3 points	8
4–6 points	3
≥7 points	4
Duration of knee pain	
1–3 years	3
>3 years	12
K/L grade	
1	1
2	3
3	5
4	6
No. of prior intraarticular injections	
0	4
1–2	3
≥3	8

* Values are the number of participants unless indicated otherwise. K/L = Kellgren/Lawrence scale for scoring radiographs.

Theme 1: impact OA has on participants' lives. The extent to which OA impacted participants' lives influenced their willingness to try different IA therapies.

Subtheme 1: activity limitations. Participants were frustrated by the limitations that OA placed on valued activities. The desire to regain function and return to these activities provided a rationale for trying an injection.

> "I was just so desperate to get back to functional. I was missing work… I wanted to run again. I was just feeling very defeated in my daily life, and I wanted to get back to functionality."

Some subjects felt that injections provided them with the opportunity to participate in special events.

> "I know [the Boston Marathon is] comin' up, so I'll… come in and get the shots so I have a tune-up for my knees and I can run."

Subtheme 2: disease severity. OA severity influenced the treatments subjects considered viable options. Injections were viewed as less appropriate for severe OA because surgical intervention would better address the underlying problem.

> "Until [the injection] doesn't work—because I'm sure that if the disease progresses, at some point that doesn't work anymore. Then I think that an injection is a temporary aid before you have surgery."

Table 3. Themes, subthemes, and selected supporting data identified through thematic analysis

Theme or subtheme	Supporting text from transcripts
Theme 1: impact OA has on participants' lives	
Activity limitations	"Since I've had my injections and for maybe 2 months after they really kicked in, I really am not as limited I'm back to running…I don't have those episodes where my knee suddenly just says I'm in charge. Not like I was having a year ago." (Subject 1, 50-60 years old, prior injections)
Injection prior to event†	"If I had a cruise coming up and my knee was acting up the week before I would be like, 'Please, inject this before I get on the ship,' so that I can walk, and talk, and dance my way around." (Subject 13, 70–80-year-old woman, no prior injections)
Disease severity	"Some people rave over it, and other people, they have advanced arthritis. That isn't gonna help them, 'cause they need more than that, but it does work for some people." (Subject 13, 70-80-year-old woman, no prior injections)
End-stage disease lessens fear†	"I'm like how worse can it be? I'm already bone-on-bone with bone spurs. There's nothing there. I mean the injections provide the cushion to keep me going for a few months, until I get another one. How much worse can it get?" (Subject 4, 50-60-year-old woman, prior injections)
Prior treatment failure	"I can do the Tylenol and the nonsteroidals when it's acting up…[but] today I have the worst knee pain I've ever had, so if somebody was giving me an injection today I'd be like, 'Where? Sign me up. Sign me up', because it was killing me…I've taken double dose [of] nonsteroidals and Tylenol so I could get myself in here." (Subject 13, 70-80-year-old woman, no prior injections)
Theme 2: attitudes and concerns about injections	
Acceptance of uncertainty	"[My doctor] said it's different for different people. Some people, they could last anywhere from 2 to 3 months. Other people say they get no relief to maybe a week or so, and that was it. She would not be able to tell me definitively how long it would work, or if it would even work." (Subject 4, 50–60-year-old woman, prior injections)
Nothing-to-lose attitude†	"If it was 50/50, I'd give it a shot…What do you got to lose?" (Subject 10, 60–70-year-old man, prior injections)
Desire to avoid or delay surgery	"I'm probably gonna go and have a series of gel injections…to me, it would be worth it rather than going straight into surgery." (Subject 6, 80-90-year-old woman, prior injections)
Desire for quick and lasting pain relief	Interviewer 1: "How would you define a success?" Subject: "Pain free." (Subject 10, 60–70-year-old man, prior injections)
Fear of dependence	"The expectation was that I probably would have to keep on gettin' [injections]. That's what it was. I don't wanna depend on that." (Subject 15, 40-50-year-old woman, no prior injections)
Fear of injections	"Thinking of the idea of an injection is—that's tough to get over. It really is." (Subject 6, 80-90-year-old woman, prior injections)
Aversion to "foreign" vs. "natural" injectables	"I also like the fact that it was my own juice going back in me." (Subject 2, 70-80-year-old man, prior injections)
Healing vs. damaging nature of injectable†	"Well, I don't want anything that's gonna cause more damage, and I think the steroids aren't gonna make my knee better." (Subject 13, 70-80-year-old woman, no prior injections)
Duration of relief	"Forget about a month [of pain relief], 6 months doesn't sound too good to me either, but it's possible depending on what I'm going through. I like the idea of a year or 2." (Subject 7, 60-70-year-old woman, no prior injections)
Temporary relief insufficient†	"I do it this way, it's gonna help for a matter of time. I want you to do somethin' [so that] it ain't gonna come back. I don't want it to come back." (Subject 15, 40-50-year-old woman, no prior injections)
Fear of side effects	"I mean you watch those commercials on TV and the guy's speaking so fast with all the potential side effects. I'm like I'd rather live with that disease than all the side effects." (Subject 1, 50-60 years old, prior injections)
Systemic spread of injectables†	"I think the main thing that would give me pause would be a high risk of allergic reaction. You get this now injected into your body…you [could] have a systemic allergic reaction." (Subject 3, 40-50-year-old woman, prior injections)
Systemic effect of oral medication†	"I manage it with Mobic a few times a week. I would be more comfortable if it took it every day, but my blood pressure would be higher, so I try to only take it when I really need it." (Subject 13, 70-80-year-old woman, no prior injections)
Theme 3: information gathering and processing	
Extent of treatment consensus among physicians	"One of the things that I would really like from my doctor is like a lot of this is—it's evidence-based, but it's not 100 percent guarantee, right? Boston has one approach. I have friends that live in Utah, and that's a different approach. People in New York have a different approach. It feels like it's not objective." (Subject 1, 50-60 years old, prior injections)
Extent of physician guidance	"Yeah, talking to other people, [they] have had different [experiences], so everybody is not the same. You just listen to [the doctor], and he's the expert and go with what he says, if it doesn't involve a needle." (Subject 14, 70-80-year-old woman, no prior injections)
Physician's recommendation†	"[My doctor] offered a steroid injections, but also, again, talked to me about the side effects of steroids, and she offered me both the Synvisc and the Euflexxa for the hyaluronic acid. I forget why she told me she thought the Euflexxa would be better…but I think she told me the Euflexxa would be what she would recommend of the two. I went with the Euflexxa." (Subject 3, 40-50-year-old woman, prior injections)

(Continued)

Table 3. (*Cont'd*)

Theme or subtheme	Supporting text from transcripts
Anecdotal vs. scientific evidence	"I would talk to the doctor, and if the doctor wasn't readily available, in the meantime I would probably use the internet. There's so much information out there." (Subject 13, 70-80-year-old woman, no prior injections)
Accuracy of injection location	"[I would like] ultrasound to guide [the] location of where to place it. Hopefully that's helpful so they do get the right spot." (Subject 10, 60-70-year-old man, prior injections)
Procedure explanation	"I was at my primary over at the [hospital], and told her, and she said gonna try the shot right then and there. I don't know what it was. It hurt. It didn't do a thing." (Subject 5, 70-80-year-old woman, prior injections)
Treatment mechanism	"It's a mechanical problem and how's it gonna be fixed? That's why I think I'm looking for a cushion of some kind, that would be absolutely ideal if you could just inject it and have it puff out." (Subject 6, 80-90-year-old woman, prior injections)
Treating pain vs. mechanical symptoms†	"I would want a cushioning, that's all I want, so it wouldn't grind and scrape. I think that's what causes the pain more than anything else." (Subject 6, 80-90-year-old woman, prior injections)
Out-of-pocket costs	"I certainly didn't want to continue to pay out-of-pocket for something that may or may not work. That's how I ended up on the cortisone for like 7 years." (Subject 8, 40-50-year-old woman, prior injections)
Theme 4: availability and access issues	
Limited effectiveness of injectables	"I've been given—usually, they introduced to me the gel injection that I told you. I was excited that it was gonna work. I did it twice and didn't do much for me, not even for a little bit. The cortisone injection didn't do much either. I didn't have any good experience about any of it so far." (Subject 11, 40-50-year-old woman, prior injections)
Variation in product availability	"I think my doctors are pretty conservative, because they didn't wanna do—they really didn't even offer me anything other than steroids and surgery." (Subject 1, 50-60 years old, prior injections)
Availability of alternate providers	"I did seek out a doctor that did PRP based on the recommendation of my physical therapist because he…he's a professional and well-versed. He said it doesn't really happen a lot around here, so you might find that it's new to this area. He helped me find some doctors that did it and went from there." (Subject 8, 40-50-year-old woman, prior injections)
Differences in insurance coverage	"In my case, my insurance wouldn't cover it even though I have good insurance. I was shocked." (Subject 10, 60-70-year-old man, prior injections)

* PRP = platelet-rich plasma.
† Subthemes.

On the other hand, having end-stage disease reduced fears about negative effects. Participants' sense that the joint could not deteriorate further made them more willing to try new injectables.

> "How much worse can it be? I'm already bone-on-bone with bone spurs. There's nothing there. I mean the injections provide the cushion to keep me going for a few months, until I get another one."

Theme 2: attitudes and concerns about injections. Participants expressed a variety of attitudes and concerns that appeared to influence their willingness to try an injection.

Subtheme 3: acceptance of uncertainty. Some participants understood that response to an injection is variable and were willing to gamble on trying a new one.

> "I think all they can really do is give you the statistics behind the science. You either decide to go for it or you don't."

Some participants described feeling that they had nothing to lose after other treatments failed.

> "At that point, I said, 'Why not give it a try, because I'm not ready for surgery yet, so why not'?"

Subtheme 4: desire to avoid or delay surgery. The possibility of delaying or avoiding TKR increased many participants' openness to IA injections.

> "My understanding of the surgery is that basically for 3 to 6 months, it's really not much fun…I'm 74. I don't know how many 3-6 monthers I have…. Do I wanna take one of those chunks and ensure that I'm miserable with no guarantee that I'll be really much better than I am now, afterwards?"

Subtheme 5: desire for quick and lasting pain relief. Participants varied in the degree of pain relief they expected from a successful injection; some defined success as total pain relief and some as rapid relief.

> "Well, basically, for quick relief, that would be what I would be expecting… and if it didn't show quick relief, there would be no more… [Quick] is within a four-hour period."

Subtheme 6: fear of dependence. Some participants chose injections to avoid dependence on NSAIDs or opioids, while others were concerned about becoming reliant on injections.

"I used to go through the big jug [of Advil], which isn't really healthy either, to be eating ibuprofen all the time."

"I mean, then you're a slave to [injections] too, and you can't use [them] that often."

Subtheme 7: fear of injections. The size of the needle and pain associated with the injection deterred a few participants.

"I never had an injection 'cause I'm scared. No, that needle was too big."

Subtheme 8: aversion to "foreign" versus "natural" injectables. Some participants expressed a preference for products they perceived as "natural," such as hyaluronic acid (HA) or platelet-rich plasma (PRP), over those they considered synthetic, such as glucocorticoids. Steroid-based products were viewed as more likely to damage the joint.

"[For] the Orthovisc, I thought, 'Well, maybe this one's better because it's more natural.'"

"I did not want straight steroid injections, because I know that steroids impede healing… When the hyaluronic acid was introduced, I did a little research on it and learned that it's something we make ourselves anyway."

Subtheme 9: duration of relief. The expected duration of pain relief was an important factor, especially in relation to cost. The longer the benefits lasted, the more favorably participants viewed it.

"Definitely the longer it lasts, the better…I was not prepared for a $90 co-pay for the med, and then 3 visits at $15 co-pays a piece… If I had to do that multiple times a year, I'd probably have it be a little different. The durability for me is important."

Some participants viewed the temporary relief offered by injections as an indication that they were masking deeper problems by temporarily relieving symptoms.

"I mean, if you're in that much pain, you need surgery, in my opinion. It's like taking a band-aid. It's only gonna give you short-term relief. It doesn't fix the problem."

Subtheme 10: fear of side effects. Many participants expressed concerns about side effects, including flare-up and joint damage.

"I would be a little concerned as to what wear and tear it would do to the joint or what wear and tear it would do traveling through the tissue."

The potential for systemic spread of the injectable or a systemic reaction concerned participants. At the same time, some expressed concerns about the systemic effects of NSAIDs or had contraindications to NSAIDs and saw injections as a more targeted solution.

"I'm wondering more, does it go into the bloodstream? How do they know they're getting that cavity in your knee and not hitting an artery?"

"…pills don't differentiate. A pill goes in your system and I don't need it in my shoulders. I just need it in my knee. An injection is—it goes right to the issue."

Subtheme 11: prior treatment failure. Participants expressed greater desire to try injections when their previous treatments failed.

"That's when I usually run in and get injections, because it just hurts so bad. The Tylenol doesn't soothe it. All my little remedies that I've tried along the way don't soothe it."

Theme 3: information gathering and processing. The information sources participants used and the ways in which they considered this information affected their expectations and opinions of injections.

Subtheme 12: extent of treatment consensus among physicians. The lack of consensus among physicians across different specialties and geographic regions of the US was unsettling for participants. They expressed a desire for unified recommendations.

"It seems like the surgeons recommend surgery, and the pain people recommend nerve blocker[s]… Rheumatology's like well, we can try this, or we can try that. There's not really an opinion about what might help."

Subtheme 13: extent of physician guidance. Some participants desired clear guidance from their physician regarding what treatment would be best for them.

"I've actually had a lotta respect for my doctors who have said, 'I wouldn't try this because I've not had good success with this…' I've been very appreciative of that, because what works for some people might not necessarily work for other people."

A physician's recommendation regarding specific products carried substantial weight in the decision-making process.

"[My doctor] thought that the Synvisc would be the best for me at the time… she made the recommendation as to which steroid I should probably try."

Subtheme 14: anecdotal versus scientific evidence. When deciding if an injection was the right treatment, participants consulted a variety of sources, including their doctors, friends or family, and the internet. They acknowledged that some of these sources were higher quality than others, but they wanted to understand the experiences of those who had gotten injections.

"I try to go to good sources, but I go to some of the sketchy sources too just to see what people say."

"[My doctor] agrees that scientifically, it doesn't work, but by the same token, he said he's referred his patients to this person to do gel injections and they've had relief, so empirically, it works for some people."

Subtheme 15: accuracy of injection location. Participants wanted information about accuracy of the procedure and were concerned about the physician's ability to inject the specific location that would provide pain relief. Imaging-guided injections assuaged these fears for some participants.

"[S]ome people say the injections should be done under fluoroscopy…some people say that doesn't matter. It would be helpful to know the precision of where the injection is going, if that's important, or if it just needs to be somewhere in the joint."

Subtheme 16: procedure explanation. Participants found it unsettling when their physician did not fully explain the injection process.

"I don't know what they shootin' in there… [The doctor] had a big old needle. I don't know what that is about. You got to explain it."

Subtheme 17: treatment mechanism. Participants' opinions of an injectable product were shaped by their understanding of whether it worked by simply relieving pain or by addressing an underlying mechanism. They preferred injectable products they perceived as improving an underlying knee problem.

"Well, I don't want anything that's gonna cause more damage… [The steroids are] gonna take the pain away, but they aren't gonna make your knee better."

For those with specific types of symptoms, pain relief was deemed an insufficient marker of success.

"I would want a cushioning. That's all I want, so it wouldn't grind and scrape."

Subtheme 18: out-of-pocket costs. Participants' decisions regarding which product to choose involved weighing each injectable's out-of-pocket costs and likelihood of success. Participants were willing to pay more if they were in more pain.

"Depending on how uncomfortable I was at the time, I think that's probably what would be the tipping point for me. But it is expensive."

"I'm like, 'I'm not gonna pay $150 for something that's probably not going to work', and that's how I ended up doing the cortisone."

Theme 4: availability and access issues. Participants' injection options were shaped by external factors, including the injectable therapies on the market, the products that physicians offered them, and insurance coverage.

Subtheme 19: limited effectiveness of injectables. Many participants were frustrated by the lack of treatment options that were certain to relieve pain. They felt they had to choose between injections with uncertain efficacies.

"It's hard because you don't know. I've had so many times where it hasn't worked, and it just gets disappointing."

Subtheme 20: variation in product availability. The injectable products that were offered to participants varied by region, hospital, and physician. Choices were limited by insurance coverage and what was available on the hospital formulary.

"I said to [my rheumatologist], I said what's there left for me? He said surgery, but you're not really ready yet. I said well, I'm gonna go look for gel. He said you better go look outside the [hospital]."

Subtheme 21: availability of alternate providers. Some participants knew of alternative injectable products, and when their physician did not offer these treatments, they were willing to find another provider who would.

"I did seek out a doctor that did PRP based on the recommendation of my physical therapist… He helped me find some doctors that did it."

Subtheme 22: differences in insurance coverage. Variation in insurance coverage limited the range of injectable products available to participants.

"My expectation is the insurance was gonna cover it because of my condition. I was really surprised that it wasn't covered. That played into my decision not to go for it."

DISCUSSION

We conducted focus groups and individual interviews with 15 patients with knee OA to better understand participants' perceived facilitators of and barriers to IA injections. We found that participants' decisions were influenced most by the following: the impact of knee OA on their lives; attitudes toward side effects, uncertain efficacy, specific features of injectables; ways in which they gathered and processed information; and the availability of different injectable products.

Several aspects of the injection decision-making process portrayed in Figure 1 merit comment. Participants' attitudes were influenced by the content and sources of information they sought, as they balanced physicians' explanations with their peers' experiences. They noted the lack of consensus among physicians regarding efficacy of different injectable products and the individuality of injection response. They relied on their physician's recommendations in conjunction with other factors to make decisions. Participants were interested in the injectable's mechanism, and some expressed fear of products they considered "unnatural." They perceived that "naturally" derived products, like PRP or HA derivatives, would promote healing better than glucocorticoids.

Discrete-choice experiments have found that the factors most pertinent to OA treatment decisions include evaluation of the benefits, risk of side effects, cost, and mode of administration (14). Other studies have highlighted the role of emotion and biases in patient decision-making (15). Our work also aligns with the framework of preference-based decision-making, which acknowledges that the right decision for any patient involves an integration of knowledge of risks and benefits with a weighting of these factors by their importance (16).

We identified themes that are consistent with prior qualitative research on OA treatments. Participants discussed how the impact of OA on their lives and the failure of prior treatments prompted them to consider injections. The influence of ineffective prior treatments has been described as a motivating factor for patients who undergo TKR as well (17). A study by Selten et al demonstrated that participants were most concerned about a treatment's efficacy, accessibility, and potential side effects (10), which are concerns that our participants shared. Those who chose injections despite these concerns indicated that after failing other treatments, an injection could be worth the potential downsides (e.g., out-of-pocket cost, potential side effects), especially if it could help them avoid surgery or return to their valued activities. Other studies have similarly found that participants expressed concerns about side effects, including the effect of glucocorticoids on cartilage (18). Consistent with our findings regarding participants' gathering and processing of information, having a good relationship with and trust in one's physician is an important factor in the OA treatment decision-making process (19,20), as is a preference for personalized treatment recommendations (10). Finally, the burden of out-of-pocket costs and concerns surrounding insurance coverage are widespread and should be part of the shared decision (10).

As this is qualitative work, these results should be viewed as hypothesis-generating. Participants were recruited from a single tertiary medical center and may not be representative of the general knee OA population. Furthermore, subjects were quite knowledgeable, which may have contributed to the achievement of thematic saturation with 15 participants; while the importance that participants attached to relevant considerations differed, the factors they considered were similar. We did not collect information about participants' socioeconomic or insurance status, each of which may influence their views of injections. Finally, while all moderators used the same guide, it is possible that having different moderators across the focus groups and interviews may have introduced bias.

These results may help guide further research and have implications for clinical practice. Additional qualitative and cross-sectional studies should be conducted to better understand these issues. Survey studies could quantify the prevalence of the attitudes toward injections documented in this study. A better understanding of the considerations that patients weigh when making injection-related decisions could help clinicians provide relevant information during shared decision-making discussions.

Our findings suggest that it would be useful for providers to help patients sort through these concerns, including the process of receiving an injection and its mechanism, content, cost, and efficacy. Given the qualitative nature of this research, we recommend that these clinical suggestions be interpreted alongside clinical guidelines and recommend that future studies investigate the issues raised here to further examine the decision-making process surrounding IA injections.

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. Lenhard, Lape, MacFarlane, Losina, Katz.
Acquisition of data. Lenhard, Lape, MacFarlane, Losina, Katz.
Analysis and interpretation of data. Lenhard, Lape, MacFarlane, Losina, Katz.

ROLE OF THE STUDY SPONSOR

Flexion Therapeutics provided funding for the study and approved the manuscript for submission. Flexion provided no input into design, data collection, data analysis, interpretation, or writing the manuscript. Publication was not contingent upon approval by Flexion.

REFERENCES

1. Deshpande BR, Katz JN, Solomon DH, Yelin EH, Hunter DJ, Messier SP, et al. Number of persons with symptomatic knee osteoarthritis in the US: impact of race and ethnicity, age, sex, and obesity. Arthritis Care Res (Hoboken) 2016;68:1743–50.

2. Bannuru RR, Osani MC, Vaysbrot EE, Arden NK, Bennell K, Bierma-Zeinstra SM, et al. OARSI guidelines for the non-surgical management of knee, hip, and polyarticular osteoarthritis. Osteoarthritis Cartilage 2019;27:1578–89.
3. Jevsevar DS. Treatment of osteoarthritis of the knee: evidence-based guideline, 2nd edition. J Am Acad Orthop Surg 2013;21:571–6.
4. Gato-Calvo L, Magalhaes J, Ruiz-Romero C, Blanco FJ, Burguera EF. Platelet-rich plasma in osteoarthritis treatment: review of current evidence. Ther Adv Chronic Dis 2019;10:2040622319825567.
5. Rutjes AW, Juni P, da Costa BR, Trelle S, Nuesch E, Reichenbach S. Viscosupplementation for osteoarthritis of the knee: a systematic review and meta-analysis. Ann Intern Med 2012;157:180–91.
6. Wehling P, Evans C, Wehling J, Maixner W. Effectiveness of intra-articular therapies in osteoarthritis: a literature review. Ther Adv Musculoskeletal Dis 2017;9:183–96.
7. Richardson SS, Schairer WW, Sculco TP, Sculco PK. Comparison of infection risk with corticosteroid or hyaluronic acid injection prior to total knee arthroplasty. J Bone Joint Surg Am 2019;101:112–8.
8. Pouli N, Das Nair R, Lincoln NB, Walsh D. The experience of living with knee osteoarthritis: exploring illness and treatment beliefs through thematic analysis. Disabil Rehabil 2014;36:600–7.
9. Ballantyne PJ, Gignac MA, Hawker GA. A patient-centered perspective on surgery avoidance for hip or knee arthritis: lessons for the future. Arthritis Rheum 2007;57:27–34.
10. Selten EM, Vriezekolk JE, Geenen R, van der Laan WH, van der Meulen-Dilling RG, Nijhof MW, et al. Reasons for treatment choices in knee and hip osteoarthritis: a qualitative study. Arthritis Care Res (Hoboken) 2016;68:1260–7.
11. Clarke V, Braun V. Thematic analysis. J Posit Psychol 2017;12:297–8.
12. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3:77–101.
13. Braun V, Clarke V. To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. Qual Res Sport Exerc Health 2019:1–16.
14. Hiligsmann M, Pinto D, Dennison E, Al-Daghri N, Beaudart C, Branco J, et al. Patients' preferences for osteoarthritis treatment: the value of stated-preference studies. Aging Clin Exp Res 2019;31:1–3.
15. Ubel PA. Beyond costs and benefits: understanding how patients make health care decisions. Oncologist 2010;15 Suppl 1:5–10.
16. Fraenkel L. Incorporating patients' preferences into medical decision making. Med Care Res Rev 2013;70 Suppl:80–93S.
17. Hsu KY, Tsai YF, Yeh WL, Chen DW, Chen CY, Wang YW. Triggers and decision-making patterns for receiving total knee arthroplasty among older adults with knee osteoarthritis: a qualitative descriptive study. J Clin Nurs 2018;27:4373–80.
18. Papandony MC, Chou L, Seneviwickrama M, Cicuttini FM, Lasserre K, Teichtahl AJ, et al. Patients' perceived health service needs for osteoarthritis (OA) care: a scoping systematic review. Osteoarthritis Cartilage 2017;25:1010–25.
19. Traumer L, Sorensen EE, Kusk KH, Skou ST. Investigating the motives of patients with knee OA undergoing a TKR: a qualitative interview study. Musculoskeletal Care 2018;16:380–7.
20. Alami S, Boutron I, Desjeux D, Hirschhorn M, Meric G, Rannou F, et al. Patients' and practitioners' views of knee osteoarthritis and its management: a qualitative interview study. PloS One 2011;6:e19634.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 3, March 2022, pp 420–426
DOI 10.1002/acr.24474

REVIEW

Increasing Ancestral Diversity in Systemic Lupus Erythematosus Clinical Studies

Jessica N. Williams,[1] Maria Dall'Era,[2] S. Sam Lim,[3] Candace H. Feldman,[1] Kathleen A. Arntsen,[4] Ashira D. Blazer,[5] Tawara Goode,[6] Joan T. Merrill,[7] Saira Sheikh,[8] Anne M. Stevens,[9] Peter E. Lipsky,[10] and Karen H. Costenbader[1]

Non-White people are more likely to develop systemic lupus erythematosus (SLE) yet are underrepresented in SLE clinical trials. The efficacy and safety of drugs may be influenced by ancestry, and ancestrally diverse study populations are necessary to optimize treatments across the full spectrum of patients. However, barriers to entry into clinical trials are amplified in non-White populations. To address these issues, a conference was held in Bethesda, Maryland, from October 15–16, 2019, entitled "Increasing Ancestral Diversity in Systemic Lupus Erythematosus Clinical Studies: Overcoming the Barriers." Conference participants included people with lupus, lupus physicians, lupus clinical trialists, treatment developers from biotechnology, social scientists, patient advocacy groups, and US government representatives (The Office of Minority Health, Centers for Disease Control and Prevention, National Institutes of Health, and the Food and Drug Administration). For all these groups, the organizers of the conference purposefully included people of non-White ancestry. Decreased participation of non-White SLE patients in clinical research was evaluated through historical, societal, experiential, and pragmatic perspectives, and several interventional programs to increase non-White patient participation in SLE and non-SLE research were described and discussed. The presentations and discussions highlighted the need for changes at the societal, institutional, research team, referring physician, and patient education levels to achieve equitable ancestral representation in SLE clinical studies.

INTRODUCTION

This article summarizes the findings of a conference entitled "Increasing Ancestral Diversity in Systemic Lupus Erythematosus Clinical Studies: Overcoming the Barriers," held in Bethesda, Maryland between October 15 and 16, 2019. The majority of the audience, 71% of the breakout session moderators, and 44% of the speakers were non-White people. The problem of decreased participation of non-White patients with systemic lupus erythematosus (SLE) in clinical research and strategies was described and discussed as shown in the present article.

BACKGROUND

Racial epidemiology of SLE and minority underrepresentation in SLE clinical trials

The prevalent SLE population in the US is 44–67% non-White (1–6). This high proportion of disease in racial minorities has been attributed to genetic and environmental factors (7). Additionally, racial minorities are more likely to have severe SLE, with SLE being the fifth leading cause of death for Black and Latinx women under the age of 25 years in the US (8). Despite being disproportionately affected by SLE, racial minorities are

Presented in part at a conference entitled "Increasing Ancestral Diversity in Systemic Lupus Erythematosus Clinical Studies: Overcoming the Barriers," held in Bethesda, Maryland, between October 15 and 16, 2019.

[1]Jessica N. Williams, MD, MPH, Candace H. Feldman, MD, ScD, Karen H. Costenbader, MD, MPH: Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts; [2]Maria Dall'Era, MD: University of California San Francisco School of Medicine; [3]S. Sam Lim, MD, MPH: Emory University, Atlanta, Georgia; [4]Kathleen A. Arntsen, BA: Lupus and Allied Diseases Association, Inc., Verona, New York; [5]Ashira D. Blazer, MD, MSCI: Langone Medical Center, New York, New York; [6]Tawara Goode, MA: Georgetown University Medical Center, Washington, DC; [7]Joan T. Merrill, MD: Oklahoma Medical Research Foundation, Oklahoma City; [8]Saira Sheikh, MD: University of North Carolina School of Medicine, Chapel Hill; [9]Anne M. Stevens, MD, PhD: University of Washington School of Medicine, Seattle, and Immunology Therapeutic Area, Janssen Research & Development LLC, Spring House, Pennsylvania; [10]Peter E. Lipsky, MD: RILITE Research Institute and AMPEL BioSolutions, Charlottesville, Virginia.

The conference summarized in this work was sponsored by Bristol Myers Squibb, Amgen, GlaxoSmithKline, Biogen, Gilead, Pfizer, EMD Serono, Genentech, Eli Lilly, AstraZeneca, Celgene, the Lupus and Allied Diseases Association, the Lupus Research Alliance, and the Lupus Foundation of America. No other disclosures relevant to this article were reported.

Address correspondence to Jessica N. Williams, MD, MPH, Emory University, 80 Jesse Hill Jr Dr SE, Atlanta, GA 30303. Email: jessica.williams2@emory.edu.

Submitted for publication September 8, 2020; accepted in revised form September 29, 2020.

typically underrepresented in SLE clinical studies. This is problematic as treatment responses may vary by race/ethnicity. For example, Black patients with lupus nephritis may have a less favorable response to intravenous cyclophosphamide (9), but a more favorable response to rituximab (10), than White patients.

A systematic review of 193 SLE randomized controlled trials (RCTs) carried out from 1997 to 2017 showed that White participants were overrepresented (51% of RCT participants and 33% of prevalent SLE cases), as were Latinx patients (21% of RCT participants and 16% of prevalent SLE cases) (11). Underrepresented groups included Black patients (14% of RCT participants and 43% of prevalent SLE cases), and to a lesser degree, Asian patients (10% of RCT participants and 13% of prevalent SLE cases). This occurred despite multiple National Institutes of Health (NIH) policies designed to increase minority representation in research that were implemented before and during the study period (12–14).

One likely contributing factor to the underrepresentation of racial minorities in SLE clinical studies is lack of diversity among rheumatologists and SLE researchers in the US. For example, according to the American College of Rheumatology (ACR) 2015 Rheumatology Workforce Study, only 0.8% of the 1,000 adult rheumatologists surveyed self-identified as a Black physician (15). Improving diversity among the rheumatology clinical and research workforce may be a critical step in the quest to achieve racial equity in SLE clinical studies, but this subject was beyond the scope of the conference *(Titilola Falasinnu, PhD, Stanford University)*.

Ancestral diversity in SLE: barriers to clinical trial enrollment

Genetic differences between racial groups may contribute to SLE treatment responses. Single-nucleotide polymorphisms (SNPs) are mutations of a single DNA base pair and account for a large amount of human genetic variation (16). Fewer SLE-associated SNPs have been identified in Black and Latinx populations compared to White populations. For example, a 2017 genome-wide association study of 11,590 individuals with SLE and 15,984 control subjects identified 58 SLE-associated SNPs in White subjects, 9 in Black subjects, and 16 in Latinx subjects (17). The study population was 66% White, 20% Black, and 14% Latinx. Comparing subjects with similar numbers of SLE-associated SNPs, White subjects had a 10-fold higher risk of having SLE than Black subjects. Fewer Black and Latinx subjects assessed in the study and the development of the SNP assay (ImmunoChip) largely in White subjects may have contributed to this seeming racial disparity in SLE-associated SNPs.

A method known as expression quantitative trait loci mapping links genetic variants to changes in gene expression in different tissues. This method may identify ancestry-dependent and ancestry-independent genes, which may in turn identify ancestry-dependent and ancestry-independent pathways (e.g., European ancestry, African ancestry, and their shared pathways). A recent study of SLE-associated SNPs mapped numerous molecular pathways (n = 1,006) unique to individuals with European ancestry and a few (n = 55) to individuals with African ancestry, as well as many shared pathways (n = 670), including interferon signaling (18). These results underscore the importance of studying SLE genetics across ancestral groups, which may elucidate drug targets that differ by ancestry.

It is important to note that there is considerable ancestral heterogeneity within self-reported racial groups in the US. For instance, the mean percentages of European ancestry among self-identified Black Americans and Latinx Americans are 24% and 65%, respectively (19). Thus, it is recommended to categorize subjects in SLE genetic studies according to genetic ancestral composition (by principal components analysis), rather than by self-reported race *(Peter E. Lipsky, MD, RILITE Research Institute and AMPEL BioSolutions)*.

UNDERSTANDING AND ADDRESSING FACTORS THAT AFFECT MINORITY ENROLLMENT IN SLE CLINICAL STUDIES

Racism in health care and medical research (Monnica Williams, PhD, University of Ottawa) and the role of cultural and linguistic competence in increasing participation of diverse populations in SLE clinical trials

Lack of trust in the research establishment is one reason for minority underrepresentation in clinical studies. For example, many Black Americans have mistrust toward clinical studies rooted in historical abuses during slavery in the US as well as due to such studies as the Tuskegee Syphilis Study, which involved withholding curative antibiotic therapy from Black male subjects who had syphilis in order to study the natural history of the disease (20,21). There are numerous other instances of mistreatment of racial minorities by medical investigators including non-anesthetized surgical experiments performed on three Black enslaved women in Alabama between 1845 and 1849 (22), a study conducted from 1946 to 1948 that involved intentionally infecting Guatemalans with sexually-transmitted diseases without informed consent (23), and an experimental hepatitis A vaccine study conducted in Lakota Sioux Indian newborns that did not disclose potential risks to parents (24) *(Tawara Goode, MA, Georgetown University)*.

Given these egregious examples, some minority patients understandably have reservations about medical research participation, which the research community should address. The first step for researchers may be to understand that culture, defined as "the learned and shared knowledge that specific groups use to generate their behavior and interpret their experience of the

world," (25) can be applied to racial and ethnic groups. It may be beneficial for SLE researchers to reflect on how their own culture, or the culture of their institution leads to racial bias in research questions and study design.

Cultural competence requires that organizations have a clearly defined, congruent set of values and principles, and demonstrate behaviors, attitudes, policies, structures, and practices that enable them to work effectively cross-culturally (26). Applying cultural and linguistic competence to research may include, but is not limited to, the steps outlined in the Georgetown University National Center for Cultural Competence Research Checklist (27), which includes the promotion of cultural and linguistic competence among research staff and the incorporation of culturally competent approaches in all steps of the research process.

EFFORTS TO INCREASE MINORITY ENROLLMENT IN NON-SLE RESEARCH STUDIES

INSPIRE Trial: a clinical trial with culturally appropriate video to increase participation of African American patients in breast cancer trials

The Increasing Participation In Research-Breast Cancer (INSPIRE-BrC) trial aimed to use a culturally targeted patient education video to increase breast cancer clinical trial enrollment among Black women (28). This intervention was successful in increasing clinical trial enrollment from 6% to 13.5% ($P < 0.001$). The INSPIRE-BrC trial demonstrated that a culturally tailored educational video could increase clinical trial participation among Black patients with breast cancer, with the added benefit of the video having the ability to be easily disseminated in a rapid, widespread fashion *(Deliya Wesley, PhD, MPH, MedStar Health Research Institute)*.

STRIDE: Strengthening Translational Research in Diverse Enrollment

The goal of the STRIDE study is to increase underrepresented minority representation in biomedical research using cultural and literacy-relevant interventions. STRIDE's aims include intervention development with community input, ongoing evaluation, and wide dissemination. Evaluations of the STRIDE intervention versus usual protocol are currently ongoing *(Stephenie Lemon, PhD, MS, University of Massachusetts)*.

EFFORTS TO INCREASE MINORITY ENROLLMENT IN SLE RESEARCH STUDIES

Pediatric experience in SLE clinical trial recruitment: APPLE trial

The Atherosclerosis Prevention in Pediatric Lupus Erythematosus (APPLE) study was a multisite clinical trial comparing the effect of treatment with atorvastatin versus placebo on intimal-medial thickness progression among 221 patients with pediatric SLE (29). This study was successful in recruiting a population that was 61% minority (27% Black, 24% Latinx, and 10% Asian). Regarding minority recruitment, there were several takeaways from the APPLE trial. Patients were more willing to participate if their treating physician discussed the study with them and was supportive of enrollment. Other helpful practices included reviewing consent forms in clinic, asking every eligible patient to participate without making assumptions about willingness or ability to participate (addressed physician implicit bias), and providing patient-centered talking points to promote trial participation *(Laura Schanberg, MD, Duke University)*.

Optimizing minority recruitment in trials: the EMBRACE experience

Belimumab was approved for the treatment of SLE by the US Food and Drug Administration (FDA) in 2011. Phase III studies did not show beneficial effects of belimumab therapy in Black patients ($n = 148$) (30,31). Given the small sample size of Black patients included in the phase III studies, the Efficacy and Safety of Belimumab in Black Race Patients with SLE (EMBRACE) study recruited patients who self-reported Black ancestry. EMBRACE did not meet its primary end point (32).

Despite multiple strategies employed to improve enrollment in EMBRACE, the trial was under-enrolled, with 448 participants who were ultimately enrolled in the trial compared to an initial enrollment target of 816 participants and a revised enrollment target of 501 participants. The investigators felt that the following enrollment strategies were not successful: 1) social media recruitment through a large SLE patient advocacy group and 2) an EMBRACE website supported by a 24-hour call center, which received few calls despite many website hits. The EMBRACE team felt that the following strategies were successful in improving enrollment: 1) enhancing physician referrals by providing study letter templates and presentation slides, 2) enhancing referrals from local patient advocacy groups via targeted funding and study letters, and 3) enhancing site capabilities (e.g., multimedia advertising support and providing study tool kits; this strategy was felt to be the most effective by the EMBRACE team) *(Susan Burriss, MS, BSN, GlaxoSmithKline)*.

Goals and activities of Centers for Disease Control and Prevention (CDC)–funded lupus projects: current efforts by the lupus community

In order to estimate SLE prevalence and incidence, the CDC funded five population-based registries targeting five racial/ethnic groups (White, Black, Latinx, Asian, and Native American/Alaskan Native) between 2003 and 2014 in Alaska, California, Georgia, Michigan, and New York. Three of these registries were expanded to form longitudinal cohorts to further study SLE

natural history, treatment, access to care, and disparities in specific populations: California for Asian/Latinx groups, Georgia for Black/White groups, and Michigan for Black/White groups (33). For example, studies performed using the Georgia SLE cohort (GOAL) found that Black patients had higher standardized mortality ratios compared to White patients (3.34 versus 2.43) (34). This work by the CDC demonstrates the importance of collecting data on race in order to better understand disparities and target future interventions *(Charles Helmick, MD, CDC)*.

ACR efforts to increase minority participation in lupus clinical trials

The ACR has implemented two interventions to increase minority participation in lupus clinical trials: Materials to Increase Minority Involvement in Clinical Trials (MIMCT; increasing physician referrals to clinical trials for Black and Latinx patients using an educational toolkit [35]), and the Lupus Clinical Trials Training Program (LuCTT; educating potential subjects about lupus clinical trials via community health workers [36]). Results thus far from MIMICT suggest that the educational toolkit improved primary care physician knowledge and the intention to refer. LuCTT accomplishments to date include partnerships between the ACR and state and local organizations, as well as training of 73 community health workers *(Sheryl McCalla, JD, ACR)*.

PALS: Patient Advocates for Lupus Studies

PALS is an ongoing peer support program wherein trained lupus patients serve as educators and advocates for clinical trial participation, focusing on diverse populations. PALS was inspired by the Patient Partners in Arthritis program, started at the University

Table 1. Qualitative findings of patient and investigator breakout sessions regarding minority participation in systemic lupus erythematosus (SLE) clinical trials

Motivators for participation, patient session	Barriers to participation, patient session	Barriers to participation, investigator session
	Patient barriers	Societal barriers
Lack of other therapeutic options	Concerns about being a "guinea pig"	Implicit racial bias of researchers and referring physicians
Recommended by a trusted physician	Lack of trust in the healthcare system (often related to negative prior personal or family experiences)	Structural racism (that may influence the ability and willingness of minorities to participate)
Trust in the research institution	Discouragement to participate from others (including loved ones and religious leaders)	Partisan political climate with scarce research dollars
Desire to help the greater SLE community (including family members and other minority patients with SLE)	Lack of familiarity with and knowledge about clinical trials	Institutional barriers (National Institutes of Health, academia, pharmaceutical industry, among others)
Personal empowerment	Being overwhelmed or in denial about SLE diagnosis	Lack of prioritization of trial diversity
Compensation	Not being asked to participate	Underdeveloped relationships with communities of color
Greater access to care	Meeting exclusion criteria despite desire to participate (e.g., lupus nephritis)	Distance of research sites from communities of color
	Referring physician barriers	Lack of outreach to remote areas (e.g., through telemedicine, social media, or mobile units)
	Lack of time to discuss clinical trials with patients	Intimidating informed consent process
	Competing clinical priorities	Rejection of uninsured and underinsured patients
	Lack of reimbursement for trial recruitment	Lack of investment in metrics important to involved parties (health for patients, relative value units for treating physicians, grants and publications for researchers, and first patient enrolled for pharmaceutical companies)
	Lack of prioritization of trial diversity	Lack of research team diversity
		Referring physicians' fear of losing patients
		Lack of communication to patients and treating physicians regarding study findings
		Patient barriers
		Desire for compensation (including childcare, transportation reimbursement)
		Desire for less strict eligibility criteria
		Language barriers
		Desire to hear about trial from treating medical doctor rather than a research team

of Texas Southwestern in the early 1990s, a program that successfully trained patients with arthritis to serve as peer educators and facilitate participation in arthritis-related research, including clinical trials (37) *(S. Sam Lim, MD, MPH, Emory University)*.

PURPLE: Programs to Address Unmet Needs and Promote Representation of all Participants in Lupus Clinical Trials using Mobile Technology for Engagement

Smartphone ownership rates are similar among Black, White, and Latinx patients, but Black and Latinx patients are more likely to look up health-related information on their smartphones (38). The PURPLE intervention was designed to increase minority enrollment in lupus clinical trials using a patient-focused digital tool delivering culturally and linguistically appropriate clinical trial education, delivered by custom-built avatars that are modeled after the patient's treating physician. PURPLE was developed by Dr. Sheikh's team at The University of North Carolina-Chapel Hill Thurston Arthritis Research Center *(Saira Sheikh, MD, University of North Carolina-Chapel Hill)*.

IMPACT for lupus: a faith-based approach to awareness and participation in lupus clinical trials

Improving Minority Participation and Awareness in Clinical Trials (IMPACT), which is funded by the US Department of Health and Human Services (HHS) Office of Minority Health (OMH), is designed to improve minority patient awareness of lupus clinical trials using a community and faith-based participatory research approach (39). IMPACT utilizes a trained navigator who acts as an intermediary between the patient, his or her church, his or her physician, the Lupus Foundation of America, and the clinical trial coordinator. The results of the three-month pilot study indicated the IMPACT program was effective in raising awareness of clinical trials among Black patients with lupus *(Patricia Davidson, MPA, Lupus Foundation of America)*.

The CLIMB Project 2019: Connecting People with Lupus to Improve Meaningful Benefits from Trials

The CLIMB project aimed to improve minority enrollment in lupus clinical trials by focusing on the informed consent process and the education of primary care clinicians and patients. Following the CLIMB educational intervention, a post-questionnaire showed an increased numbers of patients who preferred to learn about clinical trials from someone who had the same racial/ethnic background, suggesting the intervention process may have uncovered previously unstated preferences. After completing the CLIMB modules on the history and nature of clinical trials, fewer patients and clinicians stated a belief that poverty will reduce a clinical trial participant's adherence to study protocol *(Joan T. Merrill, MD, Oklahoma Medical Research Foundation)*.

Table 2. Best practices for clinicians and researchers to increase ancestral diversity in systemic lupus erythematosus (SLE) clinical studies

For clinicians:
Understand the historical exploitation of racial minorities by medical researchers and ongoing structural racism in the US, and how this has led to mistrust of clinical research among these populations.
Reflect on how personal or institutional culture may lead to racial bias in research questions and study design.
Increase recruitment of underrepresented racial minorities (Black, Latinx, and Native American/Alaskan Native physicians) to the field of rheumatology.
Discuss research studies with eligible patients in a supportive manner.
Avoid assumptions (implicit bias) about the "kind" of patients who would be willing to participate in a clinical trial.
For researchers:
Understand the historical exploitation of racial minorities by medical researchers and ongoing structural racism in the US, and how this has led to mistrust of clinical research among these populations.
Reflect on how personal or institutional culture may lead to racial bias in research questions and study design.
Encourage research careers among underrepresented minority rheumatologists through mentorship and career support.
Collect ancestral data in every research study involving human subjects.
Develop means to classify subjects in SLE genetic studies according to genetic ancestral composition, rather than by self-reported race (given the ancestral heterogeneity of self-reported racial groups in the US).
Recruit a racially diverse research study team, including those in leadership positions.
Train all research staff in cultural and linguistic competence.
Incorporate cultural and linguistic competence in every step of the research process, with a particular emphasis on informed consent and the ability of subjects to withdraw from the study at any time.
Include ancestrally diverse patients in the study planning process.
Form relationships with people and institutions within the target community to be studied.
Tailor patient recruitment and educational interventions to the target racial group, using input from members of the community.
Consider providing transportation, childcare, and monetary compensation to potential study subjects in order to enable their participation.
Stimulate a discussion with the institutional review boards concerning the topic of reimbursement versus coercion and include racially and economically diverse patients in the discussion.
Provide educational and study recruitment materials to referring physicians and patient advocacy groups.
Train racially diverse community members, including those with SLE, to educate and recruit potential SLE research subjects.

Lupus conversations modules: an academic-community partnership

The HHS OMH-funded Lupus Conversations Modules is a program designed to increase enrollment of Black patients in lupus clinical trials through academic-community partnerships in Boston and Chicago. This program involves identifying influential community members to serve as popular opinion leaders, who are trained to disseminate information about lupus and clinical trials throughout their social network. Accomplishments to date include authorship of a systematic literature review regarding minority underrepresentation in rheumatology research (40), development of curricula to train popular opinion leaders and physicians about lupus clinical trial recruitment, and development of popular opinion leader– and physician-led curricula to educate patients about clinical trials *(Candace Feldman, MD, ScD, Harvard University, and Rosalind Ramsey-Goldman, MD, DrPH, Northwestern University)*.

BREAKOUT SESSIONS: QUALITATIVE EVALUATION OF MINORITY PARTICIPATION IN SLE CLINICAL TRIALS

Three breakout sessions addressed minority participation in lupus clinical trials. One session was patient-focused and featured 20 minority patients with SLE, nine of whom had clinical trial experience, as well as three lupus physicians. The other two breakout sessions were focused on the investigator perspective, and featured clinical trial investigators, treating physicians, patients, patient advocates, and treatment developers. See Table 1 for a summary of the motivators and barriers to minority clinical trial participation that were identified during these three breakout sessions.

CONCLUSIONS

Non-White people are more likely to develop SLE and severe SLE but are underrepresented in lupus clinical trials. Efficacy and safety of treatments may be influenced by ancestry, and ancestrally diverse study populations are necessary to ensure a better understanding of optimal treatment for all patients. Impediments to enrollment of minority patients may include an unfavorable view of biomedical research because of historical injustices and ongoing structural racism and bias. To advance the field, acknowledgement of these realities and improvements in cultural and linguistic competence might help to overcome these barriers. Several interventions to increase minority participation in research for SLE and non-SLE conditions have been described in this work. Many of these interventions focused on Black patients, who are the most underrepresented in international SLE clinical trials. Going forward, to represent the people who live in the US, there is a need to include programs for the growing Latinx and Asian populations in the US, as well as Native Americans, Pacific Islanders, Alaskan Natives, and men with SLE. Table 2 summarizes best practices for clinicians and researchers to increase ancestral diversity in SLE clinical studies. In conclusion, input from patients and investigators has uncovered the need for change at the societal, institutional, research team, referring physician, and patient level in order to achieve racial equity in lupus clinical studies.

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.

REFERENCES

1. Helmick CG, Felson DT, Lawrence RC, Gabriel S, Hirsch R, Kwoh CK, et al. Estimates of the prevalence of arthritis and other rheumatic conditions in the United States. Part I. Arthritis Rheum 2008;58:15–25.
2. Feldman CH, Hiraki LT, Liu J, Fischer MA, Solomon DH, Alarcón GS, et al. Epidemiology and sociodemographics of systemic lupus erythematosus and lupus nephritis among US adults with Medicaid coverage, 2000–2004. Arthritis Rheum 2013;65:753–63.
3. Dall'Era M, Cisternas MG, Snipes K, Herrinton LJ, Gordon C, Helmick CG. The incidence and prevalence of systemic lupus erythematosus in San Francisco County, California: the California Lupus Surveillance Project. Arthritis Rheumatol 2017;69:1996–2005.
4. Izmirly PM, Wan I, Sahl S, Buyon JP, Belmont HM, Salmon JE, et al. The incidence and prevalence of systemic lupus erythematosus in New York County (Manhattan), New York: the Manhattan Lupus Surveillance Program. Arthritis Rheumatol 2017;69:2006–17.
5. Somers EC, Marder W, Cagnoli P, Lewis EE, DeGuire P, Gordon C, et al. Population-based incidence and prevalence of systemic lupus erythematosus: the Michigan Lupus Epidemiology and Surveillance Program. Arthritis Rheumatol 2014;66:369–78.
6. Lim SS, Bayakly AR, Helmick CG, Gordon C, Easley KA, Drenkard C. The incidence and prevalence of systemic lupus erythematosus, 2002-2004: the Georgia Lupus Registry. Arthritis Rheumatol 2014; 66:357–68.
7. Gergianaki I, Bortoluzzi A, Bertsias G. Update on the epidemiology, risk factors, and disease outcomes of systemic lupus erythematosus. Best Pract Res Clin Rheumatol 2018;32:188–205.
8. Yen EY, Singh RR. Brief report: lupus-an unrecognized leading cause of death in young females: a population-based study using nationwide death certificates, 2000–2015. Arthritis Rheumatol 2018;70:1251–5.
9. Dooley MA, Hogan S, Jennette C, Falk R. Cyclophosphamide therapy for lupus nephritis: poor renal survival in black Americans. Glomerular Disease Collaborative Network. Kidney Int 1997;51:1188–95.
10. Merrill JT, Neuwelt CM, Wallace DJ, Shanahan JC, Latinis KM, Oates JC, et al. Efficacy and safety of rituximab in moderately-to-severely active systemic lupus erythematosus: the randomized, double-blind, phase II/III systemic lupus erythematosus evaluation of rituximab trial. Arthritis Rheum 2010;62:222–33.
11. Falasinnu T, Chaichian Y, Bass MB, Simard JF. The representation of gender and race/ethnic groups in randomized clinical trials of individuals with systemic lupus erythematosus. Curr Rheumatol Rep 2018; 20:20.
12. National Institutes of Health. NIH policy and guidelines on the inclusion of women and minorities as subjects in clinical research.

URL: https://grants.nih.gov/policy/inclusion/women-and-minorities/guidelines.htm.

13. US Department of Health and Human Services. National Institutes of Health. Outreach notebook for the NIH guidelines on inclusion of women and minorities as subjects in clinical research. Bethesda, Maryland. NIH publication no. 03-7036. 2002.
14. US Department of Health and Human Services. National Institutes of Health. Monitoring adherence to the NIH Policy on the inclusion of women and minorities as subjects in clinical research. comprehensive report: tracking of human subjects research as reported in fiscal year 2006 and fiscal year 2007. 2008.
15. American College of Rheumatology. 2015 workforce study of rheumatology specialists in the United States. URL: https://www.rheumatology.org/portals/0/files/ACR-Workforce-Study-2015.pdf.
16. 1000 Genomes Project Consortium. A global reference for human genetic variation. Nature 2015;526:68–74.
17. Langefeld CD, Ainsworth HC, Cunninghame Graham DS, Kelly JA, Comeau ME, Marion MC, et al. Transancestral mapping and genetic load in systemic lupus erythematosus. Nat Commun 2017;8:16021.
18. Owen K, Aidukaitis B, Labonte A, Catalina M, Bachali P, Geraci N, et al. The integration of genetic data, molecular pathway analysis and differential expression to delineate the impact of ancestral differences on lupus [abstract]. Arthritis Rheumatol 2019;71.
19. Bryc K, Durand EY, Macpherson JM, Reich D, Mountain JL. The genetic ancestry of African Americans, Latinos, and European Americans across the United States. Am J Hum Genet 2015;96:37–53.
20. Heintzelman C. The Tuskegee Syphilis Study and its implications for the 21st century. The New Social Worker 2003;10.
21. Gamble VN. Under the shadow of Tuskegee: African Americans and health care. Am J Public Health 1997;87:1773–8.
22. Suite DH, La Bril R, Primm A, Harrison-Ross P. Beyond misdiagnosis, misunderstanding and mistrust: relevance of the historical perspective in the medical and mental health treatment of people of color. J Natl Med Assoc 2007;99:879–85.
23. Spector-Bagdady K, Lombardo PA. From in vivo to in vitro: how the Guatemala STD experiments transformed bodies into biospecimens. Milbank Q 2018;96:244–71.
24. Baum D. Vaccine tests racist, some Indians say. Chicago Tribune. Chicago, Illinois; 1992.
25. Gilbert J, Goode TD, Dunne C. National Center for Cultural Competence, Georgetown University Center for Child and Human Development. Cultural awareness. From the Curricula Enhancement Module Series. Washington, DC; 2007.
26. Cross TL, Bazron BJ, Dennis KW, Isaacs MR. CASSP Technical Assistance Center, Georgetown University Child Development Center. Towards a culturally competent system of care: a monograph on effective services for minority children who are severely emotionally disturbed. Washington, DC; 1989.
27. National Center for Cultural Competence, Georgetown University Center for Child and Human Development. Research—cultural and linguistic competence checklist for MCH training programs. 2010. URL: https://nccc.georgetown.edu/documents/NCCC-Research-Checklist.pdf.
28. Robinson BN, Newman AF, Tefera E, Herbolsheimer P, Nunes R, Gallagher C, et al. Video intervention increases participation of black breast cancer patients in therapeutic trials. NPJ Breast Cancer 2017;3:36.
29. Schanberg LE, Sandborg C, Barnhart HX, Ardoin SP, Yow E, Evans GW, et al. Use of atorvastatin in systemic lupus erythematosus in children and adolescents. Arthritis Rheum 2012;64:285–96.
30. Furie R, Petri M, Zamani O, Cervera R, Wallace DJ, Tegzová D, et al. A phase III, randomized, placebo-controlled study of belimumab, a monoclonal antibody that inhibits B lymphocyte stimulator, in patients with systemic lupus erythematosus. Arthritis Rheum 2011;63:3918–30.
31. Navarra SV, Guzman RM, Gallacher AE, Hall S, Levy RA, Jimenez RE, et al. Efficacy and safety of belimumab in patients with active systemic lupus erythematosus: a randomised, placebo-controlled, phase 3 trial. Lancet 2011;377:721–31.
32. Sergio Guedes Barbosa L, Saxena A, Ginzler E, Furie R, D'Cruz D, Burriss S, et al. Efficacy of belimumab in patients of black race with systemic lupus erythematosus and high disease activity or renal manifestations [abstract]. Arthritis Rheumatol 2019;71.
33. Drenkard C, Lim SS. Update on lupus epidemiology: advancing health disparities research through the study of minority populations. Curr Opin Rheumatol 2019;31:689–96.
34. Lim SS, Helmick CG, Bao G, Hootman J, Bayakly R, Gordon C, et al. Racial disparities in mortality associated with systemic lupus erythematosus — Fulton and DeKalb counties, Georgia, 2002–2016. MMWR Morb Mortal Wkly Rep 2019;68:419–22.
35. The Lupus Initiative. Materials to increase minority involvement in clinical trials (MIMICT). URL: https://thelupusinitiative.org/mimict-2/.
36. The Lupus Initiative. Lupus clinical trials training (LuCTT) program. URL: https://thelupusinitiative.org/luctt/.
37. Branch VK, Lipsky K, Nieman T, Lipsky PE. Positive impact of an intervention by arthritis patient educators on knowledge and satisfaction of patients in a rheumatology practice. Arthritis Care Res 1999;12:370–5.
38. Perrin A, Turner E. Home broadband adoption, computer ownership vary by race, ethnicity in the US. URL: https://www.pewresearch.org/fact-tank/2019/08/20/smartphones-help-blacks-hispanics-bridge-some-but-not-all-digital-gaps-with-whites/.
39. Lupus Foundation of America. Improving minorty participation and awareness in clinical trials (IMPACT). URL: https://www.lupus.org/for-researchers/improving-minority-participation-and-awareness-in-clinical-trials.
40. Lima K, Phillip CR, Williams J, Peterson J, Feldman CH, Ramsey-Goldman R. Factors associated with participation in rheumatic disease–related research among underrepresented populations: a qualitative systematic review. Arthritis Care Res (Hoboken) 2020;72:1481–9.

Arthritis Care & Research
Vol. 74, No. 3, March 2022, pp 427–432
DOI 10.1002/acr.24472

American College of Rheumatology
Empowering Rheumatology Professionals

BRIEF REPORT

Prevalence of Neuropsychiatric Lupus in Psychosis Patients Who Have Tested Positive for Antinuclear Antibodies

Michael C. Spies,[1] Johannes A. Gutjahr-Holland,[2] James V. Bertouch,[2] and Anthony M. Sammel[3]

Objective. Psychosis is a rare manifestation of neuropsychiatric systemic lupus erythematosus (NPSLE). Current guidelines do not make a recommendation regarding the use of antinuclear antibody (ANA) testing in the assessment of patients with psychosis. The present study was undertaken to determine the prevalence of NPSLE in patients with psychosis who were positive for ANAs.

Methods. A retrospective review of patients who were admitted to the mental health service of 2 metropolitan tertiary referral centers with a diagnosis of psychosis and had been tested for ANAs was conducted. A diagnosis of SLE was made when the 2019 American College of Rheumatology (ACR)/European Alliance of Associations for Rheumatology (EULAR) classification criteria were fulfilled. Attribution of psychosis-related events to NPSLE were made according to validated criteria.

Results. There were 10,205 mental health admissions with diagnoses of psychosis representing 4,766 individual patients, 911 patients (19%) were tested for ANAs, 135 (15%) of those tests returned a positive result with a titer of ≥1:160. The mean ± SD follow-up time was 47 ± 26 months. At discharge, there were 4 patients who met 2019 ACR/EULAR criteria for SLE, 2 of whom met criteria for NPSLE (2 patients had other manifestations of SLE), yielding an NPSLE prevalence of 1.5% (2 of 135) among patients who were positive for ANAs, and 0.2% (2 of 911) among all patients who underwent testing for ANAs.

Conclusion. The prevalence of NPSLE in patients with psychosis who were positive for ANAs was low, at 1.5%. The low rate of clinically significant positive results would argue against routine testing for ANAs in patients with psychosis.

INTRODUCTION

Neuropsychiatric systemic lupus erythematosus (NPSLE) is a heterogeneous group of disorders involving a range of clinic syndromes. Psychosis is 1 of the 19 neuropsychiatric syndromes defined by the American College of Rheumatology (ACR) 1999 criteria (1).

The reported prevalence of neuropsychiatric disease in SLE is highly variable (2), with the most common manifestations being headache, mood disorder, and cognitive dysfunction. Psychosis represents a smaller subset of this, with variable prevalence of 0–12%, depending on the cohort (3,4).

Patients with SLE who develop psychosis may have psychosis as part of their initial presentation (3), meaning psychosis can be a presenting feature of previously undiagnosed SLE. Assessment of psychosis involves the exclusion of organic disorders as a precipitant. Various clinical and laboratory investigations, including serum chemistry, blood count, and tests for syphilis, hepatitis C, and HIV may be performed, depending on the initial presentation (5).

Antinuclear antibodies (ANAs) are the typical serologic finding in patients with SLE, including those with NPSLE. An absence of ANAs is rare in patients with clinical lupus. ANAs are present in at least 90% of patients with psychosis as a clinical syndrome of NPSLE (3,4), and as such a negative test makes SLE unlikely.

Currently, there is limited evidence to support or refute the use of ANA testing to screen for SLE or connective tissue disease in patients with psychosis. In a retrospective study of 85 patients with psychosis who were tested for ANAs, Mantovani and colleagues found that 3 patients were ANA positive, 2 of whom were diagnosed and treated for NPSLE (6). A prospective study by

[1]Michael C. Spies, MBBS: Royal Prince Alfred Hospital, Camperdown, New South Wales, Australia; [2]Johannes A. Gutjahr-Holland, MBBS, James V. Bertouch, MBBS, MD, FRACP: Prince of Wales Hospital, Randwick, New South Wales, Australia; [3]Anthony M. Sammel, MBBS, FRACP: Prince of Wales Hospital and Prince of Wales Clinical School, University of New South Wales, Randwick, New South Wales, Australia.

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

Address correspondence to Michael C. Spies, MBBS, Department of Rheumatology, Royal North Shore Hospital, Reserve Road St Leonards, NSW 2065, Australia. Email: Michael.Spies@health.nsw.gov.au.

Submitted for publication May 9, 2020; accepted in revised form September 24, 2020.

SIGNIFICANCE & INNOVATIONS

- A new diagnosis of neuropsychiatric systemic lupus erythematosus (NPSLE) is made in <2% of psychosis patients who have tested positive for antinuclear antibodies (ANAs).
- Of all patients admitted with psychosis who were tested for ANAs, only 0.2% had a diagnosis of NPSLE.
- Testing for ANAs to screen for NPSLE in patients with psychosis is likely to have a very low yield.

Audemard-Verger et al of 100 patients found no episodes of NPSLE among 32 ANA-positive patients in a more heterogeneous group of mental health disorders (7). The American Psychiatric Association Practice guideline for the treatment of patients with schizophrenia does not make a recommendation regarding the use of testing for ANAs as a part of the assessment in patients with psychosis (5). Given the small size of previous study cohorts and discrepancy in findings, we designed a study to clarify the significance of ANA positivity in patients with psychosis.

The primary objective of this study was to determine the prevalence of NPSLE in patients admitted to a mental health service with a diagnosis of psychosis who were positive for ANAs. The secondary objectives were to determine the frequency of testing for and proportion of positive ANAs in this patient group. We also sought to determine the pattern and titers of positive ANAs, as well as the subsequent investigation, referral, and diagnosis of ANA-positive patients.

PATIENTS AND METHODS

We performed a retrospective chart review of patients admitted to the mental health service of 2 Sydney metropolitan tertiary referral centers (Prince of Wales Hospital and Royal Prince Alfred Hospital) with a diagnosis of psychosis who had been tested for ANAs. Ethics approval was provided by South Eastern Sydney Local Health Human Research Ethics Committee (reference number 18/121). Patients were identified using their electronically entered diagnoses based on the International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10) codes (8). We included ICD-10 codes for primary psychotic disorders or mood disorders with psychotic features (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24472/abstract). To ensure that patients with a final diagnosis of SLE were not missed, we also searched for ICD-10 SLE diagnoses in patients admitted under mental health (see Supplementary Table 2, online at http://onlinelibrary.wiley.com/doi/10.1002/acr.24472/abstract).

Patients included in the present study were admitted between January 1, 2010 and March 31, 2018. The electronic medical record, which records all inpatient pathology tests for that period, was then reviewed to determine whether ANA testing had been ordered. Titer was also recorded for positive results. A titer of ≥1:160 was considered a positive result, in line with international recommendations for interpreting ANA when screening for systemic autoimmune rheumatic diseases (9).

For those patients who were ANA positive, hard copy and electronic medical records were reviewed, and a prespecified data set was completed. Data were compiled and reviewed by 1 of 2 reviewers who were not part of patient care (MCS and JAG-H). Patients were considered to have SLE if they met the 2019 ACR/European Alliance of Associations for Rheumatology (EULAR) classification criteria (10).

Follow-up data were obtained from medical records of subsequent admissions. For those patients for whom follow-up data beyond 6 months were unavailable, outpatient treating clinicians were contacted to obtain the follow-up data.

Attribution of NPSLE. Decisions regarding attribution of psychosis-related events to SLE followed the Italian criteria used by Bortoluzzi et al, which incorporates the Systemic Lupus Erythematosus Disease Activity Index (11,12). A score of ≥7 was defined as criteria for attribution of a psychotic event to NPSLE, as per previous validation (13). NPSLE was confirmed if patients met the criteria during initial admission or when a subsequent diagnosis was made at follow-up.

Statistical analysis. Unpaired *t*-tests and chi square tests were used where applicable. *P* values less than 0.05 were considered significant. Direct comparison of data between hospitals using statistical tests was not performed due to potential for heterogeneity.

RESULTS

Between January 1, 2010 and March 31, 2018, there were 10,205 mental health admissions with an ICD-10 diagnosis of psychosis (n = 5,585 for site 1 and n = 4,620 for site 2), representing 4,766 individual patients. Of these individual patients, 911 (19%) were tested for ANAs, and 135 of the 911 patients (15%) were ANA positive with a titer of ≥1:160. The characteristics of these patients are summarized in Table 1.

The most common ANA pattern was speckled (in 67 of 135 patients), followed by homogeneous in 37 patients, nucleolar in 15, dense fine speckled in 6, mitotic spindle in 5, cytoplasmic in 2, centriole in 2, and centromere in 1. The most frequent ANA titer was 1:320 (n = 64 patients) followed by 1:160 (n = 42), 1:640 (n = 12), 1:1,280 (n = 10), 1:5,120 (n = 5), and 1:2,560 (n = 2).

During mental health admission, ANA-positive patients were most commonly diagnosed with schizophrenia (151 of 376 patients [40%]), followed by a diagnosis of schizoaffective (131 of 376 patients [35%]). Referral to a rheumatology or clinical

Table 1. Characteristics of patients admitted with psychosis*

	ANAs tested	ANAs not tested	*P*
Site 1†			
ANA positive, no. (%)	78 (17)	–	–
Age, years	43.6 ± 16.4	42.9 ± 15.5	0.39
Female sex, no. (%)	205 (46)	860 (43)	<0.01
No. of admissions	3.2 ± 4.1	2.1 ± 2.6	<0.01
Length of stay, days	40.3 ± 40.8	27.4 ± 32.1	<0.01
Patients with NPSLE, no. (%)	1 (0.2)	–	–
Site 2‡			
ANA positive, no. (%)	57 (12)	–	–
Age, years	38.7 ± 13.3	39.6 ± 12.7	0.18
Female sex, no. (%)	258 (56)	720 (39)	<0.01
No. of admissions	2.1 ± 2.4	2.0 ± 2.6	0.45
Length of stay, days	22.6 ± 23.6	19.0 ± 21.7	<0.01
Patients with NPSLE, no. (%)	1 (0.2)	–	–

* Values are the mean ± SD unless indicated otherwise. ANAs = antinuclear antibodies; NPSLE = neuropsychiatric systemic lupus erythematosus.
† For site 1, n = 2,451 patients, including those with ANAs tested (n = 449) and ANAs not tested (n = 2,002).
‡ For site 2, n = 2,315 patients, including those with ANAs tested (n = 462) and ANAs not tested (n = 1,853).

immunology service was sought in 20 of 135 ANA-positive patients (15%).

In the majority of the patients (85 of 135), the reason given for testing for ANAs was to determine an organic cause of psychosis. A total of 7 patients were tested for ANAs due to elevated liver function tests, 2 for reassessment of known SLE, 2 for pleural effusions, 2 for arthralgia, and 1 for acute kidney injury. In the remaining 36 patients, there was no clear rationale documented for testing for ANAs.

Histone antibody testing was positive in 3 patients, none of whom were being treated with recognized ANA-inducing medications. Four patients were taking chlorpromazine during admission, none of whom were tested for histone antibodies. No other recognized ANA-inducing medications were identified. Hepatitis C was a common comorbidity for patients with ANA positivity, with 24 of 135 patients (18%) having a documented or serologic history.

At discharge, there were 4 patients with documented SLE who also met the 2019 ACR/EULAR criteria for SLE (Table 2). Two of these patients had known SLE diagnosed >6 months prior to admission. Two patients met the predetermined attribution criteria for NPSLE and had clinical manifestations of SLE in addition to psychosis.

Patient 1 was a 71-year-old woman with known SLE who was diagnosed with delirium with psychotic features during a hospital admission for pulmonary hemorrhage, arthritis, fever, and pancytopenia related to her autoimmune disease. She met criteria for NPSLE and was treated with intravenous immunoglobulin, methylprednisolone, and mycophenolate, as well as with sodium valproate and olanzapine.

Patient 2 was a 29-year-old woman with known mixed connective tissue disease. During admission, she was reviewed by the immunology service and underwent magnetic resonance imaging (MRI) of the brain, which yielded normal results. She did not meet the criteria for NPSLE. She was treated with mirtazapine and quetiapine. At follow-up 2 years after index admission, her illness was being treated as SLE/rheumatoid arthritis overlap with adalimumab, azathioprine, and prednisone 10 mg daily.

Patient 3 was a 72-year-old woman presenting in the context of obstructive lung disease and long-standing paranoid delusions. She had alopecia, previous photosensitive rash, and elevated double-stranded DNA (dsDNA). After rheumatology inpatient consult service, she underwent computed tomography of the brain and cerebral single-photon–emission computed tomography, showing no features of NPSLE; she did not meet the criteria for NPSLE. She was treated with 40 mg of prednisone for obstructive lung disease. Follow-up data were not available.

Patient 4 was a 36-year-old woman with oral ulcers, arthritis, and a history of seizures. She met the criteria for NPSLE and was treated with intravenous methylprednisolone, hydroxychloroquine, rituximab, and risperidone. At follow-up, 46 months after admission, the patient was noncompliant with SLE-directed therapy but had no documented features of active SLE or psychosis.

One 32-year-old female patient developed psychosis in the context of known primary antiphospholipid syndrome with previous deep venous thrombosis and high positive antiphospholipid antibody titers. An MRI of the brain showed evidence of microvascular ischemia with small foci of T2 hyperintensity in the left frontal and left inferior cerebellar hemisphere. Psychosis was deemed secondary to primary antiphospholipid syndrome. She did not meet the criteria for SLE. Her treatment included warfarin and risperidone. She did not receive treatment with immunosuppressive drugs.

Discharge data were available for all patients, and long-term follow-up data were available for 113 of 135 patients (84%). The mean follow-up time was 47 ± 26 months.

Table 2. Patients meeting the ACR/EULAR 2019 criteria for SLE*

Patient no.	Age, years	Clinical features of SLE	Serologic features of SLE	Timing of psychosis	MRI brain	SPECT scan	Corticosteroid dose at time of psychosis	NPSLE criteria met	Treatment
Patient 1†	71	Arthritis, fever, AIHA, tcp, pulmonary hemorrhage	ANA 1:2,560 homogeneous, Ro52, Ro60, SSB, low C3, low C4, dsDNA, Sm, aPL	>6 months after SLE diagnosis	Mild diffuse atrophy	Not done	None	Yes	IVIG, IVMP, mycophenolate valproate, olanzapine
Patient 2†	29	Arthritis, serositis, RA/SLE overlap	ANA 1:2560 speckled, RNP, CCP, RF, dsDNA	>6 months after SLE diagnosis	Normal	Not done	7.5 mg/daily	No	Mirtazapine, quetiapine
Patient 3‡	72	Alopecia	ANA 1:320 speckled, dsDNA	Within 6 months of SLE diagnosis	Not done	No features of cerebral SLE	None	No	NA
Patient 4‡	36	Ulcers, arthritis, seizure	ANA 1:5120 homogeneous, dsDNA, low C3	Within 6 months of SLE diagnosis	Normal	Consistent with cerebral SLE	None	Yes	IVMP, rituximab, risperidone, hydroxychloroquine

* All patients (1–4) were female. ACR = American College of Rheumatology; AIHA = autoimmune hemolytic anemia; ANA = antinuclear antibody; aPLs = antiphospholipid antibodies; CCP = cyclic citrullinated peptide; dsDNA = double-stranded DNA; EULAR = European Alliance of Associations for Rheumatology; ITP = immune thrombocytopenia; IVIG = intravenous immunoglobulin; IVMP = intravenous methylprednisolone; MRI = magnetic resonance imaging; NA = not available; NPSLE = neuropsychiatric systemic lupus erythematosus; RA = rheumatoid arthritis; RF = rheumatoid factor; RNP = ribonucleoprotein antibody; SLE = systemic lupus erythematosus; SPECT = single-photon–emission computed tomography; SSB = La/SSB antibody; tcp = thrombocytopenia.
† Patients at site 2.
‡ Patients at site 1.

One patient, a 45-year-old woman, was diagnosed with a connective tissue disease at long-term follow-up. She was diagnosed with Sjögren's syndrome 6 years after her index admission for psychosis. This patient did not meet NSPLE criteria during her psychosis admission and did not display features of active psychosis at follow-up.

Based on the admission and follow-up data, the overall prevalence of NPSLE in patients admitted with psychosis who tested positive for ANAs was 2 of 135 (1.5%). This represents 0.2% (2 of 911) of all patients tested for ANAs.

DISCUSSION

While previous studies have shown the prevalence of neuropsychiatric events in patients with established SLE (4), there is a paucity of data regarding the use of testing for ANAs to screen for NPSLE in patients with psychosis (6,7) or the significance of a positive test. Our study demonstrates the prevalence of NSPLE in a group of selected inpatients with psychosis and ANA positivity to be <2%. Given that this is a patient group with severe disease requiring admission, it could be expected that in psychosis patients, more generally ANA testing may yield an even lower proportion of clinically significant results. The overall prevalence of NPSLE in those tested for ANA was 0.2%, suggesting the diagnostic yield in this context to be very low.

Only 2 new diagnoses of SLE were identified by testing for ANAs, and only 1 of them was determined to have NPSLE. Hence, the rate of finding a new diagnosis of SLE with associated NPSLE among all patients tested for ANA was 0.1%. This indicates that 1,000 psychosis patients would need to undergo ANA screening to identify 1 new case of NPSLE. The patients with previously known SLE did not have a change in their clinical approach due to ANA positivity, as clinicians were already alerted to the possibility of NPSLE as a cause of psychosis.

Importantly, other features of SLE, such as a high titer ANA, homogeneous pattern, elevated dsDNA antibodies, and low serum complement proteins, were all features of patients adjudged to have NPSLE. Absence of these or other features of SLE makes an underlying diagnosis of NPSLE unlikely (3). As psychosis itself is a manifestation of SLE, ANA testing may still be considered, especially where there are other potential manifestations of SLE.

The mean length of follow-up of >3 years, while not complete for all patients, is likely to allow adequate time for the development of subsequent underlying SLE to become apparent (4). Patients in our study often had a long-standing psychotic disorder, further extending the period for development of other SLE manifestations. Depending on the patient's ethnicity, NPSLE is usually associated with other clinical features of SLE that would become clinically apparent given the long period of data collection and follow-up (3,4).

Performing this study at 2 sites provides validity to its findings. The total number of patients with psychosis was similar at both sites. The rate of ANA testing was similar (18% and 20%), as were the characteristics of tested patients and the rate of

ANA positivity. Taken together, this suggests that the results may be more broadly applicable.

Patients tested for ANAs had more prolonged admissions than their peers at both centers. This suggests that those with recalcitrant mental health issues were more likely to have further testing. Screening for organic causes of psychosis was the most common reason given for testing ANAs. Female patients were tested more commonly at both sites, possibly due to the expected female preponderance of SLE.

Patients who were ANA positive were infrequently referred for rheumatology or clinical immunology review. The reason for this is not clear but perhaps relates to the frequently low ANA titer and lack of other clinical features found in these patients.

This study did not assess for the prevalence of ANA positivity in patients with psychosis; a prospective cohort has shown no significant difference between patients with schizophrenia and a group of healthy controls (14). There are potential confounding factors for the higher prevalence of ANA positivity in our study, the most important being selection bias by the treating clinicians. Hepatitis C infection was a common association in the patients who were ANA positive. There is an established link between hepatitis C infection and ANA positivity that may account for this finding (15). As highlighted in other prevalence studies, antipsychotics such as chlorpromazine may be another confounder for positive ANA (14).

The present study builds on previous studies that have been conflicting in their findings regarding the outcome of ANA testing in patients admitted with mental health disorders (6,7). The cohort in our study of ANA-positive patients with psychosis is significantly larger than the cohorts in previous studies and suggests that the rate of NPSLE is very low.

The main limitation of the study was a retrospective rather than prospective design. Testing of ANAs in the present study was not uniform and was performed at the discretion of the treating psychiatrist. As a result, it did not assess the prevalence of NSPLE in a general hospital inpatient psychosis population but rather the prevalence in those selected to undergo ANA testing. While this may limit generalizability, we believe it does not detract from the key finding of the study, namely that the rate of NPSLE is very low in acute hospital inpatient psychosis patients with ANA positivity. Indeed, the ANA sampling bias in this cohort may have overestimated the rate of NPSLE in ANA-positive psychosis patients, as psychiatrists generally ordered the ANA testing when they had suspicion of organic disease.

Other limitations of the present study include a lack of patient follow-up by a specialist experienced in the diagnosis of NPSLE and incomplete follow-up data, with 16% of ANA-positive patients lost to follow-up. In a low prevalence condition, this increases the risk that the data exclude important results that may significantly change the outcome.

In conclusion, in a preselected group of hospital inpatients with psychosis, the likelihood of ANA positivity being associated with NPSLE was low. Only 2 patients had a diagnosis of NPSLE, which represented 1.5% of those who were ANA positive and 0.2% of those undergoing ANA testing. Both of these patients had extra-psychiatric clinical features of SLE. The findings of this study argue against routine screening of ANAs in hospital inpatients with psychosis.

ACKNOWLEDGMENTS

The authors thank the Departments of Psychiatry at Prince of Wales Hospital and at the Royal Prince Alfred Hospital. The authors also thank Gordana Popovic, statistical consultant at the University of New South Wales, and Sean O'Neil, associate professor at Royal North Shore Hospital.

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. Spies 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. Spies, Bertouch, Sammel.
Acquisition of data. Spies, Gutjahr-Holland.
Analysis and interpretation of data. Spies, Gutjahr-Holland, Bertouch, Sammel.

REFERENCES

1. Liang MH, Corzillius M, Bae SC, Lew RA, Fortin PR, Gordon C, et al. The American College of Rheumatology nomenclature and case definitions for neuropsychiatric lupus syndromes. Arthritis Rheum 1999; 42:599–608.
2. Unterman A, Nolte JE, Boaz M, Abady M, Shoenfeld Y, Zandman-Goddard G. Neuropsychiatric syndromes in systemic lupus erythematosus: a meta-analysis. Semin Arthritis Rheum 2011;41:1–11.
3. Pego-Reigosa JM, Isenberg DA. Psychosis due to systemic lupus erythematosus: characteristics and long-term outcome of this rare manifestation of the disease. Rheumatology (Oxford) 2008;47:1498–502.
4. Hanly JG, Li Q, Su L, Urowitz MB, Gordon C, Bae SC, et al. Psychosis in systemic lupus erythematosus. Arthritis Rheum 2019;71:281–9.
5. Lehman AF, Lieberman JA, Dixon LB, McGlashan TH, Miller AL, Perkins DO, et al. Practice guideline for the treatment of patients with schizophrenia. Am J Psychiatry 2004;161 Suppl:1–56.
6. Mantovani C, Louzada-Junior P, Nunes EA, de Figueiredo FP, Oliveira GR, Del-Ben CM. Antinuclear antibodies testing as a routine screening for systemic lupus erythematosus in patients presenting first-episode psychosis. Early Interv Psychiatry 2012;6:322–5.
7. Audemard-Verger A, Comby E, Nathou C, Sultan A, Frémont M, Baldolli A, et al. Is it relevant to screen young women hospitalized in psychiatric department for neuropsychiatric systemic lupus erythematosus (NPSLE)? A prospective study of 100 psychiatric inpatients. Medicine (Baltimore) 2016;95:e5288.
8. International Statistical Classification of Diseases and Related Health Problems, tenth revision. Geneva: World Health Organization; 1994.
9. Agmon-Levin N, Damoiseaux J, Kallenberg C, Sack U, Witte T, Herold M, et al. International recommendations for the assessment of autoantibodies to cellular antigens referred to as anti-nuclear antibodies. Ann Rheum Dis 2014;73:17–23.
10. Aringer M, Costenbader K, Daikh D, Brinks R, Mosca M, Ramsey-Goldman R, et al. 2019 European League Against Rheumatism/American College of Rheumatology classification criteria for systemic lupus erythematosus. Arthritis Rheumatol 2019;71:1400–12.

11. Bortoluzzi A, Scire CA, Bombardieri S, Caniatti L, Conti F, De Vita S, et al. Development and validation of a new algorithm for attribution of neuropsychiatric events in systemic lupus erythematosus. Rheumatology (Oxford) 2015;54:891–8.
12. Gladman DD, Ibañez D, Urowitz MB. Systemic lupus erythematosus disease activity index 2000. J Rheumatol 2002;29:288–91.
13. Bortoluzzi A, Fanouriakis A, Appenzeller S, Costallat L, Scirè CA, Murphy E, et al. Validity of the Italian algorithm for the attribution of neuropsychiatric events in systemic lupus erythematosus: a retrospective multicentre international diagnostic cohort study. BMJ Open 2017;7:e015546.
14. Van Mierlo HC, De Witte L, Derksen RH, Otten HG. The prevalence of antinuclear antibodies in patients with schizophrenia spectrum disorders: results from a large cohort study. NPJ Schizophr 2015;1:15013.
15. Gilman AJ, Le AK, Zhao C, Hoang J, Yasukawa LA, Weber SC, et al. Autoantibodies in chronic hepatitis C virus infection: impact on clinical outcomes and extrahepatic manifestations. BMJ Open Gastroenterol 2018;5:e000203.

Arthritis Care & Research
Vol. 74, No. 3, March 2022, pp 433–441
DOI 10.1002/acr.24451

American College of Rheumatology
Empowering Rheumatology Professionals

Evolution of Systemic Sclerosis–Associated Interstitial Lung Disease One Year After Hematopoietic Stem Cell Transplantation or Cyclophosphamide

Jacopo Ciaffi,[1] Nina M. van Leeuwen,[2] Maaike Boonstra,[2] Lucia J. M. Kroft,[2] Anne A. Schouffoer,[3] Maarten K. Ninaber,[2] Tom W. J. Huizinga,[2] and Jeska K. de Vries-Bouwstra[2]

Objective. Hematopoietic stem cell transplantation (HSCT) and cyclophosphamide (CYC) are treatment options for progressive systemic sclerosis associated with interstitial lung disease (SSc-ILD). The aims of our retrospective observational study were to evaluate: 1) the evolution of SSc-ILD in SSc patients treated with HSCT (assessed by high-resolution computed tomography [HRCT]; a group of patients treated with CYC was included as frame of reference); 2) how results of pulmonary function tests (PFTs) are associated with HRCT findings; and 3) which factors predict ILD reduction.

Methods. We semiquantitatively scored total ILD extent, reticulations, and ground-glass opacities (GGO) scores at baseline and at the 1-year HRCTs of SSc patients treated with HSCT or CYC. Linear association between changes in HRCT scores and PFT results and predictors of ILD improvement were studied.

Results. We included 51 patients (those treated with HSCT [n = 20] and those treated with CYC [n = 31]). The mean change in total ILD score was –5.1% (95% confidence interval [95% CI] –10.2, 0.0) in the HSCT treatment group ($P = 0.050$), and –1.0% (95% CI –4.3, 2.3) in the CYC treatment group ($P = 0.535$). For all patients, the evolution of HRCT scores was weakly associated with relative changes in PFT results. In univariate logistic regression, higher ground-glass opacities, higher total ILD, and lower single-breath diffusing capacity for carbon monoxide scores at baseline predicted improvement of ILD extent after treatment, but a multivariable model could not be built to assess independency of predictors.

Conclusion. One year after treatment with HSCT, a nonsignificant but clear reduction of SSc-ILD extent was observed. Changes in PFT results were associated with changes in HRCT scores but the correlation was weak and cannot be considered conclusive.

INTRODUCTION

Systemic sclerosis (SSc) is a complex connective tissue disease characterized by autoimmunity, vasculopathy, and fibrosis of skin and internal organs (1). Prevalence of SSc associated with interstitial lung disease (SSc-ILD) is estimated to be between 35% and 52% of SSc patients (2), but presence of lung abnormalities assessed by high-resolution computed tomography (HRCT) of the thorax has been described in up to 80% of SSc patients (3).

Complementarily to pulmonary function tests (PFTs), HRCT is the recommended screening tool to detect ILD in patients with SSc (4). With HRCT, ground-glass opacities (GGO), reticulations, subpleural nodules, and honeycombing, which are the characteristic features of ILD, can be recognized. It is also possible to identify the ILD pattern that, in SSc, is predominantly nonspecific interstitial pneumonia (5,6). Moreover, distinguishing patients with or without extensive lung involvement has significant treatment implications, underlining the prognostic relevance of HRCT (7).

In general, there are 2 treatment regimens most frequently adopted for SSc-ILD. Based on the results of Scleroderma Lung Study I (SLS-I) and II (SLS-II) (8,9), an induction phase with cyclophosphamide (CYC) that is followed or not followed by maintenance therapy with mycophenolate mofetil (MMF) or azathioprine (10), or, alternatively, MMF alone, are often chosen. The role of

[1]Jacopo Ciaffi, MD: Leiden University Medical Center, Leiden, The Netherlands, Azienda Policlinico of Modena, University of Modena, and Reggio Emilia, Modena, Italy, and IRCCS Istituto Ortopedico Rizzoli, Bologna, Italy; [2]Nina M. van Leeuwen, MD, Maaike Boonstra, MD, Lucia J. M. Kroft, MD, PhD, Maarten K. Ninaber, MD, Tom W. J. Huizinga, MD, PhD, Jeska K. de Vries-Bouwstra, MD, PhD: Leiden University Medical Center, Leiden, The Netherlands; [3]Anne A. Schouffoer, MD, PhD: Leiden University Medical Center, Leiden, The Netherlands and Haga Teaching Hospital, the Hague, The Netherlands.

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

Address correspondence to Jacopo Ciaffi, MD, Leids Universitair Medisch Centrum, Stafcentrum Reumatologie, C1-R, Postbus 9600, 2300 RC, Leiden, The Netherlands. Email: j.ciaffi@lumc.nl.

Submitted for publication January 29, 2020; accepted in revised form September 8, 2020.

SIGNIFICANCE & INNOVATIONS

- Analyzing patients with systemic sclerosis (SSc) who were followed in our cohort through serial high-resolution computed tomography (HRCTs), we observed a clear reduction in the extent of interstitial lung disease (ILD) 1 year after hematopoietic stem cell transplantation (HSCT).
- Literature on HSCT in SSc mainly derives from randomized controlled trials. Our work contributes real-world data about the evolution of SSc-associated ILD after HSCT.
- Our findings suggest a weak correlation between changes in HRCT and modifications of pulmonary function tests.

biologic therapies, such as tocilizumab (11) or rituximab (12), is not yet supported by strong evidence, and the advent of antifibrotic agents promises to enrich the narrow armamentarium for SSc-ILD (13). According to the latest update of the European Alliance of Associations for Rheumatology (EULAR) recommendations for the treatment of SSc (14) and after 3 positive randomized controlled trials (RCTs) (15–17), hematopoietic stem cell transplantation (HSCT) is the other option to be considered for SSc-ILD in selected patients with rapidly progressive disease who are at risk of organ failure (after thorough screening).

The effect of CYC on the evolution of ILD assessed through HRCT (HRCT-ILD) in SSc patients has been evaluated in 2 analyses of the SLS trials (18,19), while the potential of HSCT to modify the progression of ILD extent has been investigated in 2 studies involving a limited number of patients (15,20). The main purpose of our research is to contribute additional real-world data retrospectively describing the evolution of SSc-ILD in patients treated according to the latest EULAR recommendations (14). Specifically, we are interested in the course of ILD in patients treated with autologous HSCT. Patients who received conventional immunosuppressive therapy with CYC were included as frame of reference to evaluate whether the changes in ILD that were observed after HSCT are different from what can be expected with intravenous CYC. The secondary objectives were to investigate the strength of correlation between HRCT evolution and changes in PFT results and which patient-related factors can predict ILD reduction in response to immunosuppressive therapy.

PATIENTS AND METHODS

Patients. The studied population was composed of adult patients enrolled in the Leiden Combined Care in SSc (CCISS) cohort (21). Data were collected between 2004 and 2019. To be included in the study, all participants had to meet the following criteria: 1) fulfill the 2013 American College of Rheumatology/EULAR classification criteria for SSc (22); 2) have evidence of ILD on baseline HRCT; 3) have been treated either with autologous HSCT or with intravenous CYC for at least 6 months; and 4) have high-quality HRCT images at baseline and after treatment completion available for scoring.

Research on the CCISS cohort is approved by the Ethics Committee of Leiden University Medical Center (LUMC) (approval number P09.003), and all patients gave written informed consent. The study was conducted in accordance with the principles of the Declaration of Helsinki.

HRCT scoring. Baseline and follow-up HRCT scans were acquired in supine position and at maximal inspiration, using a standardized protocol at the radiology department of the LUMC. In all patients, HRCT scans were obtained contiguously with a slice thickness of 1–2 mm. Baseline HRCTs performed less than 6 months before the start of therapy and follow-up HRCTs obtained from 6 to 18 months after HSCT procedure or after first CYC pulse were selected for scoring. These HRCTs were performed as standard annual evaluation in the context of the dedicated SSc care pathway (21).

Two investigators (1 thoracic radiologist [LJMK] and 1 rheumatologist [AAS]), who are experienced in evaluation of chest imaging in patients with connective tissue diseases, independently scored all HRCTs blinded for patients' clinical characteristics, treatment history, and pulmonary function. For each patient, baseline and follow-up HRCTs were directly compared. Scoring of total ILD extent was performed according to the simple visual semiquantitative system described by Goh et al (7). The following 5 levels were examined: 1) origin of great vessels, 2) main carina, 3) pulmonary venous confluence, 4) halfway between the third and fifth section, and 5) immediately above the right hemi-diaphragm. In each of the 5 sections, total ILD extent was estimated at the nearest 5% approximation. The method proposed by Goh et al (7) would then provide separate estimations of GGO and reticular pattern (RET) extents, multiplying the relative proportions of GGO and RET by the total ILD extent in each level. In order to be more adherent to what is observed in clinical practice, we decided not to evaluate GGO and RET in the 5 levels as relative proportions but rather as extents at the nearest 5% approximation. Considering that the GGO and RET may overlap in the same portion of lung parenchyma, a single area could be scored as involved by both. As a consequence, the ILD extent would not necessarily correspond to the mere summation of the 2 components. Global scores for each of the 3 variables were calculated as the mean of the scores obtained at the 5 levels. The mean of the 2 global scores of each variable, computed by the 2 readers (LJMK and AAS), resulted in the total scores, namely GGO, RET, and total ILD scores. Discrepancies above 10% in any of the variables' global scores were discussed between the 2 readers to reach consensus. Changes at follow-up were computed as absolute percentages compared to baseline scores.

A study by Goldin et al (19) demonstrated that decreases or increases of at least 4% in quantitative ILD scores in the lobe of

maximal involvement and of 2% or more in the whole lung were considered to respectively identify significant worsening or improvement in ILD extent in SSc patients, while scores that remained within these limits represented stable disease; however, those thresholds were based on quantitative HRCT texture-based computer-aided analysis. By contrast, the scoring method proposed by Goh et al (7) is visual and, to our knowledge, no study has validated its longitudinal application. Therefore, we collegially agreed that a difference of 5% between pre- and posttreatment HRCTs, corresponding to the approximated minimum change visually identifiable at each HRCT level, could be reliable to distinguish patients with different ILD evolution. Improvement or progression were thus defined, respectively, as absolute reduction or increase above 5% in total ILD score, while stability was defined as those patients in whom no more than 5% ILD extent modification was identified at follow-up.

Finally, since the differential diagnosis of GGO is wide-ranging, other possible causes of ILD were always considered for all patients with SSc. In Leiden, all SSc patients are assessed for pulmonary arterial hypertension (PAH) at least annually. The screening is based on the European Society of Cardiology/European Respiratory Society recommendations (23), and, since 2013, the DETECT algorithm (24) is systematically applied to all patients in order to identify individuals that should proceed to right-sided heart catheterization. Moreover, since opportunistic and nonopportunistic infections can cause GGO, all HRCTs were evaluated with a multidisciplinary approach involving thoracic radiologists, pulmonologists, rheumatologists, and, in case of suspicion of infectious processes, also infectious disease specialists. Bronchoalveolar lavage, however, was performed only in patients with symptoms suggestive of respiratory tract infection. Although lung biopsy would provide a definitive diagnosis, the procedure is not routinely performed in SSc patients at the LUMC.

Clinical and laboratory data. Demographic data and clinical characteristics were collected at baseline, including disease subset, autoantibody positivity, presence of PAH, and mean modified Rodnan skin thickness score. Screening for PAH and HRCTs were preformed contextually as part of the 2-day care pathway for the follow-up of SSc patients. Disease duration was defined as time since the onset of the first sign or symptom attributable to SSc that was different from Raynaud's phenomenon. Results of baseline and follow-up PFTs that were closest to the corresponding HRCT dates were collected, provided that the time interval between HRCTs and PFTs did not exceed 3 months. All measurements were obtained at the pulmonology department of the LUMC. Forced vital capacity (FVC), forced expiratory volume in 1 second (FEV_1), and single-breath diffusing capacity for carbon monoxide (DLco) corrected for hemoglobin were included in the study. All physiologic values were reported as percentages of the predicted reference values, in accordance with published standards (25,26). Patients receiving 6 or 12 CYC pulses were analyzed as a single group because they were comparable in terms of demographic characteristics and HRCT scores at baseline, and at follow-up there was no statistically significant difference in mean changes of HRCT scores and PFT results.

Table 1. Characteristics of patients*

Baseline characteristics	HSCT (n = 20)	CYC (n = 31)	*P*
Female sex, no. (%)	10 (50)	24 (77)	0.043†
Age, years	46.5 ± 10.1	51 ± 12.8	0.189
Disease duration, median (IQR) years	2.5 (1.2–5.4)	1.8 (0.7–4.4)	0.380
dcSSc, no. (%)	18 (90)	19 (61)	0.025†
ATA, no. (%)	14 (70)	17 (55)	0.279
ILD, no. (%)	20 (100)	31 (100)	1
PAH, no. (%)	0	1 (3)	0.417
MRSS	23.2 ± 12.2	13.5 ± 10.4	0.003†
HRCT scores, %			
Ground-glass score	20.3 ± 13.6	21.5 ± 12.7	0.737
Reticular pattern score	13.9 ± 10.8	14.8 ± 10.3	0.772
Total ILD score	26.8 ± 14.6	25.1 ± 13.9	0.679
Pulmonary function tests, % predicted ‡			
FVC	77.8 ± 18.1	80.2 ± 16.7	0.633
FEV_1	75.7 ± 15.3	82.1 ± 15.8	0.165
DLco	53.4 ± 19.2	53.3 ± 11.6	0.985

* Values are the mean ± SD unless indicated otherwise. ATA = anti–topoisomerase I antibodies; CYC = cyclophosphamide; dcSSc = diffuse cutaneous systemic sclerosis; DLco = single-breath diffusing capacity for carbon monoxide; FEV_1 = forced expiratory volume in 1 second; FVC = forced vital capacity; HRCT = high-resolution computed tomography; HSCT = hematopoietic stem cell transplantation; ILD = interstitial lung disease; IQR = interquartile range; MRSS = modified Rodnan skin thickness score; PAH = pulmonary arterial hypertension.
† Significant.
‡ In 2 patients, baseline FVC and FEV_1 (n = 1 in CYC group) or DLco (n = 1 in CYC group) were not available, while at follow-up FVC and FEV_1 had not been obtained in 1 patient in the CYC group.

Table 2. Difference of pretreatment and 12-month posttreatment HRCT scores and PFTs within 2 groups*

	Changes within groups			
	Pretreatment	Posttreatment	Difference (95% CI), %†	*P*
HSCT				
HRCT scores, %				
Ground glass opacities score	20.3 ± 13.6	14.1 ± 8.7	−6.2 (−11.0, −1.4)	0.015‡
Reticular pattern score	13.9 ± 10.8	13.5 ± 10.3	−0.4 (−1.7, 0.9)	0.542
Total ILD score	26.8 ± 14.6	21.7 ± 11.8	−5.1 (−10.2, 0.0)	0.050
PFTs, %				
FVC	77.8 ± 18.1	84.7 ± 19.2	+6.9 (3.5, 10.4)	<0.001‡
FEV_1	75.7 ± 15.3	82.2 ± 15.9	+6.5 (2.6, 10.4)	0.002‡
DL_{CO}	53.4 ± 19.2	55.1 ± 15.0	+1.7 (−2.8, 6.3)	0.431
CYC				
HRCT scores, %				
Ground glass opacities score	21.5 ± 12.7	19.8 ± 12.3	−1.7 (−5.0, 1.6)	0.301
Reticular pattern score	14.8 ± 10.3	16.2 ± 10.9	+1.4 (0.0, 2.8)	0.053
Total ILD score	25.1 ± 13.9	24.1 ± 13.8	−1 (−4.3, 2.3)	0.535
PFTs, %				
FVC	80.2 ± 16.7	84.6 ± 19.7	+4.4 (−0.5, 9.4)	0.077
FEV_1	82.1 ± 15.8	86.5 ± 17.6	+4.4 (−0.2, 9.0)	0.061
DL_{CO}	53.3 ± 11.6	55.6 ± 13.1	+2.3 (−1.4, 6.0)	0.209

* Values are the mean ± SD unless indicated otherwise. 95% CI = 95% confidence interval; PFTs = pulmonary function tests (see Table 1 for other definitions).
The posttreatment period was 12 months after HSCT or CYC treatment initiation.
† The reported differences are absolute percentages.
‡ Significant.

Minimum clinically important difference (MCID) for FVC changes in SSc-ILD has been estimated using data from the SLS trials. It corresponds to a range from 3.0% to 5.3% for improvement and from –3.0% to –3.3% for worsening (27). However, this MCID was obtained at group level in the SLS trials, whereas in the present study we primarily aimed at assessing changes in individual patients. Therefore, we applied a more stringent 5% cutoff also to FVC changes, in order to identify patients with improved, stable, or worsened ventilatory function.

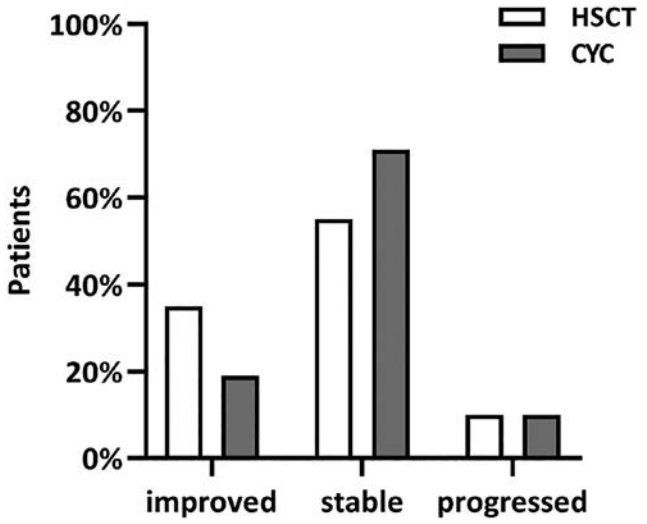


Figure 1. Proportion of improved, stable, or progressed patients 1 year after treatment with hematopoietic stem cell transplantation (HSCT) or cyclophosphamide (CYC). Improvement or progression were defined as absolute changes >5% in total interstitial lung disease (ILD) score, while stability identified patients with no more than 5% ILD extent modification at follow-up.

The stratification method described by Goh et al (7) was applied to define, at baseline, patients with extensive or limited lung disease. In particular, patients with an ILD extent of ≤10% were categorized as having limited disease and patients with >30% as having extensive disease. For cases with an ILD extent of >10% and ≤30%, an FVC of ≥70% or <70% stratified the patient, respectively, in the limited or extensive disease group.

Statistical analysis. Statistical analysis was performed using SPSS, version 23. Baseline demographic, clinical, and laboratory characteristics were expressed using descriptive statistics, with mean ± SD or median (interquartile range [IQR]) reported when appropriate. Differences in baseline characteristics between patients treated with CYC or HSCT were analyzed. These characteristics were compared using the 2-sample *t*-test and Mann–Whitney U test, respectively, for normally and non-normally distributed continuous variables, and with the chi-square test for categorical variables. Pre- to posttreatment changes in HRCT scores and PFT results were compared within groups using paired sample *t*-tests, and differences in means were reported with 95% confidence intervals (95% CIs). In the combined population, Spearman's rho (ρ) was used to evaluate the correlation between absolute changes in HRCT scores and relative changes in PFT results, whereas univariate logistic regression was used to assess baseline characteristics predictive of improvement in ILD extent. *P* values less than 0.05 were considered statistically significant.

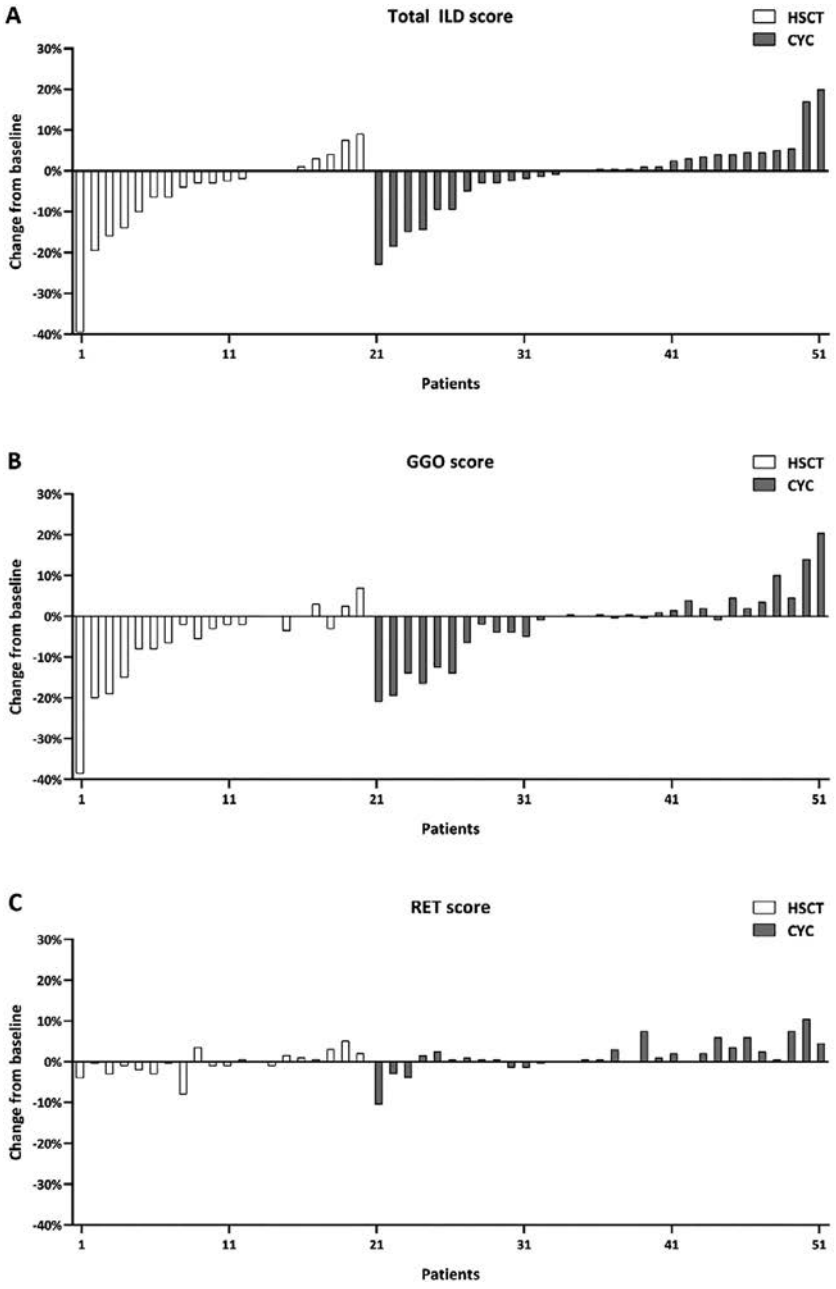


Figure 2. Evolution of high-resolution computed tomography scores 1 year after treatment shown in descending order from maximum improvement to maximum progression of interstitial lung disease (ILD) score. Individual changes from baseline to follow-up of total (ILD) score (**A**), ground-glass opacities (GGOs) score (**B**), and reticular pattern (RET) score (**C**) are shown. Bars represent individual patients. HSCT = hematopoietic stem cell transplantation; CYC = intravenous cyclophosphamide.

RESULTS

Patient characteristics. In order to have the 1-year follow-up HRCTs available, we included patients treated until May 2018. Of the patients enrolled in the CCISS cohort, 40 received HSCT. Of these patients who received HSCT, 20 were not included in the present study. The main reasons for exclusion were absence of ILD at baseline HRCT (n = 7) and unavailability of HRCT images stored in digital format and suitable for scoring (n = 9). Moreover, 3 patients died before the follow-up HRCT could be obtained (treatment-related complications, including sepsis and multiple organ failure [n = 2] and disease progression [n = 1]), and in 1 case scoring accuracy was compromised by the presence of radiation-induced pulmonary fibrosis. As a result, 20 patients (10 men, 10 women) treated with HSCT were included in the study. As frame of reference, 31 patients (7 men, 24 women) who were treated with 6 monthly pulses (n = 17) or 12 monthly pulses (n = 14) of CYC (750 mg/m^2) were studied. Baseline characteristics are

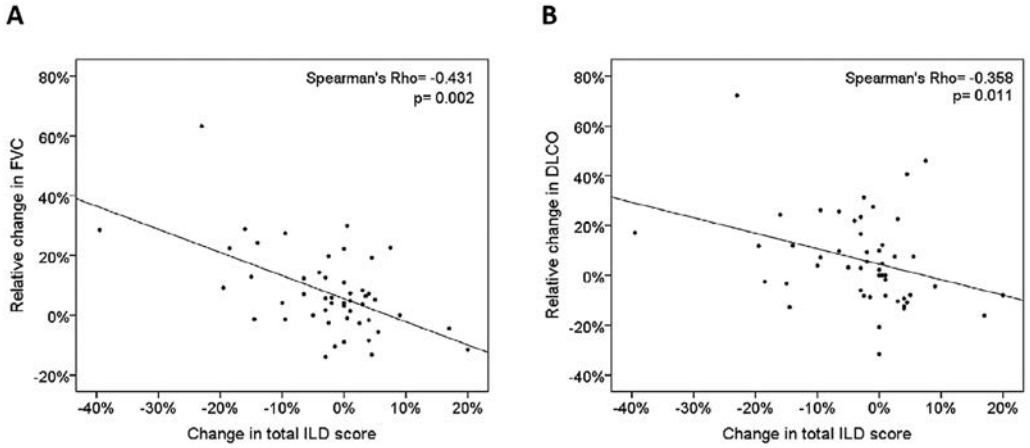


Figure 3. Correlations between evolution of interstitial lung disease (ILD) and changes in pulmonary function tests. Scatter plots of bivariate correlation, in the combined population, between changes in total ILD and relative changes in forced vital capacity (FVC) (**A**) or single-breath diffusing capacity for carbon monoxide (DLCO) (**B**). At follow-up compared with baseline, changes in total ILD score were calculated as absolute differences, and changes in FVC and DLCO were calculated as relative changes.

presented in Table 1. In the population, the mean age at treatment was 49.2 ± 11.9 years and median disease duration was 1.9 years (IQR 0.9–4.4 years). In the HSCT group, there were more men ($P = 0.043$), more patients with diffuse cutaneous SSc ($P = 0.025$), and higher mean mRSS ($P = 0.003$). Thirteen of the 20 patients treated with HSCT had received ≥1 immunosuppressive therapies before HSCT (CYC [n = 7], methotrexate [n = 5], and glucocorticoids [n = 6]). In the CYC group, 18 patients had been treated with at least 1 immunosuppressant before CYC was administered (methotrexate [n = 10], MMF [n = 7], glucocorticoids [n = 12], azathioprine [n = 4], and rituximab [n = 2]). Although it cannot be excluded that these therapies had a long-term effect on ILD, all had already been discontinued when baseline HRCTs included in the present study were obtained and all patients had progressive ILD.

Baseline HRCTs and PFTs. In the whole population, at baseline, the mean total ILD score was 25.8% ± 14.0%, the mean GGO score was 21.0% ± 12.9% and the mean RET score was 14.4% ± 10.4%. According to the staging system proposed by Goh et al (7), which includes combining HRCT scores and PFT results, 12 patients in the HSCT group (60%) and 12 in the CYC group (39%) were classified as having extensive lung disease while, respectively, 8 (40%) and 18 (58%) had limited disease. In 1 patient in the CYC group, pretreatment FVC was missing, and the staging system could not be applied. In the studied population, the mean baseline FVC and DLCO were, respectively, 79.2% ± 17.2% and 53.3% ± 14.9%.

The HRCT scores and the results of PFTs at baseline in patients treated with HSCT or CYC were numerically comparable. No statistically significant difference was detected between the 2 groups (Table 1).

HRCT and PFT changes at follow-up. Follow-up HRCTs were obtained 11.4 ± 3 months after treatment initiation. In the HSCT group, the mean changes in HRCT scores were –5.1% (95% CI –10.2, 0.0; $P = 0.050$) for total ILD, –6.2% (95% CI –11.0, –1.4; $P = 0.015$) for GGO, and –0.4% (95% CI –1.7, 0.9; $P = 0.542$) for RET (Table 2). In the CYC group, mean changes in HRCT scores were –1.0 % (95% CI –4.3, 2.3; $P = 0.535$) for total ILD, –1.7% (95% CI –5.0, 1.6; $P = 0.301$) for GGO, and +1.4% (95% CI 0.0, 2.8; $P = 0.053$) for RET (Table 2). Using the defined cutoff of >5% change in HRCT score to define ILD extent improvement, stability, or progression, we found that 35% (n = 7) of patients in the HSCT group and 19% (n = 6) in the CYC group were categorized as improved, while ILD extent remained stable respectively in 55% of patients in the HSCT group (n = 11) and 71% (n = 22) in the CYC group, and showed progression in 10% of patients in both groups (n = 3 in CYC and n = 2 in HSCT) (Figure 1). Details regarding changes in each HRCT score at individual patient level are provided in Figure 2. The HRCT scan of a patient with marked ILD improvement after HSCT is shown (see Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24451/abstract).

One year after HSCT, the mean FVC increased by 6.9% (95% CI 3.5, 10.4; $P < 0.001$) and the mean DLCO by 1.7% (95% CI –2.8, 6.3; $P = 0.431$). In the CYC group, the mean FVC increased by 4.4% (95% CI –0.5, 9.4; $P = 0.077$) and the mean DLCO by 2.3% (95% CI –1.4, 6.0; $P = 0.209$). Data are presented in Table 2.

Applying the defined cutoff of 5% change in FVC% to define improvement, stability, or progression of ventilatory function, we observed that 55% (n = 11) of patients in the HSCT and 36% (n = 11) in the CYC group experienced FVC% improvement after treatment, respectively 40% (n = 8) and 42% (n = 13) were

Table 3. Predictors of improvement in ILD extension at follow-up HRCT*

Variables	ILD improvement at follow-up HRCT	
	OR (95% CI)	*P*
Baseline GGO score	1.20 (1.08, 1.34)	0.001†
Baseline reticular pattern score	1.02 (0.95, 1.08)	0.662
Baseline total ILD score	1.12 (1.04, 1.21)	0.003†
Baseline FVC	0.96 (0.92, 1.01)	0.092
Baseline FEV_1	0.97 (0.93, 1.01)	0.150
Baseline DLco	0.91 (0.84, 0.97)	0.008†
Disease duration	0.76 (0.54, 1.07)	0.113
Age	1.02 (0.97, 1.08)	0.453
DcSSc	1.24 (0.31, 4.95)	0.756
Baseline mRSS	0.98 (0.93, 1.04)	0.540
ATA positivity	1.64 (0.43, 6.26)	0.472
Female sex	1.94 (0.46, 8.28)	0.368

* Univariate logistic regression in the combined population analyzing patients with >5% reduction in interstitial lung disease (ILD) assessed through high-resolution computed tomography (HRCT) extension (13 of 51 patients). 95% CI = 95% confidence interval; ATA = anti–topoisomerase I antibody; dcSSc = diffuse cutaneous systemic sclerosis; DLco = single-breath diffusing capacity for carbon monoxide; FEV_1 = forced expiratory volume in 1 second; FVC = forced vital capacity; GGO = ground-glass opacity; OR = odds ratio.
† Significant.

stable, while in 5% (n = 1) and 16% (n = 5) of cases FVC% worsened. Change in FVC% could not be calculated in 2 CYC patients due to missing data.

Correlation of HRCT and PFTs changes. In a pooled analysis of all patients, relative change in FVC was correlated with changes in the mean total ILD score ($\rho = -0.431$, $P = 0.002$), in the GGO score ($\rho = -0.354$, $P = 0.013$), and in the RET score ($\rho = -0.424$, $P = 0.002$). Also, relative changes in DLco were associated with changes in the mean total ILD score ($\rho = -0.358$, $P = 0.011$) and in the RET score ($\rho = -0.368$, $P = 0.009$), but not with the GGO score ($\rho = -0.266$, $P = 0.062$). Correlations between the evolution of ILD and changes in FVC or DLco are shown in Figure 3.

Predictors of improvement in the combined population. As predefined, 13 patients (n = 7 in HSCT and n = 6 in CYC group) were categorized as improved in the combined population. Baseline characteristics of improvers and nonimprovers are shown (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24451/abstract). Using HRCT-ILD improvement as the dependent variable, we assessed baseline characteristics that were predictive of a more favorable response to immunosuppression. In univariate logistic regression (Table 3), the baseline GGO score (OR 1.20 [95% CI 1.08, 1.34], $P = 0.001$), total ILD score (OR 1.12 [95% CI 1.04, 1.21], $P = 0.003$), and pretreatment DLco (OR 0.91 [95% CI 0.84, 0.97], $P = 0.008$) predicted improvement in total ILD score at follow-up. As only 13 patients were categorized as improvers (and also taking into consideration the multicollinearity between predictors), a multivariable logistic regression model could not be built, thus precluding the possibility to analyze the independency of the improvement predictors.

DISCUSSION

We performed a retrospective observational study to describe the evolution of ILD for 1 year in SSc patients receiving HSCT. A group of patients treated with intravenous CYC was included as frame of reference. The main finding of our study is that the mean ILD extent, which was evaluated through HRCTs of the thorax, decreased in both groups. In particular, the mean ILD extent decreased by 5.1% in the HSCT group and by 1.0% in the CYC group. Secondly, we showed an association between modifications of HRCT-ILD and changes in PFT results. However, although statistically significant, the correlations between the evolution of ILD and variations of PFT results, and in particular of DLco, were weak in all analyses. Thirdly, studying which baseline factors might influence reduction of ILD extent in response to immunosuppressive therapy, the independency of improvement predictors could not be assessed and results cannot be considered conclusive.

Our findings indicate that, when SSc patients are treated in accordance with current recommendations (14), stabilization or even improvement of ILD is a reachable goal in the majority of cases. Only 10% of patients showed progression of HRCT-ILD, and, in our opinion, the mean reduction in ILD extent observed after HSCT is indicative of a clear effect. Conversely, the minimal change in total ILD extent observed after CYC can conceivably be explained by measurement variability. Comparing the findings in the HSCT group in our study with previous research, in a small RCT, Burt et al (15) demonstrated how diseased-lung volume assessed through volumetric chest CT was significantly decreased 1 year after HSCT in SSc patients. A pilot study by Launay et al (20) reported a rapid improvement of SSc-ILD 6 months after HSCT that was mainly driven by the reversion of GGO, but, in some patients, was transient over time when longer-term follow-up HRCTs were analyzed. Nonetheless, also the stabilization of ILD extent that was observed in patients treated with CYC is in line with the available literature evidence. Post hoc analyses of the SLS-I and SLS-II trials (18,19) described the effect of CYC in preventing progression of ILD extent. Moreover, both after HSCT and after CYC, an improvement of PFTs can be expected (16,28). Our findings of a potential efficacy of HSCT and also of a role of CYC in modifying the evolution of SSc-ILD are thus consistent with prior studies.

Combining data from the 2 treatment groups, we observed a significant correlation between modifications in HRCT scores and changes of FVC and DLco at follow-up. Although we could only analyze a limited number of patients, and the correlations that we found are weak, our results suggest that a relationship

between HRCT findings and functional impairments can be identified in SSc patients (29). The role of PFTs in screening, diagnosis, and severity assessment of SSc-ILD is well established (4). Additionally, our data are in line with the recent evidence-based guidance for the identification and management of SSc-ILD (4), delineating the relevance of using PFTs alongside HRCT to monitor progression of the disease.

As a secondary aim of our study, we analyzed which pretreatment disease-related or patient-related characteristics predicted better response to immunosuppressive therapy. Due to the limited number of patients included, our results cannot be considered conclusive but, in univariate analysis, patients with higher GGO score, higher total ILD score, and lower DLco at baseline were more likely to experience an improvement of HRCT-ILD 1 year after treatment, while no correlation with RET scores was shown. Interestingly, a retrospective analysis of the SLS-I trial (30) identified more severe reticular changes at baseline as a predictor of better response to CYC in terms of FVC improvement. However, it should be noted that this study (30) used a different HRCT scoring method and severity of GGO was apparently not considered in the regression model. In the current study, we included patients treated with HSCT, which were not present in the SLS-I trial. As a consequence, it is difficult to compare our results and the SLS-I trial with regard to the contribution of GGO and reticulation on ILD evolution. Moreover, we could not build a multivariable model and the independency of the predictors could not be investigated.

Our study is not without limitations. First, it has a retrospective observational design. The majority of patients were not allocated randomly to HSCT or CYC, and, within the CYC group, not all patients received the same number of pulses. However, this method was consistent with the purpose of our study. Literature on the role of HSCT in SSc-ILD is mostly derived from RCTs, with a few contributions from observational cohorts (20,31). We aimed to investigate the evolution of ILD in a real-world scenario, where SSc patients are treated on the basis of clinical decisions. As a result, patients in the HSCT group had more severe skin involvement but, despite this, our study was focused exclusively on lung disease and specifically in patients treated with HSCT, with the CYC group included as frame of reference. Moreover, all posttreatment investigations were performed according to standard procedures, independently of patients' symptoms. Secondly, we could only include 51 patients, and the limited population size eventually prevented the possibility to conduct in-depth analyses. Furthermore, the application of visual semiquantitative scoring methods might be questionable, but it is important to emphasize that all HRCT images were centrally acquired using the same protocol and were scored by 2 experienced investigators with consensus reached when needed. Hence, we believe that the evaluation of HRCTs was accurate. Finally, we decided to limit the study to the first year of follow-up, thus preventing the possibility to make considerations about longer-term effects of HSCT and CYC.

In conclusion, this study shows that, after 1 year, 90% of patients treated with HSCT or CYC present stable or reduced ILD extent, but the improvement in HRCT-ILD observed after HSCT was indicative of a favorable effect that did not emerge in the CYC group. Our study contributes real-world data from a considerable number of patients treated with HSCT, and screening optimization and further research might help to redefine the role of HSCT (which is currently considered a rescue therapy) as an effective option to prevent the progression of, and even potentially reverse, SSc-ILD.

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. Ciaffi 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. Ciaffi, van Leeuwen, Boonstra, Kroft, Schouffoer, Ninaber, Huizinga, de Vries-Bouwstra.

Acquisition of data. Ciaffi, van Leeuwen, Boonstra, Kroft, Schouffoer, Ninaber, Huizinga, de Vries-Bouwstra.

Analysis and interpretation of data. Ciaffi, van Leeuwen, Boonstra, Kroft, Schouffoer, Ninaber, Huizinga, de Vries-Bouwstra.

REFERENCES

1. Denton CP, Khanna D. Systemic sclerosis. Lancet 2017;390:1685–99.
2. Bergamasco A, Hartmann N, Wallace L, Verpillat P. Epidemiology of systemic sclerosis and systemic sclerosis-associated interstitial lung disease. Clin Epidemiol 2019;11:257–73.
3. Desai SR, Veeraraghavan S, Hansell DM, Nikolakopolou A, Goh NS, Nicholson AG, et al. CT features of lung disease in patients with systemic sclerosis: comparison with idiopathic pulmonary fibrosis and nonspecific interstitial pneumonia. Radiology 2004;232:560–7.
4. Hoffmann-Vold A-M, Maher TM, Philpot EE, Ashrafzadeh A, Barake R, Barsotti S, et al. The identification and management of interstitial lung disease in systemic sclerosis: evidence-based European consensus statements. Lancet Rheumatol 2020;2:e71–83.
5. Wells AU, Hansell DM, Corrin B, Harrison NK, Goldstraw P, Black CM, et al. High resolution computed tomography as a predictor of lung histology in systemic sclerosis. Thorax 1992;47:738–42.
6. Fischer A, Swigris JJ, Groshong SD, Cool CD, Sahin H, Lynch DA, et al. Clinically significant interstitial lung disease in limited scleroderma: histopathology, clinical features, and survival. Chest 2008;134:601–5.
7. Goh NS, Desai SR, Veeraraghavan S, Hansell DM, Copley SJ, Maher TM, et al. Interstitial lung disease in systemic sclerosis: a simple staging system. Am J Respir Crit Care Med 2008;177:1248–54.
8. Tashkin DP, Elashoff R, Clements PJ, Goldin J, Roth MD, Furst DE, et al. Cyclophosphamide versus placebo in scleroderma lung disease. N Engl J Med 2006;354:2655–66.
9. Tashkin DP, Roth MD, Clements PJ, Furst DE, Khanna D, Kleerup EC, et al. Mycophenolate mofetil versus oral cyclophosphamide in scleroderma-related interstitial lung disease (SLS ii): a randomised controlled, double-blind, parallel group trial. Lancet Respir Med 2016;4:708–19.
10. Nagaraja V, Denton CP, Khanna D. Old medications and new targeted therapies in systemic sclerosis. Rheumatology (Oxford) 2015;54:1944–53.
11. Khanna D, Denton CP, Lin CJ, van Laar JM, Frech TM, Anderson ME, et al. Safety and efficacy of subcutaneous tocilizumab in systemic

sclerosis: results from the open-label period of a phase II randomised controlled trial (faSScinate). Ann Rheum Dis 2018;77:212–20.

12. Boonstra M, Meijs J, Dorjee AL, Marsan NA, Schouffoer A, Ninaber MK, et al. Rituximab in early systemic sclerosis. RMD Open 2017;3:e000384.
13. Distler O, Highland KB, Gahlemann M, Azuma A, Fischer A, Mayes MD, et al. Nintedanib for systemic sclerosis-associated interstitial lung disease. N Engl J Med 2019;380:2518–28.
14. Kowal-Bielecka O, Fransen J, Avouac J, Becker M, Kulak A, Allanore Y, et al. Update of EULAR recommendations for the treatment of systemic sclerosis. Ann Rheum Dis 2017;76:1327–39.
15. Burt RK, Shah SJ, Dill K, Grant T, Gheorghiade M, Schroeder J, et al. Autologous non-myeloablative haemopoietic stem-cell transplantation compared with pulse cyclophosphamide once per month for systemic sclerosis (assist): an open-label, randomised phase 2 trial. Lancet 2011;378:498–506.
16. Van Laar JM, Farge D, Sont JK, Naraghi K, Marjanovic Z, Larghero J, et al. Autologous hematopoietic stem cell transplantation vs intravenous pulse cyclophosphamide in diffuse cutaneous systemic sclerosis: a randomized clinical trial. JAMA 2014;311:2490–8.
17. Sullivan KM, Goldmuntz EA, Keyes-Elstein L, McSweeney PA, Pinckney A, Welch B, et al. Myeloablative autologous stem-cell transplantation for severe scleroderma. N Engl J Med 2018;378:35–47.
18. Goldin J, Elashoff R, Kim HJ, Yan X, Lynch D, Strollo D, et al. Treatment of scleroderma-interstitial lung disease with cyclophosphamide is associated with less progressive fibrosis on serial thoracic high-resolution CT scan than placebo: findings from the Scleroderma Lung Study. Chest 2009;136:1333–40.
19. Goldin JG, Kim GH, Tseng CH, Volkmann E, Furst D, Clements P, et al. Longitudinal changes in quantitative interstitial lung disease on computed tomography after immunosuppression in the Scleroderma Lung Study II. Ann Am Thorac Soc 2018;15:1286–95.
20. Launay D, Marjanovic Z, de Bazelaire C, Florea L, Zohar S, Keshtmand H, et al. Autologous hematopoietic stem cell transplant in systemic sclerosis: quantitative high resolution computed tomography of the chest scoring. J Rheumatol 2009;36:1460–3.
21. Meijs J, Schouffoer AA, Ajmone Marsan N, Kroft LJ, Stijnen T, Ninaber MK, et al. Therapeutic and diagnostic outcomes of a standardised, comprehensive care pathway for patients with systemic sclerosis. RMD Open 2016;2:e000159.
22. Van den Hoogen F, Khanna D, Fransen J, Johnson SR, Baron M, Tyndall A, et al. 2013 classification criteria for systemic sclerosis: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 2013;65:2737–47.
23. Galiè N, Humbert M, Vachiery JL, Gibbs S, Lang I, Torbicki A, et al. 2015 ESC/ERS guidelines for the diagnosis and treatment of pulmonary hypertension: the joint task force for the diagnosis and treatment of pulmonary hypertension of the European Society of Cardiology (ESC) and the European Respiratory Society (ERS): endorsed by: Association for European Paediatric and Congenital Cardiology (aepc), International Society for Heart and Lung Transplantation (ISHLT). Eur Respir J 2015;46:903–75.
24. Coghlan JG, Denton CP, Grünig E, Bonderman D, Distler O, Khanna D, et al. Evidence-based detection of pulmonary arterial hypertension in systemic sclerosis: the DETECT study. Ann Rheum Dis 2014;73:1340–9.
25. Quanjer PH, Stanojevic S, Cole TJ, Baur X, Hall GL, Culver BH, et al. Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. Eur Respir J 2012;40:1324–43.
26. Laszlo G. European standards for lung function testing: 1993 update. Thorax 1993;48:873–6.
27. Kafaja S, Clements PJ, Wilhalme H, Tseng CH, Furst DE, Kim GH, et al. Reliability and minimal clinically important differences of forced vital capacity: results from the scleroderma lung studies (SLS-I and SLS-II). Am J Respir Crit Care Med 2018;197:644–52.
28. Van den Hombergh WM, Simons SO, Teesselink E, Knaapen-Hans HK, van den Hoogen FH, Fransen J, et al. Intravenous cyclophosphamide pulse therapy in interstitial lung disease associated with systemic sclerosis in a retrospective open-label study: influence of the extent of inflammation on pulmonary function. Clin Rheumatol 2018;37:2715–22.
29. Tashkin DP, Volkmann ER, Tseng CH, Kim HJ, Goldin J, Clements P, et al. Relationship between quantitative radiographic assessments of interstitial lung disease and physiological and clinical features of systemic sclerosis. Ann Rheum Dis 2016;75:374–81.
30. Roth MD, Tseng CH, Clements PJ, Furst DE, Tashkin DP, Goldin JG, et al. Predicting treatment outcomes and responder subsets in scleroderma-related interstitial lung disease. Arthritis Rheum 2011;63:2797–808.
31. Van Bijnen S, de Vries-Bouwstra J, van den Ende CH, Boonstra M, Kroft L, Geurts B, et al. Predictive factors for treatment-related mortality and major adverse events after autologous haematopoietic stem cell transplantation for systemic sclerosis: results of a long-term follow-up multicentre study. Ann Rheum Dis 2020;79:1084–9.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 3, March 2022, pp 442–450
DOI 10.1002/acr.24488

Relationship Between Gastrointestinal Transit, Medsger Gastrointestinal Severity, and University of California–Los Angeles Scleroderma Clinical Trial Consortium Gastrointestinal Tract 2.0 Symptoms in Patients With Systemic Sclerosis

Zsuzsanna H. McMahan,[1] Ana E. Tucker,[2] Jamie Perin,[3] Elizabeth R. Volkmann,[4] Subhash Kulkarni,[1] Harvey A. Ziessman,[1] Pankaj J. Pasricha,[1] and Fredrick M. Wigley[1]

Objective. Systemic sclerosis (SSc)–associated gastrointestinal (GI) complications are attributed to a variety of factors, including diet, microbiota dysbiosis, or GI transit abnormalities. Our objective was to examine the contribution of abnormal GI transit to SSc Medsger GI severity scores and/or University of California Los Angeles Scleroderma Clinical Trial Consortium Gastrointestinal Tract (UCLA GIT) 2.0 symptoms.

Methods. Patients with SSc and GI symptoms (n = 71) and healthy controls (n = 18) underwent whole gut transit (WGT) scintigraphy to assess transit from the esophagus to the colon. The presence of delayed transit and percent emptying in each GI region were measured. We compared the WGT measurements between categories of the Medsger GI severity score (0–4) and across UCLA GIT 2.0 domains and total score (0–3).

Results. A total of 88% of patients had >1 abnormal region of the gut on WGT scintigraphy. All patients requiring total parenteral nutrition had delayed small bowel transit, compared to only approximately 11% of patients in other Medsger GI severity groups ($P \leq 0.01$). Severe colonic transit delays were more likely in patients with Medsger GI scores of 3 (pseudo-obstruction and/or malabsorption) compared to other Medsger GI groups ($P = 0.02$). Seventy-percent of these patients had ≤30% colonic emptying at 72 hours. Modest associations were noted between gastroesophageal reflux disease symptoms and delayed esophageal ($r = –0.31$, $P = 0.05$) and gastric emptying ($r = –0.32$, $P = 0.05$).

Conclusion. These data are important in providing evidence that SSc bowel disease affects transit of GI content and that delay in transit accounts in part for both bowel symptoms and Medsger GI severity. Prospective studies examining the benefit of early therapeutic intervention targeting GI transit abnormalities in patients at high risk for severe GI complications are needed.

INTRODUCTION

Systemic sclerosis (SSc) is a complex disease characterized by autoimmunity, progressive vasculopathy, and excess deposition of collagen due to aberrant fibroblast function in the skin and internal organs (1,2). The gastrointestinal (GI) tract is the most commonly identified internal organ involved in SSc, with approximately 90% of patients affected (3). GI manifestations in patients with SSc are variable in terms of symptoms, complications, time course, and regions affected (4,5). A number of factors may contribute to GI symptoms and severity, including diet, microbiota dysbiosis, or abnormalities in GI transit (6,7). Pathologic findings and previous physiologic studies implicate bowel dysfunction leading to dysmotility, yet these studies have not clearly determined the clinical impact

Supported by the Martha McCrory Professorship and the Scleroderma Research Foundation, Chresanthe Stauraluakis Memorial Discovery Fund. Dr. McMahan's work was supported by the NIH/National Institute of Arthritis and Musculoskeletal and Skin Diseases (K23-AR-071473), the Scleroderma Research Foundation, and the Jerome L. Greene Foundation.

[1]Zsuzsanna H. McMahan, MD, MHS, Subhash Kulkarni, PhD, MS, Harvey A. Ziessman, MD, Pankaj J. Pasricha, MBBS, MD, Fredrick M. Wigley, MD: Johns Hopkins University, Baltimore, Maryland; [2]Ana E. Tucker, MD: Johns Hopkins Bayview Medical Center, Baltimore, Maryland; [3]Jamie Perin, PhD, MS: Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland; [4]Elizabeth R. Volkmann, MD, MS: University of California–Los Angeles.

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

Address correspondence to Zsuzsanna H. McMahan, MD, MHS, 5200 Eastern Avenue, Suite 5200, Mason F. Lord Building, Center Tower, Baltimore, MD 21224. Email: zmcmaha1@jhmi.edu.

Submitted for publication April 23, 2020; accepted in revised form October 13, 2020.

SIGNIFICANCE & INNOVATIONS

- Delayed gastrointestinal (GI) transit in specific regions of the gut plays a role in severe GI complications in patients with scleroderma.
- Severe GI phenotypes in systemic sclerosis associate with delayed transit in the small bowel and colon on whole gut transit studies.
- Several regions of the GI tract in patients with scleroderma may be affected by dysmotility simultaneously, the knowledge of which can impact our understanding and approach to targeted clinical therapies for scleroderma patients with GI disease.

of abnormalities in GI transit of food or content, particularly in the lower bowel (8–12). Understanding the relationship between GI severity, symptoms, and abnormal GI transit may allow for a more targeted approach in the management of such patients with regards to the selection and application of distinct therapies (13–16). For example, some medications, such as octreotide, primarily impact small bowel motility, whereas others, like prucalopride or linaclotide, have a more significant impact on large bowel motility (17–19).

Whole gut transit (WGT) scintigraphy is a tool used to objectively assess GI transit from the esophagus to the colon. It uses the passage of radioisotopes ingested as a solid and liquid meal through the gut to determine the extent and severity of the transit abnormalities (20). The results of WGT scintigraphy can help accurately define the regions of the gut affected by dysmotility as well as categorize transit severity (14).

We hypothesized that abnormal GI transit would associate with the severity of bowel dysfunction and specific GI clinical complaints. We used WGT studies in conjunction with the Medsger GI severity score and University of California Los Angeles Scleroderma Clinical Trial Consortium Gastrointestinal Tract (UCLA GIT) 2.0 instrument to evaluate these associations (21–23). Identifying such GI abnormalities that associate with poor outcomes would facilitate the application of targeted therapies and the study of earlier initiation of GI interventions in high-risk subgroups.

PATIENTS AND METHODS

Patients. All patients were from the Johns Hopkins Scleroderma Center and met the 2013 American College of Rheumatology/European League Against Rheumatism criteria for SSc (24). Patients were part of a prospectively enrolled GI cohort of patients evaluated in the Johns Hopkins Scleroderma Center (GI Assessment Protocol [GAP] cohort). Whole gut scintigraphy studies were obtained as part of clinical care in patients who had symptoms of significant upper GI disease or symptoms of both lower and upper GI dysfunction. At the clinical visit, significant symptoms of GI dysfunction were defined as early satiety, nausea/vomiting, unintentional weight loss, distension, bloating, diarrhea, and/or constipation as determined by the treating physician. To include patients from across the spectrum of GI disease, WGT studies on minimally symptomatic (e.g., mild heartburn alone) or asymptomatic SSc patients were obtained as part of a research protocol. All study patients were evaluated during their routine clinical visits at the Johns Hopkins Scleroderma Center. Written informed consent was obtained from all patients. The current study was approved by the Johns Hopkins Institutional Review Board.

Clinical phenotyping of the SSc patients. The Johns Hopkins Scleroderma Center database collects demographic and detailed clinical data from patients at their first clinical encounter and every 6 months at subsequent follow-up clinical visits. Age and disease duration were calculated from the date of birth and the date of the first SSc-associated symptom (Raynaud's or non–Raynaud's phenomenon) to the date of the WGT study, respectively. Patients were identified as having limited or diffuse SSc based on the extent of skin tightness (25). To define SSc phenotypes associated with specific GI dysmotility patterns and GI severity, the maximum clinical severity scores were used. The presence of a myopathy was denoted on the basis of an elevated creatine phosphokinase with evidence of electromyography supportive of myopathy, magnetic resonance imaging with evidence of muscle edema, or muscle biopsy consistent with myopathy (22). The muscle severity score was also used to classify the degree of associated proximal muscle weakness and was based on the following scale collected in our database: 0 = full strength, 1 = ability to lift upper or lower extremities against gravity with some resistance, 2 = ability to lift upper or lower extremities against gravity only, 3 = ability to move upper or lower extremities but not against gravity, and 4 = requiring ambulatory aids to walk (22,26).

Cardiac involvement was determined by the Medsger severity scale and was considered present with a score of 1 or greater (0 = normal, 1 = evidence of conduction defect on electrocardiogram or left ventricular ejection fraction [LVEF] of 45–49% on echocardiogram, 2 = evidence of arrhythmia on electrocardiogram or LVEF of 40–44%, 3 = clinical signs of left or right heart failure or arrhythmia requiring treatment with medication or intervention) (22,26). To capture the clinical phenotype, the minimum measurements from the forced vital capacity and single breath diffusing capacity of carbon monoxide pulmonary lung function testing and maximum measurements from the estimated right ventricular systolic pressure (measured by transthoracic echocardiogram) were used for the analysis (27). Sicca symptoms were defined as the presence of at least 1 of the following: dry eyes for >3 months, the use of artificial tears 3 times daily, dry mouth for >3 months, swollen salivary glands, the necessity of liquids for swallowing due to dry mouth, and/or the sensation of sand or gravel in the eyes (28). Evidence of patient-reported GI

symptoms was determined by the UCLA GIT 2.0 survey from the time closest to the WGT study.

Autoantibody profile. SSc autoantibodies (anti–Scl-70, anticentromere, anti–RNA polymerase III [anti-RNAP]) were determined for patients with available serum using a commercially available Euroline immunoblot assay (Scleroderma [Nucleoli] Profile Euroline IgG; Euroimmun). Moderate-to-high titers of autoantibodies, as determined by the manufacturer's cutoffs, were considered positive.

Control population. The control population of 18 patients with WGT studies was obtained from Johns Hopkins Nuclear Medicine. These individuals were recruited through in-house advertisements and were interviewed and screened with the aid of the Mayo Clinic Research Questionnaire. The accepted controls had no history of GI disorders or prior surgery, were not taking any medications, did not smoke or abuse alcohol (no more than 2–3 drinks per week) and were screened for GI disease through a standard questionnaire. Individuals without a history or symptoms of GI disease were enrolled in this cohort (29).

Instruments. *The UCLA GIT 2.0 instrument.* Each scale has a weighted subscore, with a 3-point categorical response (0–3) used to evaluate all items, excluding items 15 and 31 in the diarrhea and constipation categories, respectively, which rely on a score of 0 or 1.

Modified Medsger GI severity score. Physician-reported GI symptom severity was classified using the modified Medsger severity score (22). The score is composed of 5 categories that include: 1) score 0 = normal (no GI symptoms); 2) score 1 = requiring gastroesophageal reflux disease (GERD) medications, including an H2 blocker, proton pump inhibitor or prokinetic, or an abnormal bowel series; 3) score 2 = requiring high dose GERD medications (defined as greater than the lowest daily dose or a proton pump inhibitor plus a prokinetic drug) and/or having small bowel dilation on radiography; 4) score 3 = episodes of pseudo-obstruction or malabsorption syndrome; and 5) score 4 = severe GI dysmotility requiring either supplemental enteral or total parenteral nutrition (TPN).

WGT study. Upon study entry, WGT scintigraphy was obtained in all patients. Three days prior to the study, patients were instructed to refrain from taking promotility agents, stool softeners, opiates, benzodiazepines, or antibiotics (see Supplementary Tables 1 and 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24488/abstract). Patients were instructed to begin fasting at midnight prior to the study. WGT scintigraphy required that the patient consume a standard amount of radiolabeled ^{111}In water for the esophageal portion and the liquid gastric emptying parts of the study. The patient then consumed a radiolabeled ^{99m}Tc standard egg meal as part of the solid gastric emptying study. Anterior and posterior standing images were obtained by a gamma camera at standard times by increments of hours (1, 2, 4, 6, 24, 48, and 72) to track the transit of meals through the esophagus, stomach, and small and large intestines. Gamma cameras were placed at the front and back of the patient to monitor counts of radiation. A standard validated formula (geometric mean) was used to correct for soft tissue attenuation. Transit and emptying times were measured for each anatomic region of the gut. The standardized ranges of normal and abnormal transit, percent emptying at a given time in each region, and continuous transit times in controls were described previously (29,30).

Statistical analysis. We first used cross-sectional analysis to compare the clinical and demographic features of patients in the GAP cohort with patients in the Johns Hopkins Scleroderma Center cohort to determine whether the GAP cohort is representative overall of the scleroderma patients seen in our center. We performed chi-square or Fisher's exact tests to evaluate for associations between dichotomous clinical and demographic variables.

We then sought to determine whether physician-scored GI clinical severity as measured by the Medsger severity score (e.g., pseudo-obstruction/malabsorption, TPN dependence, etc.) is related to the presence of abnormal transit in distinct anatomical GI regions or to the extent of GI transit delays. We described WGT transit study data within each category of Medsger GI severity using both the dichotomous (presence or absence of dysmotility) and continuous data (percent emptying in an anatomical region and transit time). To determine whether GI symptoms were associated with specific GI transit abnormalities, we estimated the association between GI symptom scores (i.e., reflux, distention, diarrhea, etc.) of the UCLA GIT 2.0 and continuous measures of GI transit from the WGT studies using Spearman's rank correlations. We performed Fisher's exact tests to assess for the proportion of abnormal transit by region (e.g., esophagus, stomach, small bowel, and colon) in each category of the Medsger GI severity score. We also calculated the median (interquartile range [IQR]) for regional transit using the Medsger GI score, and compared the transit times of each region (e.g., esophagus) for a trend across Medsger severity categories using linear regression. Pearson's correlations were estimated for continuous variables, and Spearman's correlations were estimated for highly skewed continuous variables. Student's *t*-tests were used to examine differences between the means of continuous variables between 2 groups. Stata software was used to perform the analyses. A *P* value of less than or equal to 0.05 was considered statistically significant.

RESULTS

Clinical characteristics of the study cohort relative to other patients in the Johns Hopkins Scleroderma Center. Between October 2014 and February 2019, 71 patients who met criteria for SSc with GI symptoms were evaluated at the Johns Hopkins Scleroderma Center and completed WGT scintigraphy and the UCLA GIT 2.0 survey (GAP cohort). Patient

mean ± SD age with first manifestation of symptoms was 41 ± 14 years. Of these patients, 85.9% were female and 78.6% were White. The median disease duration of SSc was 6.5 years (IQR 2.6–17.6), and mean ± SD body mass index was 25.9 ± 7.3 kg/m^2. Limited cutaneous disease was present in 70% of patients. In the cohort, 18.2% of patients had evidence of cardiac involvement (Medsger cardiac score >1), and 54.6% had evidence of lung involvement (Medsger lung score >1). In addition, 14.1% of patients had evidence of myopathy, 41.8% had a Raynaud's phenomenon severity score >1, and 71.6% had sicca symptoms. Table 1 summarizes the demographic features of the cohort.

During the same time period, 1,445 SSc patients were seen in our center and did not enroll in this study. To determine whether the study cohort was representative of the rest of the Scleroderma Center cohort, we compared clinical and serologic characteristics between these 2 groups. Patients in the study group were largely comparable to the patients in the SSc cohort, though patients in the study group had a longer disease duration (from onset of first symptoms to baseline visit: 6.5 versus 4.2 years; $P < 0.01$). In the study group there were fewer patients with mild GI disease (Medsger GI severity score of 1: 22.4% versus 42.2%; $P < 0.01$) and more patients with severe GI disease (Medsger GI severity score of 3: 14.9% versus 5.5%; $P < 0.01$). In the study group, there was less severe Raynaud's phenomenon (Medsger score ≥2 in 41.8% versus 55.0%; $P = 0.03$) and a better forced vital capacity (79.9% versus 72.8%; $P < 0.01$), which is to be expected in a cohort of patients with predominantly limited cutaneous disease (31). Anticentromere antibodies were more commonly present in our study cohort (45% versus 27%; $P < 0.01$), while anti–RNAP III antibodies were significantly less prevalent (3% versus 17%; $P < 0.01$). Given that GI disease is known to be less severe among patients with anti-RNAP antibodies, as well as the higher prevalence of limited cutaneous disease in our cohort, these findings were not surprising (32–35). The distribution of other clinical features and serologies were otherwise comparable between the 2 groups.

GI characteristics of the cohort. *GI transit measured by whole gut scintigraphy is significantly different between SSc*

Table 1. Characteristics of the SSc patients with and without WGT in the Johns Hopkins Scleroderma Center cohort*

Characteristic	WGT (n = 71)	No WGT (n = 1,445)	*P*
Age at first symptoms, mean ± SD	41 ± 14	42 ± 15	0.59
Disease duration: 1st symptoms to baseline, median (IQR)	6.5 (2.6–17.6)	4.2 (1.6–11.8)	<0.01†
Male	10/71 (14.1)	222/1,442 (15.4)	0.77
Body mass index, mean ± SD	25.9 ± 7.3	NA	–
White	55/70 (78.6)	1,097/1,441 (76.1)	0.64
Ever smoker	6/53 (11.3)	101/958 (10.5)	0.86
SSc type limited cutaneous disease	49/70 (70.0)	891/1,423 (62.6)	0.25
Maximum Medsger GI score at first visit			
0	5/67 (7.5)	118/1,443 (8.2)	1.00
1	15/67 (22.4)	609/1,443 (42.2)	<0.01
2	34/67 (50.8)	615/1,443 (42.6)	0.19
3	10/67 (14.9)	80/1,443 (5.5)	<0.01
4	3/67 (4.5)	21/1,443 (1.5)	0.09
Cardiac involvement >1	10/55 (18.2)	318/131 (24.1)	0.31
Myopathy	9/64 (14.1)	314/1,427 (22.0)	0.13
Sicca	48/67 (71.6)	1,019/1,441 (70.7)	0.87
Raynaud's phenomenon severity >1	28/67 (41.8)	794/1,443 (55.0)	0.03†
Lung involvement >1	30/55 (54.6)	837/1,304 (64.0)	0.15
Cancer	16/69 (23.2)	256/1,445 (17.7)	0.25
Death	2/70 (2.9)	93/1,132 (8.2)	0.17
Pulmonary function parameters			
FVC % predicted, mean ± SD	79.9 ± 23.0	72.8 ± 19.9	<0.01†
DLco % predicted, mean ± SD	66.1 ± 26.4	64.2 ± 23.7	0.54
RVSP by echo, mean ± SD mm Hg	31.2 ± 6.8	34.7 ± 19.1	0.30
Antibodies			
Scl-70 (i.e., topoisomerase I)	10/62 (16)	291/1,162 (25)	0.11
Centromere	28/62 (45)	318/1,162 (27)	<0.01†
RNA polymerase III	2/62 (3)	193/1,162 (17)	<0.01†
Ro52	15/62 (24)	308/1,162 (27)	0.69
ThTo	4/62 (7)	91/1,162 (8)	1.00
U3 RNP	3/62 (5)	78/1,162 (7)	0.79
Ku	5/62 (8)	45/1,162 (4)	0.10
PMScl	2/62 (3)	32/1,162 (3)	0.70

* Values are the number/total number (%) unless indicated otherwise. DLco = diffusing capacity of carbon monoxide; FVC = forced vital capacity; GI = gastrointestinal; IQR = interquartile range; NA = not applicable; RVSP = right ventricular systolic pressure; SSc = systemic sclerosis; WGT = whole gut transit.
† Statistically significant.

Table 2. Objective GI involvement in SSc using the Whole Gut Transit study*

Region of the gut	WGT in SSc (n = 71)	WGT in controls (n = 18)	*P*
Esophagus			
Abnormal, no./total (%)	41/70 (59)	2/17 (12)	<0.01†
Esophageal transit time, sec.	22 (11–30)	10 (8–12)	<0.01†
Esophageal % emptying at 10 sec.	80 (62–88)	92 (86–93)	<0.01†
Stomach			
Liquid			
Abnormal, no./total (%)	16/71 (23)	1/18 (6)	0.18
Delayed T1/2, minutes	18 (13–22)	16 (11–20)	0.17
Solid			
Abnormal, no./total (%)	13/71 (18)	1/18 (6)	0.28
% emptying at 2 hours	61 (45–75)	84 (65–87)	<0.01†
% emptying at 4 hours	95 (88–98)	98 (97–99)	<0.01†
Small bowel			
Abnormal, no./total (%)	10/70 (14)	1/18 (6)	0.45
% emptying at 6 hours	73 (58–82)	72 (62–77)	0.53
Colon			
Abnormal, no./total (%)	38/69 (55)	7/18 (39)	0.22
% emptying at 72 hours	48 (0–87)	84 (60–94)	0.02†

* Values are the median (interquartile range) unless indicated otherwise. Normal ranges: esophageal transit time ≥15 seconds; esophageal emptying at 10 seconds ≥83%; normal liquid T1/2 ≤74 minutes; solid emptying 2 hours ≥40%; solid emptying 4 hours ≥90%; normal small bowel transit time at 6 hours ≥49%; normal % colonic emptying at 72 hours ≥67%. GI = gastrointestinal; sec. = seconds; SSc = systemic sclerosis; WGT = whole gut transit.
† Statistically significant.

patients and controls. As this was the first study to measure GI transit in a large population of patients with SSc using WGT scintigraphy, we first compared transit times and percent emptying between patients with SSc and controls. As expected, SSc patients had a significantly higher prevalence of abnormal esophageal function than the control group (59% versus 12%; $P < 0.01$). The median esophageal transit time was significantly delayed in SSc patients when compared to controls (22 seconds versus 10 seconds; $P < 0.01$), as was the median esophageal percentage emptying at 10 seconds (80% versus 92%; $P < 0.01$). Gastric emptying as measured by the percent emptying of solids at both 2 hours (61% versus 84%; $P < 0.01$) and 4 hours (95% versus 98%; $P < 0.01$) was significantly delayed in SSc patients compared to controls. Delayed small bowel transit was more common among SSc patients compared to controls, though the number of abnormal studies was small, and the difference in the prevalence of this abnormality was not statistically significant (14% versus 6%; $P = 0.45$). The percent colonic emptying at 72 hours was also significantly less in SSc patients compared to controls (48% versus 84%; $P = 0.02$). Table 2 summarizes the whole gut scintigraphy findings in both cohorts.

Delayed GI transit in specific parts of the gut associates with severe SSc GI complications. To determine whether distinct GI transit abnormalities associate with specific clinical SSc GI complications, we examined the prevalence of delayed transit in the esophagus, stomach, small bowel, and colon (measured by WGT) within each category of Medsger GI severity. We then compared transit times/percent emptying in the esophagus, stomach, small bowel, and colon across each category of the Medsger GI severity score (Table 3).

The esophagus and colon were most frequently abnormal on WGT scintigraphy across all categories of the Medsger GI score. Evidence of delayed esophageal transit was noted in the majority of patients without symptoms (Medsger score 0: ~60%). (11) In patients with mild symptoms of GERD (Medsger 1), less than one-half had evidence of delayed esophageal transit. In contrast, more patients with refractory GERD (Medsger 2) had both delayed esophageal transit and gastroparesis compared to patients in the Medsger 1 group, though the difference was not statistically significant.

Small bowel transit delay was rare among most Medsger GI severity groups (scores 0–3), as each of these Medsger GI groups had an estimated 11% of patients with evidence of small bowel transit delay. In contrast, among the most severe Medsger category of GI disease (Medsger 4, TPN dependence), 3 of 3 (100%) had small bowel involvement. Table 3 summarizes the association between the Medsger severity scores and WGT study findings.

Among patients scored as having recurrent pseudo-obstruction and malabsorption (Medsger GI score of 3), colonic transit was severely delayed, with a median percent emptying of 3.5% at 72 hours on WGT, which was lower than the other groups collectively ($P = 0.02$). In addition, within this group of patients, almost 1 of 3 (27%) had no colonic emptying (0%) at 72 hours, which was not the case for any other Medsger GI severity score. When looking across all groups of Medsger GI severity by linear regression, we found that there was a trend toward more severe disease and lower percent colonic emptying at 72 hours ($P = 0.07$). Finally, when comparing severe and none-to-moderate GI disease (Medsger 3 or 4 versus 0–2, respectively), patients with more severe disease had a lower mean percent

Table 3. Association between Whole Gut Transit study and Medsger severity scores*

	Score 0	Score 1	Score 2	Score 3	Score 4	Significance	Normal
Esophagus							
Abnormal, no./total (%)	3/5 (60)	6/15 (40)	19/33 (58)	9/10 (90)	2/3 (66)	0.15†	NA
ETT, seconds	19 (14–29)	11 (9–30)	26 (11–30)	29 (16–30)	29 (4–30)	0.17	<15 sec.
% emptying	81 (77–89)	84 (62–90)	81 (64–88)	69 (56–84)	77 (77–90)	0.51	≥83%
Stomach							
Abnormal solid emptying, no./total (%)	0/5 (0)	2/15 (13)	7/34 (21)	3/10 (30)	0/3 (0)	0.50†	NA
Liquid half time	17 (15–17)	16 (13–21)	18 (12–22)	21 (18–58)	22 (12–34)	0.17‡	≤25 min.
% solid 2 hours	53 (48–63)	70 (44–87)	54 (45–78)	61 (59–65)	97 (55–97)	0.41‡	≥40%
% solid 4 hours	94 (89–98)	93 (90–97)	96 (86–97)	95 (82–98)	86 (86–86)	0.72‡	≥90
Small bowel							
Abnormal, no./total (%)	0/4 (0)	2/15 (13)	4/34 (12)	1/10 (10)	3/3 (100)	0.02†	NA
% emptying at 6 hours	79 (68–92)	66 (58–86)	76 (56–82)	68 (56–82)	28 (16–40)	0.03§	≥49%
Colon							
Abnormal, no./total (%)	3/4 (75)	10/15 (67)	18/33 (55)	8/10 (80)	2/3 (66)	0.52†	NA
% emptying at 72 hours	81 (57–91)	53 (0–85)	60 (16–88)	3.5 (0–32)	18 (0–76)	0.07‡	≥67%
Disease duration from 1st symptom, years	4.0 (4–5)	10 (6–18)	12 (7–24)	20 (10–26)	12 (7–27)	0.07‡	NA

* Values are the median (interquartile range) unless indicated otherwise. Medsger severity scores defined as: 1 = requiring gastroesophageal reflux disease (GERD) medications; 2 = refractory reflux requiring high-dose GERD medications and/or evidence of small bowel dilation on radiography; 3 = pseudo-obstruction and/or malabsorption syndrome; 4 = total parenteral nutrition required. ETT = esophageal transit time; min. = minutes; NA = not applicable; sec. = seconds.
† Significance determined by Fisher's exact test.
‡ Significance in trend across Medsger score determined by linear regression.
§ Statistically significant. Significance in trend across Medsger score determined by linear regression.

emptying of the colon at 72 hours compared to those with none-to-moderate disease (27% versus 53%; $P = 0.04$).

GI symptoms (GIT 2.0) associate with GI transit delays by WGT in SSc. The median interval between the collection of the UCLA GIT 2.0 survey and the WGT study was –0.08 (IQR –5.08, 8.75) months. Using the UCLA GIT 2.0 scale, greater reflux scores showed a modest association with longer esophageal transit time ($r = 0.27$, $P = 0.05$), slower percentage of esophageal emptying at 10 seconds ($r = –0.31$, $P = 0.05$) and delayed gastric emptying at 3 hours ($r = –0.34$, $P = 0.05$). Esophageal transit time was positively associated with GIT diarrhea scores ($r = 0.37$, $P < 0.05$). Patient-reported symptoms of distention and bloating were inversely associated with percent gastric emptying at 3 hours ($r = –0.27$, $P = 0.06$) (Table 4). Gastric emptying at 3 hours was also inversely associated with a higher (more severe) total GIT score ($r = –0.30$, $P < 0.05$) and a trend toward worse patient-reported social well-being ($r = –0.26$, $P = 0.06$). However, symptoms determined by the constipation domain of the GIT did not show significant associations with objective findings of delayed colonic transit on WGT studies. Table 4 summarizes the correlation between WGT results and patient-reported symptoms as measured by the UCLA GIT 2.0 survey scores.

DISCUSSION

In this study, we sought to examine whether abnormal GI transit contributes to GI severity and symptoms in SSc. We found

Table 4. Spearman's correlation table between Whole Gut Transit study and GI symptoms (UCLA GIT 2.0)*

	Reflux	Distension/bloating	Soilage	Diarrhea	Social†	Emotional†	Constipation	Total score
ETT	0.27‡	0.16	0.02	0.37§	0.23	0.14	–0.06	0.17
E10s	–0.31‡	0.00	–0.06	–0.25	–0.15	–0.04	0.08	–0.10
Stomach 1 hour	–0.16	–0.09	0.01	0.00	–0.02	–0.04	0.04	–0.12
Stomach 2 hours	–0.20	–0.19	0.08	–0.12	–0.09	–0.09	–0.08	–0.14
Stomach 3 hours	–0.34‡	–0.27¶	–0.01	–0.22	–0.26	–0.03	–0.04	–0.30§
Stomach 4 hours	–0.14	–0.12	0.09	0.03	–0.12	–0.05	0.05	–0.09
Small bowel	0.03	–0.22	0.04	–0.14	–0.23	–0.25	–0.12	–0.20
Large bowel	–0.01	–0.01	–0.20	0.25	0.08	–0.21	0.12	–0.06

* University of California Los Angeles Scleroderma Clinical Trial Consortium Gastrointestinal Tract (UCLA GIT) 2.0 patient-reported scores were compared to findings on whole gut transit studies using Spearman's rank-order correlation. ETT = esophageal transit time; E10s = percentage of esophageal emptying at 10 seconds; Stomach (x) hour = solid emptying of the stomach at (x) hours; Small bowel = percentage of small bowel emptying at 6 hours; Large bowel = percentage of colonic emptying at 72 hours.
† Social impact of gastrointestinal (GI) symptoms; emotional impact of GI symptoms.
‡ $P = 0.05$.
§ $P < 0.05$.
¶ $P = 0.06$.

that patients with pseudo-obstruction and/or malabsorption syndrome are more likely to have severe colonic transit delays, with one-third of such patients having almost no colonic emptying at 72 hours. We also determined that patients on TPN are significantly more likely to have small bowel involvement when compared to other Medsger GI severity groups. Patients with Medsger GI scores representative of more significant upper GI symptoms (Medsger GI score of 1 or 2) are more likely to have dysmotility of both the esophagus and/or stomach. These data are important in providing insight on the impact of transit defects on SSc GI complications. Finally, we determined that patient symptoms as measured by the UCLA GIT 2.0 are associated with delayed transit.

In this study, small bowel involvement was significantly more prevalent among SSc patients with the most severe GI disease requiring TPN. Interestingly, these patients were not more likely than other groups to have gastroparesis and were less likely than other groups to have delayed colonic transit. Though the number of patients was small in this analysis, the association highlights the importance of small bowel function in optimizing nutrition in SSc. Recognizing the high prevalence of small bowel transit delays in this group of patients with severe disease and a high morbidity and mortality also emphasizes the need for the earlier application of targeted clinical therapies that positively influence small bowel transit, such as octreotide (18).

The finding that severe colonic transit delays were more likely in patients with pseudo-obstruction (Medsger score 3) and/or malabsorption syndrome compared to other Medsger GI groups ($P = 0.02$) is also interesting. The majority of patients in our study with pseudo-obstruction and/or malabsorption (70%) had ≤30% emptying of the colon at 72 hours (normal ≥67%), less than one-half of what would normally be expected. This finding reflects the importance of abnormal colonic motility in SSc, either as a marker of more generalized dysmotility or as a direct contributor to pseudo-obstruction via upstream reflexes. This finding is consistent with recent studies showing that colonic dysfunction leads to significant morbidity and mortality in SSc and lends to the hypothesis that early treatment of patients with delayed colonic transit with promotility agents, such as prucalopride, may help prevent this complication (36,37).

We also demonstrated that symptoms based on UCLA GIT 2.0 scores correlate with objective GI transit abnormalities in patients with SSc. We identified a moderate inverse association between UCLA GIT 2.0 GERD domain scores and esophageal transit times and gastric emptying. These findings bordered on statistical significance and were only modest, possibly because symptoms of heartburn, dysphagia, and regurgitation (captured in the GERD domain) in SSc may also be due to other causes, such as a hypotensive lower esophageal sphincter or gastroparesis, neither of which would necessarily affect esophageal transit time. Thus, patients with highly symptomatic GERD may have normal esophageal transit. In addition, symptoms of distention and bloating were significantly associated with delayed gastric transit at 3 hours. These results were similar to prior studies that demonstrated association with epigastric fullness and prolonged gastric emptying (12).

Finally, we also confirmed that SSc can affect several regions of the GI tract simultaneously, most commonly the esophagus and colon, and that a negative test in 1 region does not preclude abnormalities in other regions. Prior studies have demonstrated a high correlation between delayed gastric and esophageal emptying (12,38). Furthermore, a significant correlation was noted between the rate of gastric emptying and abnormal esophageal transit values, suggesting that worsening severity could coexist and extend between regions (12,38). These results suggest that delayed gastric emptying can lead to reflux and possible delays in esophageal transit. Determining early on whether symptoms of GERD are occurring in the presence of significant gastric transit delays may lead to more effective symptom management (e.g., combining metoclopramide with a standard GERD regimen).

Our study has several strengths and limitations. This is the first study to assess WGT in a large SSc population. The strengths of our study include evaluating WGT using a diverse cohort of well-characterized patients. We intentionally enriched our cohort with patients who had more severe GI disease so as to learn about the impact of abnormal transit on less frequently observed, but more severe SSc GI complications. We correlated the results of WGT studies with validated patient- and physician-reported outcome measures used to assess GI severity and symptoms of GI dysfunction. From the standpoint of limitations, there is a known lack of standardization in whole gut scintigraphy protocols and interpretation, which may affect reliability of results when compared between centers (39). The time interval between our symptom surveys and the WGT study also limits the interpretation of our findings, as a subset of surveys were collected retrospectively. Our study also did not address how disease-modifying agents impact whole gut scintigraphy results in SSc patients, which may merit further investigation through future studies. Finally, we recognize that transit studies are only 1 measure of motility and may fail to capture dysmotility at either an earlier stage or in some other form (e.g., lack of gastric accommodation).

WGT studies revealed that delayed transit in the small bowel and colon are associated with more severe GI complications in SSc, as currently defined by the Medsger scale. However, GI dysmotility often involves >1 region of the gut in scleroderma, and therefore more comprehensive testing may be indicated in symptomatic patients. WGT studies correlate well with the localization of symptoms in SSc (upper versus lower), and when combined with patient and physician-reported GI severity scores they may contribute to a more comprehensive approach in assessing severity of GI disease in SSc. Future studies examining the benefit of early therapeutic intervention targeting GI transit abnormalities in patients at high risk for severe GI complications are warranted.

ACKNOWLEDGMENTS

We thank Adrianne Woods, Margaret Sampedro, and Madeline Myers for their help with data acquisition for this work.

AUTHOR CONTRIBUTIONS

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

Study conception and design. McMahan, Tucker, Wigley.

Acquisition of data. McMahan, Wigley.

Analysis and interpretation of data. McMahan, Tucker, Perin, Volkmann, Kulkarni, Ziessman, Pasricha, Wigley.

REFERENCES

1. Gabrielli A, Avvedimento EV, Krieg T. Scleroderma. N Engl J Med 2009;360:1989–2003.
2. Pattanaik D, Brown M, Postlethwaite BC, Postlethwaite AE. Pathogenesis of systemic sclerosis. Front Immunol 2015;6:272.
3. Sjogren RW. Gastrointestinal motility disorders in scleroderma. Arthritis Rheum 1994;37:1265–82.
4. Sallam H, Mcneary T, Chen J. Systematic review: pathophysiology and management of gastrointestinal dysmotility in systemic sclerosis (scleroderma). Aliment Pharmacol Ther 2006;23:691–712.
5. Forbes A, Marie I. Gastrointestinal complications: the most frequent internal complications of systemic sclerosis. Rheumatology (Oxford) 2009;48 Suppl 3:iii36–9.
6. Steen VD, Medsger TA Jr. Severe organ involvement in systemic sclerosis with diffuse scleroderma. Arthritis Rheum 2000;43:2437–44.
7. Shreiner AB, Murray C, Denton C, Khanna D. Gastrointestinal manifestations of systemic sclerosis. J Scleroderma Relat Disord 2016; 1:247.
8. Ebert EC. Esophageal disease in scleroderma. J Clin Gastroenterol 2006;40:769–75.
9. Arana-Guajardo AC, Barrera-Torres G, Villarreal-Alarcón MÁ, Vega-Morales D, Esquivel-Valerio JA. Esophageal symptoms and their lack of association with high-resolution manometry in systemic sclerosis patients. Reumatol Clin (Engl Ed) 2019;15:165–9.
10. Lock G, Holstege A, Lang B, Schölmerich J. Gastrointestinal manifestations of progressive systemic sclerosis. Am J Gastroenterol 1997; 92:763–71.
11. Thonhofer R, Siegel C, Trummer M, Graninger W. Early endoscopy in systemic sclerosis without gastrointestinal symptoms. Rheumatol Int 2012;32:165–8.
12. Wegener M, Adamek RJ, Wedmann B, Jergas M, Altmeyer P. Gastrointestinal transit through esophagus, stomach, small and large intestine in patients with progressive systemic sclerosis. Dig Dis Sci 1994; 39:2209–15.
13. Al-Dhaher FF, Pope JE, Ouimet JM. Determinants of morbidity and mortality of systemic sclerosis in Canada. Semin Arthritis Rheum 2010;39:269–77.
14. Kirby DF, Chatterjee S. Evaluation and management of gastrointestinal manifestations in scleroderma. Curr Opin Rheumatol 2014;26: 621–9.
15. Nagaraja V, McMahan ZH, Getzug T, Khanna D. Management of gastrointestinal involvement in scleroderma. Curr Treatm Opt Rheumatol 2015;1:82–105.
16. Sadik R, Stotzer PO, Simrén M, Abrahamsson H. Gastrointestinal transit abnormalities are frequently detected in patients with unexplained GI symptoms at a tertiary centre. Neurogastroenterol Motil 2008;20:197–205.
17. Vigone B, Caronni M, Severino A, Bellocchi C, Baldassarri AR, Fraquelli M, et al. Preliminary safety and efficacy profile of prucalopride in the treatment of systemic sclerosis (SSc)-related intestinal involvement: results from the open label cross-over PROGASS study. Arthritis Res Ther 2017;19.
18. Soudah HC, Hasler WL, Owyang C. Effect of octreotide on intestinal motility and bacterial overgrowth in scleroderma. N Engl J Med 1991;325:1461–7.
19. Camilleri M, Piessevaux H, Yiannakou Y, Tack J, Kerstens R, Quigley EM, et al. Efficacy and safety of prucalopride in chronic constipation: an integrated analysis of six randomized, controlled clinical trials. Dig Dis Sci 2016;61:2357–72.
20. Solnes LB, Sheikhbahaei S, Ziessman HA. EnsurePlus as an alternative to the standardized egg gastric-emptying meal. Clin Nucl Med 2019;44:459–61.
21. Khanna D, Hays RD, Maranian P, Seibold JR, Impens A, Mayes MD, et al. Reliability and validity of the University of California, Los Angeles Scleroderma Clinical Trial Consortium Gastrointestinal Tract Instrument. Arthritis Rheum 2009;61:1257–63.
22. Medsger TA, Silman AJ, Steen VD, Black CM, Akesson A, Bacon PA, et al. A disease severity scale for systemic sclerosis: development and testing. J Rheumatol 1999;26:2159–67.
23. McMahan ZH, Paik JJ, Wigley FM, Hummers LK. Determining the risk factors and clinical features associated with severe gastrointestinal dysmotility in systemic sclerosis. Arthritis Care Res (Hoboken) 2018;70:1385–92.
24. Van Den Hoogen F, Khanna D, Fransen J, Johnson SR, Baron M, Tyndall A, et al. 2013 classification criteria for systemic sclerosis: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 2013;65:2737–47.
25. LeRoy EC, Medsger J. Criteria for the classification of early systemic sclerosis. J Rheumatol 2001;28:1573–6.
26. Medsger TA, Bombardieri S, Czirjak L, Scorza R, Della Rossa A, Bencivelli W. Assessment of disease severity and prognosis. Clin Exp Rheumatol 2003;21:S42–6.
27. Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general U.S. population. Am J Respir Crit Care Med 1999;159:179–87.
28. Blochowiak K, Olewicz-Gawlik A, Polanska A, Nowak-Gabryel M, Kociecki J, Witmanowski H, et al. Oral mucosal manifestations in primary and secondary Sjögren syndrome and dry mouth syndrome. Postepy Dermatol Alergol 2016;33:23–7.
29. Antoniou AJ, Raja S, El-Khouli R, Mena E, Lodge MA, Wahl RL, et al. Comprehensive radionuclide esophagogastrointestinal transit study: methodology, reference values, and initial clinical experience. J Nucl Med 2015;56:721–7.
30. Ziessman HA, Chander A, Clarke JO, Ramos A, Wahl RL. The added diagnostic value of liquid gastric emptying compared with solid emptying alone. J Nucl Med 2009;50:726–31.
31. Walker UA, Tyndall A, Czirják L, Denton C, Farge-Bancel D, Kowal-Bielecka O, et al. Clinical risk assessment of organ manifestations in systemic sclerosis: a report from the EULAR Scleroderma Trials And Research group database. Ann Rheum Dis 2007;66:754–63.
32. Jaeger VK, Wirz EG, Allanore Y, Rossbach P, Riemekasten G, Hachulla E, et al. Incidences and risk factors of organ manifestations in the early course of systemic sclerosis: a longitudinal EUSTAR study. Assassi S, ed. PLoS One 2016;11:e0163894.
33. Patterson KA, Roberts-Thomson PJ, Lester S, Tan JA, Hakendorf P, Rischmueller M, et al. Interpretation of an extended autoantibody profile in a well-characterized Australian systemic sclerosis (scleroderma) cohort using principal components analysis. Arthritis Rheumatol 2015;67:3234–44.

34. Dein E, Kuo PL, Hong YS, Hummers LK, Mecoli CA, McMahan ZH. Evaluation of risk factors for pseudo-obstruction in systemic sclerosis. Semin Arthritis Rheum 2019;49:405–10.
35. Shah AA, Laiho M, Rosen A, Casciola-Rosen L. Protective effect against cancer of antibodies to the large subunits of both RNA polymerases I and III in scleroderma. Arthritis Rheumatol 2019;71:1571–9.
36. Brandler JB, Sweetser S, Khoshbin K, Babameto M, Prokop LJ, Camilleri M. Colonic manifestations and complications are relatively underreported in systemic sclerosis: a systematic review. Am J Gastroenterol 2019;114:1847–56.
37. Valenzuela A, Li S, Becker L, Fernandez-Becker N, Khanna D, Nguyen L, et al. Intestinal pseudo-obstruction in patients with systemic sclerosis: an analysis of the Nationwide Inpatient Sample. Rheumatology (Oxford) 2016;55:654-8.
38. Zikos TA, Clarke JO, Triadafilopoulos G, Regalia KA, Sonu IS, Fernandez-Becker NQ, et al. A positive correlation between gastric and esophageal dysmotility suggests common causality. Dig Dis Sci 2018;63:3417–24.
39. Odunsi ST, Camilleri M. Selected interventions in nuclear medicine: gastrointestinal motor functions. Semin Nucl Med 2009;39:186–94.

Arthritis Care & Research
Vol. 74, No. 3, March 2022, pp 451–460
DOI 10.1002/acr.24482

American College of Rheumatology
Empowering Rheumatology Professionals

Improvement of Functioning and Health With Ixekizumab in the Treatment of Active Nonradiographic Axial Spondyloarthritis in a 52-Week, Randomized, Controlled Trial

Jessica A. Walsh,[1] Marina N. Magrey,[2] Xenofon Baraliakos,[3] Kentaro Inui,[4] Meng-Yu Weng,[5] Ennio Lubrano,[6] Désirée van der Heijde,[7] Annelies Boonen,[8] Lianne S. Gensler,[9] Vibeke Strand,[10] Jürgen Braun,[3] Theresa Hunter,[11] Xiaoqi Li,[11] Baojin Zhu,[11] Luis León,[11] David Marcelino Sandoval Calderon,[11] and Uta Kiltz[3]

Objective. To evaluate the effect of ixekizumab on self-reported functioning and health in patients with active nonradiographic axial spondyloarthritis (SpA).

Methods. COAST-X was a randomized, controlled trial conducted in patients with nonradiographic axial SpA over 52 weeks. Participants were randomized at a ratio of 1:1:1 to receive 80 mg of ixekizumab subcutaneously every 4 weeks or 2 weeks or placebo for 52 weeks. Self-reported functioning and health end points included the Medical Outcomes Study Short Form 36 (SF-36) health survey, Assessment of Spondyloarthritis International Society (ASAS) health index, and European Quality of Life-5 Dimensions-5 Level (EQ-5D-5L) health-utility descriptive system.

Results. Compared to placebo, ixekizumab treatment resulted in improvement of SF-36 physical component summary scores from baseline, with a score of 4.7 improving to 8.9 with ixekizumab therapy every 4 weeks ($P < 0.05$) and a score of 9.3 with ixekizumab therapy every 2 weeks ($P < 0.01$); the greatest improvements were observed in the domains of physical functioning, role-physical, and bodily pain at weeks 16 and 52. A higher proportion of patients receiving ixekizumab therapy every 2 weeks reported ≥3 improvements based on the ASAS health index from baseline to weeks 16 and 52 ($P < 0.05$). Significantly more patients receiving ixekizumab every 4 weeks reported improvements in "good health status" on the ASAS health index (ASAS score of ≤5) at weeks 16 and 52 ($P < 0.05$). Patients receiving ixekizumab reported improvements on the EQ-5D-5L compared to those who received placebo at week 16 (0.11 versus 0.17 for patients receiving treatment every 4 weeks and 0.19 for patients receiving treatment every 2 weeks; $P < 0.05$), which remained consistent at week 52. There were no clinical meaningful differences in responses based on the ixekizumab dosing regimen for patients who received ixekizumab therapy every 2 weeks or every 4 weeks.

Conclusion. In patients with nonradiographic axial SpA, therapy with ixekizumab was superior to placebo in the improvement of self-reported functioning and health at weeks 16 and 52.

INTRODUCTION

Axial spondyloarthritis (SpA) is a chronic inflammatory disease affecting mainly the axial skeleton (1). The term "axial SpA" encompasses patients with either radiographic axial SpA, which is also referred to as ankylosing spondylitis (AS), or nonradiographic axial SpA, which is defined by a diagnosis of axial SpA with the absence of definite sacroiliitis on radiograph (2). Among all patients with axial SpA, the proportion of patients with nonradiographic axial SpA varies. Ranges from 40% to 60% have been

ClinicalTrials.gov identifier: NCT02757352) and EudraCT database no. 2015-003938-27.

Supported by Eli Lilly and Company.

[1]Jessica A. Walsh, MD: University of Utah School of Medicine and Salt Lake City Veterans Affairs Medical Center, Salt Lake City, Utah; [2]Marina N. Magrey, MD: MetroHealth Medical Center and Case Western Reserve University School of Medicine, Cleveland, Ohio; [3]Xenofon Baraliakos, MD, PhD, Jürgen Braun, MD, PhD, Uta Kiltz, MD: Ruhr-University Bochum, Bochum, Germany, and Osaka City University Graduate School of Medicine, Osaka, Japan; [4]Kentaro Inui, MD: Osaka City University Graduate School of Medicine, Osaka, Japan; [5]Meng-Yu Weng, MD: National Cheng Kung University Medical College and Hospital, Tainan, Taiwan; [6]Ennio Lubrano, MD, PhD: University of Molise, Campobasso, Italy; [7]Désirée van der Heijde, MD, PhD: Leiden University Medical Center, Leiden, The Netherlands; [8]Annelies Boonen, MD, PhD: Maastricht University Medical Center and Caphri Research Institute, Maastricht University, Maastricht, The Netherlands; [9]Lianne S. Gensler, MD: University of California, San Francisco; [10]Vibeke Strand, MD: Stanford University School of Medicine, Palo Alto, California; [11]Theresa Hunter, PhD, Xiaoqi Li, PhD, Baojin Zhu, PhD, Luis León, PhD, David Marcelino Sandoval Calderon, MD: Eli Lilly and Company, Indianapolis, Indiana.

SIGNIFICANCE & INNOVATIONS

- Self-reported functioning and health measurements are important in understanding the impact of treatment from the perspective of the patient.
- Ixekizumab improves overall functioning and health in patients with nonradiographic axial spondyloarthritis, measured by Study Short Form 36, Assessment of Spondyloarthritis International Society health index, and European Quality of Life-5 Dimensions-5 Level health-utility descriptive system.
- Improved overall functioning and health is reported regardless of treatment regimen (80 mg ixekizumab every 2 weeks or 80 mg ixekizumab every 4 weeks).

reported (3–5). The burden of the disease is similar between individuals with nonradiographic axial SpA and individuals with AS (6–8). Patients with these conditions have comparable levels of pain, fatigue, and morning stiffness and have a patient profile characterized by impaired physical function and work productivity and an overall reduction in functioning and health.

Pharmacologic treatment with nonsteroidal antiinflammatory drugs (NSAIDs) is recommended for patients with axial SpA as a first-line treatment for improving back pain and stiffness (1,9,10). Second-line treatment comprises biologic disease-modifying antirheumatic drugs (bDMARDs) such as tumor necrosis factor inhibitors (TNFi) (9). However, patients with nonradiographic axial SpA had limited approved therapeutic options until recently. In the US, certolizumab pegol (11), ixekizumab, and secukinumab are the only biologic currently approved by the US Food and Drug Administration for the treatment of nonradiographic axial SpA, whereas in Europe, adalimumab (12), certolizumab pegol (13), etanercept (14), and golimumab (15) have been approved by the European Medicines Agency for the treatment of this disease. Approximately 60% of patients with nonradiographic axial SpA are treated with bDMARDs; however, patients often switch to another biologic due to inadequate response or intolerance (16–18). Thus, there remains a significant unmet need for patients with nonradiographic axial SpA.

Ixekizumab is an immunoglobulin G4 monoclonal antibody that selectively targets interleukin-17A with high affinity and has recently been approved in the US and European Union for the treatment of patients with active AS and nonradiographic axial SpA (19,20). The present study, COAST-X, investigated the efficacy and safety of ixekizumab in the nonradiographic axial SpA population. Ixekizumab had beneficial effects on disease activity, and it is important to note that these effects translated to improvement in the overall functioning and health of our study population. Here, we present results on self-reported functioning and health overall in individuals with nonradiographic axial SpA as measured by the Medical Outcomes Study 36-Item Short Form (SF-36) health survey, the Assessment of Spondyloarthritis International Society (ASAS) health index, and European Quality of Life-5 Dimensions (EQ-5D) descriptive system through 52 weeks of treatment.

PATIENTS AND METHODS

Study design. The COAST-X study is a phase III multicenter, randomized, controlled trial (RCT) with a 52-week duration, evaluating the efficacy and safety of ixekizumab in patients with active nonradiographic axial who are bDMARD-naive. Study protocol was reviewed and approved by applicable local ethics review boards. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committees at all sites where these studies were conducted. The RCT follows the principles of good clinical practice, standards set by the International Council for Harmonization, and local laws and regulations and conducted in accordance with the Declaration of Helsinki and its later amendments or comparable ethical standards. All enrolled patients provided written informed consent prior to study

Dr. Walsh has received consulting fees from AbbVie, Celgene, Novartis, UCB, Amgen, Pfizer and Eli Lilly and Company (less than $10,000 each) and research support from those companies. Dr. Magrey has receiving consulting fees from Eli Lily and Company, Janssen, Pfizer, and UCB (less than $10,000 each). Dr. Baraliakos has received consulting fees from AbbVie, Bristol Myers Squibb, Celgene, Janssen, MSD, Novartis, Pfizer, Roche, and UCB (< $10,000) and research support from those companies. Dr. Inui has received speaking fees from Chugai Pharmaceutical, Mitsubishi Tanabe Pharma, Daiichi Sankyo, Ono Pharmaceutical, Janssen, Pfizer, Eisai, Eli Lilly and Company, Novartis, and UCB (less than $10,000 each) and research support from those companies. Dr. Weng has received consulting fees from Eli Lilly and Company, Janssen, Novartis, Pfizer, and Roche (less than $10,000 each). Dr. Lubrano has received consulting fees from AbbVie, Bristol Meyers Squibb, Eli Lilly and Company, Janssen, MSD, Novartis, and Pfizer(less than $10,000 each). Dr. van der Heijde has received consulting fees from AbbVie, Amgen, Astellas, AstraZeneca, BMS, Boehringer Ingelheim, Celgene, Cyxone, Daiichi, Eisai, Eli Lilly and Company, Galapagos, Gilead, GlaxoSmithKline, Janssen, Merck, Novartis, Pfizer, Regeneron, Roche, Sanofi, Takeda, and UCB Pharma (less than $10,000 each) and is Director of Imaging Rheumatology at Imaging Rheumatology International. Dr. Boonen has received consulting fees from Eli Lilly and Company, Novartis, and UBC (less than $10,000 each) and has received research support from AbbVie and Celgene. Dr. Gensler has consulting fees from AbbVie, Eli Lilly and Company, GSK, Gilead, and Pfizer (less than $10,000 each) and from UCB and Novartis (more than $10,000 each). Dr. Strand has received consulting fees from AbbVie, Amgen, Arena, AstraZeneca, Bayer, Bioventus, BMS, Boehringer Ingelheim, Celltrion, Corrona, Crescendo/Myriad, EMD Serono, Equillium, Flexion, Genentech/Roche, Glenmark, GSK, Horizon, Inmedix, Janssen, Kypha, Eli Lilly, Merck, Novartis, Pfizer, Regeneron Pharmaceuticals Inc., Samsung, Samumed, Sandoz, Sanofi, Servier, Setpoint, TwoXARand, and UCB (less than $10,000 each). Dr. Braun has received consulting fees from AbbVie, Amgen, Celgene, Eli Lilly and Company, Janssen, MSD, Novartis, Pfizer, Sanofi, and UCB (less than $10,000 each). Drs. Hunter, Li, Zhu, León, and Sandoval own stocks or stock options in Eli Lilly and Company. Dr. Kiltz has received consulting fees from AbbVie, Amgen, Biogen, Eli Lilly and Company, Hexal, Janssen, MSD, Novartis, Pfizer, Roche, and UCB (less than $10,000 each) and research support from those companies.

Address correspondence to Uta Kiltz, MD, Rheumazentrum Ruhrgebiet, Ruhr-University Bochum, Claudiusstrasse 45, 44649 Herne, Germany. Email: uta.kiltz@elisabethgruppe.de.

Submitted for publication March 9, 2020; accepted in revised form September 4, 2020.

participation. Data sets generated and/or analyzed during the present study are available from the corresponding author upon reasonable request.

Study participants. Inclusion criteria have been previously detailed (21). Briefly, eligible patients were ages 18 years or older with an established diagnosis of axial SpA by a physician who fulfilled the ASAS classification criteria for nonradiographic axial SpA (22). Patients meeting the radiographic criterion of definite sacroiliitis according to the modified New York criteria (according to central reading by 2 readers and an adjudicator in case of a discrepancy) were excluded (23). Patients were also required to have disease activity at screening and at baseline (defined as having a Bath Ankylosing Spondylitis Disease Activity Index score of ≥4 and total back pain score of ≥4 on a 0–10 scale), an inadequate response to 2 or more NSAIDs or a history of intolerance to NSAIDs, and no prior treatment with bDMARDs. Patients were also required to have objective signs of inflammation, which was defined as evidence of sacroiliitis on magnetic resonance imaging (MRI; central reading by 2 readers and an adjudicator in case of a discrepancy) and/or elevated C-reactive protein (CRP) levels (>5 mg/liter). Active sacroiliitis on MRI was determined using the ASAS definition (22,24). Participants were allowed to continue background medications, including NSAIDs, conventional synthetic DMARDs (csDMARDs; methotrexate, hydroxychloroquine, and sulfasalazine), glucocorticoids, and analgesics that may be allowed if treated at a stable dose for at least 4 weeks prior to baseline randomization. If used, csDMARDs were not to be used in any combination with other csDMARDs.

Interventions. COAST-X interventions have been previously described (21). Briefly, patients were randomly assigned at a ratio of 1:1:1 to receive subcutaneous injections of ixekizumab (80 mg) every 4 weeks, subcutaneous injections of ixekizumab (80 mg) every 2 weeks, or placebo every 2 weeks. At week 0, patients assigned to ixekizumab treatment regimens were randomly assigned at a 1:1 ratio to receive a starting dose of either 80 mg ixekizumab or 160 mg ixekizumab (2 injections of 80 mg each). To maintain blinding of the study participants, all patients received 2 injections at week 0 and 1 injection every 2 weeks during the remainder of the blinded treatment dosing period. Ixekizumab and its matching placebo were visually indistinguishable from each other. Starting at week 16, patients were able to switch to open-label ixekizumab every 2 weeks or subsequent TNFi treatment (after receiving open-label treatment every 2 weeks for at least 8 weeks) if their disease activity required escalation of treatment at investigator discretion with no specific predefined criteria. Patients who had switched to open-label treatment continued to be followed up during the study. Patients, investigators, and all other personnel involved in the conduct of the study were blinded to individual treatment assignments through the 52-week blinded period. For patients who switched to open-label treatment with ixekizumab every 2 weeks, the study site personnel, patient, and study team remained blinded to the initial randomization.

Outcome measures. The effects of ixekizumab on functioning and health were assessed using 3 secondary major end points: SF-36, the ASAS health index, and EQ-5D-5 level (EQ-5D-5L). Assessments were recorded at weeks 0 (baseline), 4, 8, 16, 36, and 52 with the SF-36 and ASAS health index and at weeks 0, 16, and 52 with the EQ-5D-5L.

The SF-36 is a 36-item patient-administered measure designed as a short, generic assessment of health including the following domains: physical functioning (PF), role-physical (RP), bodily pain (BP), general health (GH), vitality (VT), role-emotional (RE), social functioning (SF), and mental health (MH) (25,26). Domain scores range from 0 to 100, with higher scores indicating better levels of function and/or better health. The physical component summary (PCS) and mental component summary (MCS) scores are calculated based on differential weighting of normalized and z-transformed 8 domain scores with normative scores of 50. Domain scores are answered based on Likert-type scales of 1 to 5. Version 2 of the SF-36 (the acute version) utilizes a 1-week recall period and has been used in the COAST-X study (25). Domains (scale 0–100, with higher scores indicating better health) were used in the spydergrams (27) as well as changes in the least squares mean (LSM) from baseline in PSC and MSC scores (Figure 1). T scores for SF-36 domains or component scores are based on the general US population norms of 2009. The calculation of age/gender–matched norms for each domain in the spydergrams (Figure 1) are based on 1998 US population norms and matched to the distribution of the protocol population.

The ASAS health index is a disease-specific health index designed to assess global functioning and health in patients with SpA. It covers areas of physical, emotional, and social functioning based on categories summarized in the ASAS/World Health Organization International Classification of Functioning, Disability, and Health core set for AS (28). This 17-item instrument has sum scores ranging from 0 (good health) to 17 (poor health) (29). Each item consists of one question that the patient needs to respond to with either "I agree" (score 1), "I do not agree" (score 0), or "not applicable" (only for items 7 and 8). If the patients choose "not applicable," the sum score is analyzed based on $n = 16$ or $n = 15$. A score of "1" is given where the item is affirmed, indicating adverse health. All item scores were summed to yield a total score or index (29). An improvement of ≥3 from baseline on the ASAS health index represents a clinically meaningful change and attaining a "good health status" is defined by having a score of ≤5 (30).

The EQ-5D-5L provides societal preferences for health states (health utility) based on 5 dimensions of health: mobility,

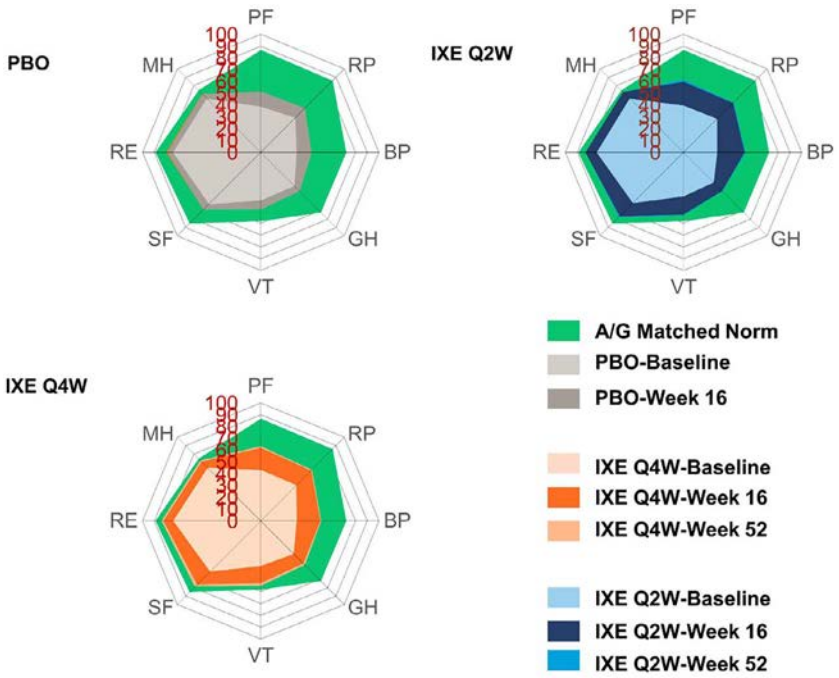


Figure 1. Medical Outcomes Study Short Form 36 (SF-36) health survey domain scores at baseline, week 16, and week 52 in the intent-to-treat population of the COAST-X trial. Subgroups included patients who received 80 mg of ixekizumab every 2 weeks (IXE Q2W), patients who received 80 mg of ixekizumab every 4 weeks (IXE Q4W), and patients who received placebo (PBO). Spydergrams depict modified baseline observation carried forward SF-36 domain scores (0–100 scale) and US age- and gender-matched normative values (A/G Matched Norm). SF-36 age- and gender-matched norms are based on the 1998 US population norms and patient counts for each age and gender distribution of the protocol population. BP = bodily pain; GH = general health MH = mental health; PF = physical functioning; RE = role-emotional; RP = role-physical; SF = social functioning; VT = vitality.

self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension can be scored on a 5-level scale: no problems, slight problems, moderate problems, severe problems, and extreme problems. The patient's completed EQ-5D-5L descriptive system was converted into a societal utility value using an available UK population–based algorithm providing a health-utility index score between −0.59 (very severe health, i.e., "worse than death") and 1.0 (perfect health [continuous variable]) (31).

Statistical analysis. Efficacy analyses were conducted on the intent-to-treat population regardless of the starting dose. The primary analysis for continuous outcomes (e.g., the SF-36 and ASAS health index) used a mixed-effects repeated measures model with treatment, geographic region (Europe and non-Europe), screening of MRI/CRP status, baseline value, visit, baseline value by visit interaction, and treatment by visit interaction as fixed factors at week 0 (baseline), 4, 8, 16, 36, and 52. When using mixed-effects repeated measures modeling, there was no prior imputation for missing data. Analyses for ASAS health index responses and good health status used logistic regression, which included treatment, geographic region, and MRI/CRP status at baseline. For continuous outcomes of EQ-5D-5L, analysis of covariance models included treatment, geographic region, screening MRI/CRP status, and baseline value. Modified baseline observation carried forward (BOCF) for missing data imputation was used with the EQ-5D-5L. For patients determined to be treatment nonresponders at the discretion of investigators who had treatment switched to open-label ixekizumab every 2 weeks, only data up to switching were included in the analyses, with data afterward treated as missing with nonresponder imputation. In patients who discontinued the study drug due to an adverse event, modified BOCF was used. In patients who discontinued the study drug for any other reason, the last nonmissing observation before discontinuation was carried forward. Patients who were randomized without at least 1 post-baseline observation were not included in the modified BOCF analysis except for those discontinuing study treatment due to the occurrence of an adverse event.

Subgroup analysis was conducted for all functioning and health end points of the proportion of patient achieving an ASAS criteria for 40% improvement (ASAS40) response at week 16 using the intent-to-treat population. A logistic regression model with treatment, subgroup, and the interaction of subgroup by treatment included as factors was used for analysis. Treatment group differences were evaluated within each category of the subgroup using Fisher's exact test, regardless of whether the interaction was statistically significant. Missing data was imputed using nonresponder imputation.

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

Variable	Ixekizumab, 80 mg every 4 weeks (n = 96)	Ixekizumab, 80 mg every 2 weeks (n = 102)	Placebo (n = 105)
Age, years	40.9 ± 14.5	40.0 ± 12.0	39.9 ± 12.4
Female sex, no. (%)	46 (48)	53 (52)	61 (58)
BMI, kg/m^2	27.6 ± 5.4	27.3 ± 5.7	27.0 ± 5.8
Race, no. (%)			
White	80 (83)	83 (81)	76 (73)
Asian	13 (14)	11 (11)	17 (16)
Other	3 (3)	8 (8)	11 (11)
Positive for HLA–B27, no. (%)	71 (75)	73 (72)	77 (74)
Age at onset of axial SpA, years	30.1 ± 9.7	29.8 ± 9.5	30.1 ± 9.8
Duration of nonradiographic SpA symptoms, years	11.3 ± 10.7	10.6 ± 10.1	10.1 ± 8.3
Concomitant baseline medication, no. (%)			
NSAIDs	81 (84)	95 (93)	96 (91)
Methotrexate	17 (18)	15 (15)	17 (16)
Sulfasalazine	23 (24)	27 (26)	21 (20)
Glucocorticoids	8 (8)	20 (20)	14 (13)
SF-36 PCS score	33.5 ± 7.4	31.9 ± 7.5	32.6 ± 8.2
SF-36 MCS score	47.2 ± 11.8	47.7 ± 12.8	48.3 ± 11.7
ASAS health index score	8.6 ± 3.4	9.6 ± 3.4	9.0 ± 3.7
EQ-5D-5L score†	0.49 ± 0.23	0.44 ± 0.25	0.47 ± 0.22

* Values are the mean ± SD except where indicated otherwise. Percentages were calculated based on the number of patients with non-missing values. ASAS = Assessment of Spondyloarthritis International Society; BMI = body mass index; EQ-5D-5L = European Quality of Life-5 Dimensions-5 Level; MCS = mental component summary; NSAIDs = nonsteroidal antiinflammatory drugs; PCS = physical component summary; SF-36 = Medical Outcomes Study Short Form 36 health survey; SpA = spondyloarthritis.
† EQ-5D-5L UK population-based Index Score.

RESULTS

Of the 303 patients with nonradiographic axial SpA who were enrolled in the present study, 96 received ixekizumab every 4 weeks, 102 received ixekizumab every 2 weeks, and 105 received placebo (Table 1). Baseline characteristics and disease activity values at baseline were similar between the randomized treatment groups. Study patients had a mean ± SD age of 40.3 ± 13.0 years; 53% (160) of 303 patients were female, and 79% (239) were White. Patients had a mean ± SD body mass index of 27.3 ± 5.6 kg/m^2. Disease duration since nonradiographic axial SpA diagnosis was mean ± SD 10.7 ± 9.7 years, and the mean ± SD age at onset of disease was 30.0 ± 9.6 years. The proportion of HLA–B27–positive patients was 73.7% (221 of 300). The proportion of the 303 patients receiving concomitant baseline medications included the following: NSAIDs (89.8% [272]), methotrexate (16.2% [49]), sulfasalazine (23.4% [71]), and glucocorticoids (13.9% [42]).

Baseline mean ± SD scores were 32.6 ± 7.7 for SF-36 PCS score, 47.8 ± 12.1 for SF-36 MCS score, 9.1 ± 3.6 for ASAS health index score, and 0.47 ± 0.23 for EQ-5D-5L. By week 52, 62 (59%) of 105 patients who received placebo had switched to open-label ixekizumab treatment every 2 weeks as compared to 40 (42%) of 96 patients who received ixekizumab every 4 weeks and 42 (41%) of 102 patients who received ixekizumab every 2 weeks. At week 52, 34 (32%) of 105 patients who received placebo had completed the full 52-week placebo-controlled period receiving double-blind study medication compared to 52 (54%) of 96 patients who received ixekizumab every 4 weeks and 52 (51%) of 102 patients who received ixekizumab every 2 weeks.

Greater improvements in all patient-reported outcomes, including physical function and health status, were reported in both ixekizumab treatment groups versus the placebo group at weeks 16 and 52 (Figures 2–5), measured by LSM changes in SF-36 PCS scores from baseline to week 4 (4.3 for patients receiving ixekizumab every 4 weeks and 5.2 for patients receiving ixekizumab every 2 weeks compared to 2.0 for patients receiving placebo; $P = 0.015$ and $P < 0.001$, respectively), and improvements continued through week 52 of the trial (8.9 for patients receiving ixekizumab every 4 weeks and 9.3 for patients receiving ixekizumab every 2 weeks compared to 4.7 for patients receiving placebo; $P = 0.012$ and $P = 0.006$, respectively) (Figure 2). Statistically significant improvement was reported in SF-36 MCS score at week 36 in the patients who received ixekizumab every 4 weeks (a mean score of 5.33 for the ixekizumab group compared to a mean score of 2.35 for the placebo group; $P = 0.035$), with nonsignificant improvements also noted at other time points (data not shown). The beneficial effect of ixekizumab treatment on SF-36 domains at weeks 16 and 52 are shown in Figure 1, whereas modest improvements compared to baseline and age- and gender-matched norms were reported in the placebo group (Figure 1). The largest improvements were

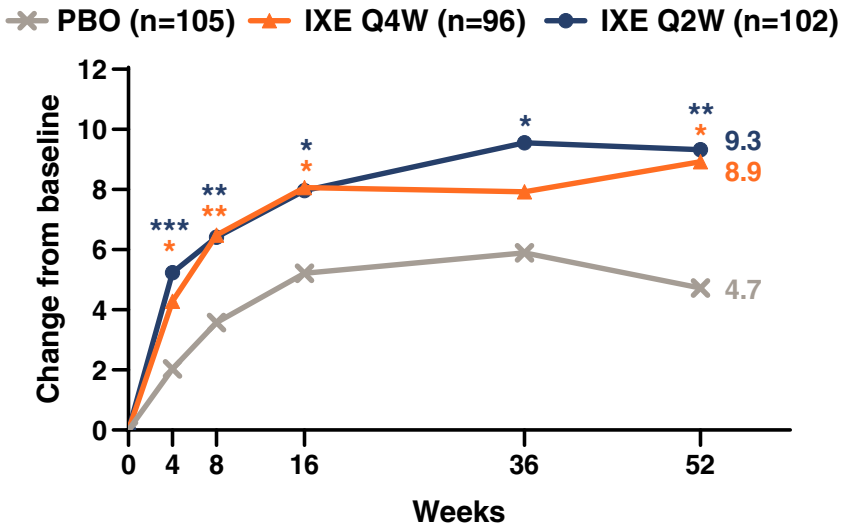


Figure 2. Medical Outcomes Study Short Form 36 health survey physical component summary scores, with least squares mean change from baseline observed in the intent-to-treat population of the COAST-X trial. Subgroups included patients who received 80 mg of ixekizumab every 2 weeks (IXE Q2W), patients who received 80 mg of ixekizumab every 4 weeks (IXE Q4W), and patients who received placebo (PBO). Comparisons between the ixekizumab treatment groups and the placebo group were made using a mixed-effects model for repeated measures. * = $P < 0.05$; ** = $P < 0.01$; *** = $P < 0.001$.

reported in PF, RP, BP, and SF domains. Improvements in RE, VT, SF, and MH domains approached those seen in matched US normative values.

At baseline, ASAS health index scores were symmetrically distributed with a median score of 9.0 (Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24482/abstract). At week 16, score distributions shifted to a median score of 6.0 (Supplementary Figure 1). Patients treated with ixekizumab every 2 weeks showed significant improvements in ASAS health index scores at week 16 (−2.74 for the patients who received ixekizumab every 2 weeks versus −1.76 for the patients who received placebo; $P = 0.023$), with numerically greater improvements in ASAS health index changes from baseline in both ixekizumab groups compared to the placebo group through week 52 (Figure 3).

ASAS health index improvements of ≥3 from baseline to week 16 were reported by 40.4% of the patients who received

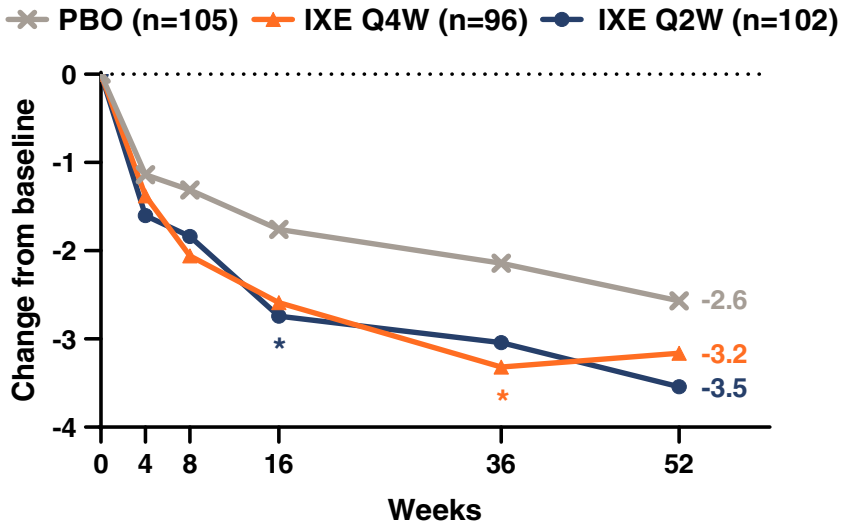


Figure 3. Assessment of Spondyloarthritis International Society health index least squares mean change from baseline in the intent-to-treat population of the COAST-X trial. Subgroups included patients who received 80 mg of ixekizumab every 2 weeks (IXE Q2W), patients who received 80 mg of ixekizumab every 4 weeks (IXE Q4W), and patients who received placebo (PBO). Comparisons between the ixekizumab treatment groups and the placebo group were made using a mixed-effects model for repeated measures. * = $P < 0.05$.

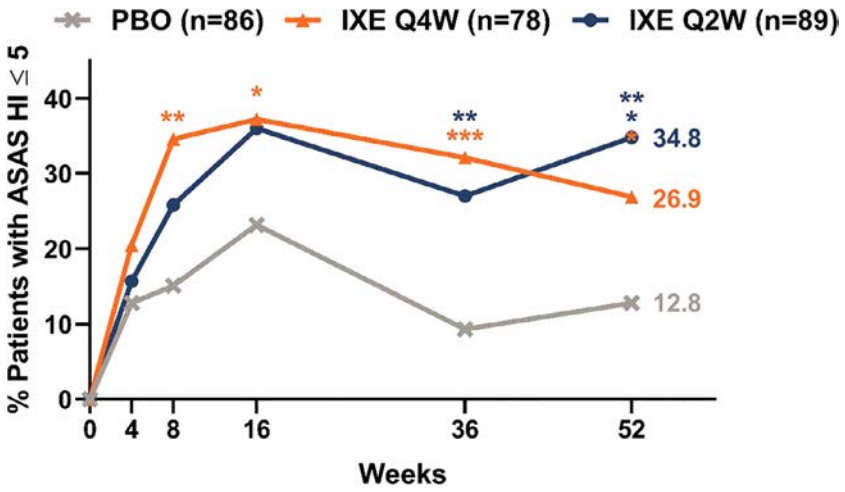


Figure 4. Percentage of patients achieving an Assessment of Spondyloarthritis International Society health index (ASAS HI) score of ≤5, indicating "good health status," in the intent-to-treat population of the COAST-X trial. Missing data were imputed with a nonresponder imputation. Subgroups included patients who received 80 mg of ixekizumab every 2 weeks (IXE Q2W), patients who received 80 mg of ixekizumab every 4 weeks (IXE Q4W), and patients who received placebo (PBO). Comparisons between the ixekizumab treatment groups and the placebo group were made using a logistic regression model. * = $P < 0.05$; ** = $P < 0.01$; *** = $P < 0.001$.

ixekizumab every 4 weeks ($P = 0.198$) and 49.0% of the patients who received ixekizumab every 2 weeks ($P = 0.017$) versus 32 (31.7%) of 101 patients who received placebo. Improvements on the ASAS health index were also observed at week 52 in 31 (33.0%) of 94 patients who received ixekizumab every 4 weeks ($P = 0.027$) and 35 (34.3%) of 102 patients who received ixekizumab every 2 weeks ($P = 0.02$) compared to 19 (18.8%) of 101 patients who received placebo (Supplementary Figure 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24482/abstract). Proportions of patients reporting "good health status" (ASAS health index score of ≤5 with a baseline score of >5) at week 16 included 29 (37.2%) of 78 patients who received ixekizumab treatment every 4 weeks ($P = 0.034$) and 32 (36.0%) of 89 patients who received ixekizumab therapy every 2 weeks compared to 19 (22.1%) of 86 patients who received placebo (Figure 4). At week 52, responses of "good health status" on the ASAS health index were reported by 21 (26.9%) of 78 patients who received ixekizumab every 4 weeks ($P = 0.02$) and 31 (34.8%) of 89 patients who received ixekizumab every 23 weeks ($P < 0.001$) compared to 11 (12.8%) of 86 patients who received placebo. A significantly higher score of ASAS health index responses of ≤5 was reported in the patient group that received ixekizumab therapy every 4 weeks compared to the patient group that received placebo from week 8 (34.6% [27 of 78] versus 15.1% [13 of 86]; $P = 0.005$).

Patients in each ixekizumab treatment group reported greater increases in health utility scores compared to the patients in the placebo group, as measured by the EQ-5D-5L (Figure 5). At week 16, patients treated with ixekizumab reported significant

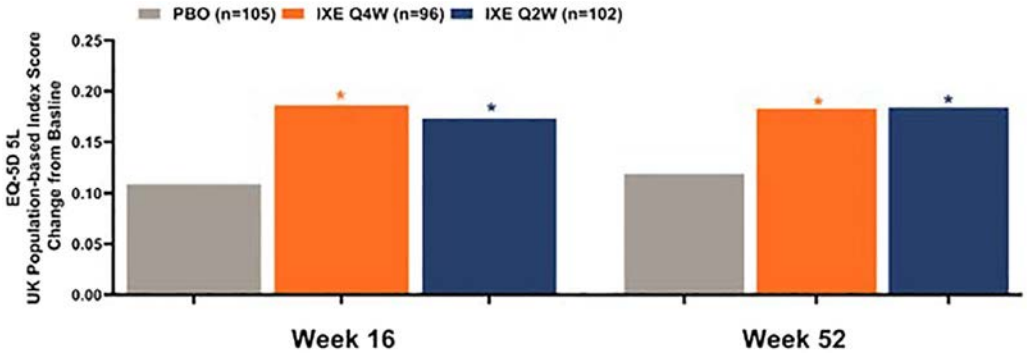


Figure 5. European Quality of Life-5 Dimensions-5 Level (EQ-5D-5L) UK population–based index scores, with the least squares mean change from baseline in the intent-to-treat population of the COAST-X trial. Missing data were imputed using modified baseline observation carried forward. Subgroups included patients who received 80 mg of ixekizumab every 2 weeks (IXE Q2W), patients who received 80 mg of ixekizumab every 4 weeks (IXE Q4W), and patients who received placebo (PBO). Comparisons between the ixekizumab treatment groups and the placebo group were made using a logistic regression model. * = $P < 0.05$.

improvements on the EQ-5D-5L compared to patients who received placebo (0.19 for patients who received ixekizumab every 4 weeks and 0.17 for patients who received ixekizumab every 2 weeks versus 0.11 for patients who received placebo; $P = 0.011$ and $P = 0.033$, respectively, for the ixekizumab groups and placebo group), and changes maintained at week 52 (0.18 for both patients treated with ixekizumab every 4 weeks and every 2 weeks versus 0.12 for patients who received placebo; $P = 0.041$ and $P = 0.036$, respectively, for the ixekizumab groups and placebo group).

Changes in self-reported functioning and health outcomes were further analyzed in a subgroup analysis of ASAS40 responders (n = 95) versus nonresponders (n = 198) at week 16 (Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24482/abstract). Significantly greater improvements were reported in LSM changes from baseline in SF-36 PCS scores (placebo: 14.4 versus 3.8, ixekizumab every 4 weeks: 13.6 versus 4.4, and ixekizumab every 2 weeks: 13.6 versus 4.4; $P < 0.001$), ASAS health index scores (placebo: −4.5 versus −1.3, ixekizumab every 4 weeks: −4.7 versus −1.3, and ixekizumab every 2 weeks: −4.5 versus −1.8; $P < 0.001$), ASAS health index responses of ≤5 (placebo: 46.7% versus 17.9%, ixekizumab every 4 weeks: 79.2% versus 18.5%, and ixekizumab every 2 weeks: 68.8% versus 18.9%; $P < 0.05$, $P < 0.001$, and $P < 0.01$, respectively), ASAS health index improvements of ≥3 (placebo: 68.4% versus 24.7%, ixekizumab every 4 weeks: 69.7% versus 24.6%, and ixekizumab every 2 weeks: 70.7% versus 36.8%; $P < 0.01$), and LSM changes from baseline on the EQ-5D-5L (placebo: 0.31 versus 0.09, ixekizumab every 4 weeks: 0.29 versus 0.11, and ixekizumab every 2 weeks: 0.33 versus 0.08; $P < 0.001$).

DISCUSSION

Nonradiographic axial SpA is a chronic inflammatory disease that affects the functioning and health of patients in a similar fashion to AS (radiographic axial SpA). The efficacy of ixekizumab as reported in this 52-week placebo-controlled trial illustrate clinically relevant and statistically significant differences as measured by the SF-36, ASAS health index, and EQ-5D-5L. ASAS40 responders reported greater improvements compared to nonresponders across all end points assessed. Patterns of response appeared similar between the 2 dosing regimens, although the study was not designed to statistically compare dosing groups.

Compared to age- and gender-matched population norms, impairments in function and health were present at baseline and largest in the SF-36 physical domains, and significant improvements in SF-36 PCS scores were reported at nearly all time points in patients who received ixekizumab treatment compared to patients who received placebo. Baseline mental domain scores approached age- and gender-matched matched norms; yet despite small margins for improvement, improvements in SF-36 MCS scores were numerically greater with ixekizumab therapy compared to placebo.

Importantly, the improvements in functioning and health, as measured by the SF-36 PCS, occurred rapidly, with statistically significant improvements at the first time point assessed (week 4) between the ixekizumab and placebo groups. In contrast, statistically significant improvements in functioning and health, as measured by the ASAS health index, were first observed at slightly later time points (weeks 8–16). Within limitations of comparison, we might speculate that emotional aspects and the financial impact of disease, which are included in the ASAS health index, are less likely to change quickly after treatment initiation and thus may be less sensitive to early improvement.

The positive overall functioning and health outcomes reported by patients with nonradiographic axial SpA in this 52-week placebo-controlled trial are consistent with results from phase III placebo-controlled studies with anti-TNF agents in nonradiographic axial SpA (12,15,32–35). However, 29–48% of patients with nonradiographic axial SpA still have active disease (based on ASAS20 responses at week 12) despite TNFi treatment (12–15); therefore, alternative treatments for TNFi are valuable.

The main strength of the present study was the sizeable patient numbers included in each group, which provided valuable information of the efficacy of ixekizumab on self-reported health and functioning outcomes in patients with nonradiographic axial SpA through week 52 of the trial. A limitation of the study is the lack of data in patients who had been previously exposed to TNFi. Another limitation of the study is that patients were allowed to switch to open-label therapy with no prespecified switching criteria. Switching to open-label therapy occurred only at the discretion of the principal investigator, which accounts for a significant proportion of patients assessed as "nonresponders."

In conclusion, ixekizumab was superior to placebo in improving overall functioning and health in patients with nonradiographic axial SpA at week 16 and 52. Ixekizumab therapy every 4 weeks and every 2 weeks was effective in showing significant levels of improvement in the study patients. These findings demonstrate that ixekizumab is effective in improving the overall functioning and health of patients affected with active nonradiographic axial SpA.

ACKNOWLEDGMENTS

The authors thank the study participants, caregivers, and investigators.

AUTHOR CONTRIBUTIONS

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

Study conception and design. van der Heijde, Braun, Hunter, León, Sandoval Calderon.

Acquisition of data. Inui, Braun, Li.
Analysis and interpretation of data. Walsh, Magrey, Baraliakos, Inui, Weng, Lubrano, van der Heijde, Boonen, Gensler, Strand, Braun, Hunter, Li, Zhu, León, Sandoval Calderon, Kiltz.

ROLE OF THE STUDY SPONSOR

Sponsorship for this study and article processing charges were funded by Eli Lilly and Company. Eli Lilly and Company contributed to study design, data collection, data analysis, data interpretation, manuscript preparation, and the decision to submit the paper for publication. An advisory committee was involved in the study design and data interpretation, together with authors from Eli Lilly and Company (Indianapolis, Indiana). Authors had full access to all group-level data in the study but not individual-level data that would risk unblinding those authors who were also study investigators. Authors had final responsibility for the decision to submit for publication. Elsa Mevel, PhD, provided writing and editorial assistance for this article, and funding for this assistance was provided by the study sponsor.

ADDITIONAL DISCLOSURES

Authors Hunter, Li, Zhu, León, and Sandoval are employees of Eli Lilly and Company.

REFERENCES

1. Sieper J, Poddubnyy D. Axial spondyloarthritis. Lancet 2017;390: 73–84.
2. Rudwaleit M, van der Heijde D, Landewe R, Listing J, Akkoc N, Brandt J, et al. The development of Assessment of SpondyloArthritis international Society classification criteria for axial spondyloarthritis (part II): validation and final selection. Ann Rheum Dis 2009;68: 777–83.
3. Strand V, Rao SA, Shillington AC, Cifaldi MA, McGuire M, Ruderman EM. Prevalence of axial spondyloarthritis in United States rheumatology practices: Assessment of SpondyloArthritis International Society criteria versus rheumatology expert clinical diagnosis. Arthritis Care Res (Hoboken) 2013;65:1299–306.
4. Rudwaleit M, Sieper J. Referral strategies for early diagnosis of axial spondyloarthritis. Nat Rev Rheumatol 2012;8:262–8.
5. Protopopov M, Poddubnyy D. Radiographic progression in non-radiographic axial spondyloarthritis. Expert Rev Clin Immunol 2018; 14:525–33.
6. Boonen A, Sieper J, van der Heijde D, Dougados M, Bukowski JF, Valluri S, et al. The burden of non-radiographic axial spondyloarthritis. Semin Arthritis Rheum 2015;44:556–62.
7. Mease PJ, Heijde DV, Karki C, Palmer JB, Liu M, Pandurengan R, et al. Characterization of patients with ankylosing spondylitis and non-radiographic axial spondyloarthritis in the US-Based Corrona Registry. Arthritis Care Res (Hoboken) 2018;70:1661–70.
8. Kiltz U, Baraliakos X, Karakostas P, Igelmann M, Kalthoff L, Klink C, et al. Do patients with non-radiographic axial spondylarthritis differ from patients with ankylosing spondylitis? Arthritis Care Res (Hoboken) 2012;64:1415–22.
9. Van der Heijde D, Ramiro S, Landewe R, Baraliakos X, Van den Bosch F, Sepriano A, et al. 2016 update of the ASAS-EULAR management recommendations for axial spondyloarthritis. Ann Rheum Dis 2017;76:978–91.
10. Ward MM, Deodhar A, Gensler LS, Dubreuil M, Yu D, Khan MA, et al. 2019 update of the American College of Rheumatology/Spondylitis Association of America/Spondyloarthritis research and treatment network recommendations for the treatment of ankylosing spondylitis and nonradiographic axial spondyloarthritis. Arthritis Rheumatol 2019;71:1599–613.
11. Deodhar A, Gensler LS, Kay J, Maksymowych WP, Haroon N, Landewe R, et al. A fifty-two-week, randomized, placebo-controlled trial of certolizumab pegol in nonradiographic axial spondyloarthritis. Arthritis Rheumatol 2019;71:1101–11.
12. Sieper J, van der Heijde D, Dougados M, Mease PJ, Maksymowych WP, Brown MA, et al. Efficacy and safety of adalimumab in patients with non-radiographic axial spondyloarthritis: results of a randomised placebo-controlled trial (ABILITY-1). Ann Rheum Dis 2013;72: 815–22.
13. Landewe R, Braun J, Deodhar A, Dougados M, Maksymowych WP, Mease PJ, et al. Efficacy of certolizumab pegol on signs and symptoms of axial spondyloarthritis including ankylosing spondylitis: 24-week results of a double-blind randomised placebo-controlled Phase 3 study. Ann Rheum Dis 2014;73:39–47.
14. Dougados M, van der Heijde D, Sieper J, Braun J, Maksymowych WP, Citera G, et al. Symptomatic efficacy of etanercept and its effects on objective signs of inflammation in early nonradiographic axial spondyloarthritis: a multicenter, randomized, double-blind, placebo-controlled trial. Arthritis Rheumatol 2014;66:2091–102.
15. Sieper J, van der Heijde D, Dougados M, Maksymowych WP, Scott BB, Boice JA, et al. A randomized, double-blind, placebo-controlled, sixteen-week study of subcutaneous golimumab in patients with active nonradiographic axial spondyloarthritis. Arthritis Rheumatol 2015;67:2702–12.
16. Navarro-Compan V, Plasencia-Rodriguez C, de Miguel E, Diaz Del Campo P, Balsa A, Gratacos J. Switching biological disease-modifying antirheumatic drugs in patients with axial spondyloarthritis: results from a systematic literature review. RMD Open 2017;3: e000524.
17. Deodhar A, Yu D. Switching tumor necrosis factor inhibitors in the treatment of axial spondyloarthritis. Semin Arthritis Rheum 2017;47: 343–50.
18. Glintborg B, Ostergaard M, Krogh NS, Tarp U, Manilo N, Loft AG, et al. Clinical response, drug survival and predictors thereof in 432 ankylosing spondylitis patients after switching tumour necrosis factor alpha inhibitor therapy: results from the Danish nationwide DANBIO registry. Ann Rheum Dis 2013;72:1149–55.
19. Liu L, Lu J, Allan BW, Tang Y, Tetreault J, Chow CK, et al. Generation and characterization of ixekizumab, a humanized monoclonal antibody that neutralizes interleukin-17A. J Inflamm Res 2016;9:39–50.
20. Toussirot E. Ixekizumab: an anti- IL-17A monoclonal antibody for the treatment of psoriatic arthritis. Expert Opin Biol Ther 2018;18: 101–7.
21. Deodhar A, van der Heijde D, Gensler LS, Kim TH, Maksymowych WP, Ostergaard M, et al. Ixekizumab for patients with non-radiographic axial spondyloarthritis (COAST-X): a randomised, placebo-controlled trial. Lancet 2020;395:53–64.
22. Rudwaleit M, Jurik AG, Hermann KG, Landewe R, van der Heijde D, Baraliakos X, et al. Defining active sacroiliitis on magnetic resonance imaging (MRI) for classification of axial spondyloarthritis: a consensual approach by the ASAS/OMERACT MRI group. Ann Rheum Dis 2009; 68:1520–7.
23. 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.
24. Lambert RG, Bakker PA, van der Heijde D, Weber U, Rudwaleit M, Hermann KG, et al. Defining active sacroiliitis on MRI for classification of axial spondyloarthritis: update by the ASAS MRI working group. Ann Rheum Dis 2016;75:1958–63.
25. Ware JE, Jr. SF-36 health survey update. Spine (Phila Pa 1976) 2000; 25:3130–9.

26. Ware JE, Kosinski M. Interpreting SF-36 summary health measures: a response. Qual Life Res 2001;10:405–13.

27. Strand V, Crawford B, Singh J, Choy E, Smolen JS, Khanna D. Use of "spydergrams" to present and interpret SF-36 health-related quality of life data across rheumatic diseases. Ann Rheum Dis 2009;68:1800–4.

28. Boonen A, Braun J, van der Horst Bruinsma IE, Huang F, Maksymowych W, Kostanjsek N, et al. ASAS/WHO ICF Core Sets for ankylosing spondylitis (AS): how to classify the impact of AS on functioning and health. Ann Rheum Dis 2010;69:102–7.

29. Kiltz U, van der Heijde D, Boonen A, Cieza A, Stucki G, Khan MA, 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.

30. Kiltz U, van der Heijde D, Boonen A, Akkoc N, Bautista-Molano W, Burgos-Vargas R, et al. Measurement properties of the ASAS Health Index: results of a global study in patients with axial and peripheral spondyloarthritis. Ann Rheum Dis 2018;77:1311–7.

31. Szende A OM, Devlin N. EQ-5D value sets: inventory, comparative review and user guide: Springer; 2007.

32. Maksymowych WP, Dougados M, van der Heijde D, Sieper J, Braun J, Citera G, et al. Clinical and MRI responses to etanercept in early non-radiographic axial spondyloarthritis: 48-week results from the EMBARK study. Ann Rheum Dis 2016;75:1328–35.

33. Sieper J, Kivitz A, van Tubergen A, Deodhar A, Coteur G, Woltering F, et al. Impact of certolizumab pegol on patient-reported outcomes in patients with axial spondyloarthritis. Arthritis Care Res (Hoboken) 2015;67:1475–80.

34. Sieper J, Kivitz A, van Tubergen A, Deodhar A, Szegvari B, Nurminen T, et al. Long-term maintenance of improvements in patient-reported outcomes with certolizumab pegol in patients with axial spondyloarthritis, including ankylosing spondylitis and non-radiographic axial spondyloarthritis: 96-week results of the Rapid-Axspa study. Value Health 2014;17:A385–6.

35. Van der Heijde D, Joshi A, Pangan AL, Chen N, Betts K, Mittal M, et al. ASAS40 and ASDAS clinical responses in the ABILITY-1 clinical trial translate to meaningful improvements in physical function, health-related quality of life and work productivity in patients with non-radiographic axial spondyloarthritis. Rheumatology (Oxford) 2016; 55:80–8.

Arthritis Care & Research
Vol. 74, No. 3, March 2022, pp 461–467
DOI 10.1002/acr.24490

American College of Rheumatology
Empowering Rheumatology Professionals

Unique Considerations for the Management of Gout in the Hmong Population: Examining Tertiary Encounters at a Large Regional Health Care System

Alison Lerman,[1] Elie Gertner,[2] Terese A. DeFor,[3] Morgan Brown,[3] and Jay Desai[4]

Objective. To evaluate demographic characteristics, care encounters, comorbidities, and clinical differences in Hmong and non-Hmong patients with gout.

Methods. Using retrospective chart review, all inpatient encounters (Hmong versus non-Hmong) were reviewed from 2014 to 2017. Acute or chronic gout was the primary or secondary diagnosis for the encounter.

Results. Hmong gout patients were on average 11 years younger than non-Hmong patients, but after adjustment for age, sex, and type of encounter, they had similar rates of hypertension, diabetes mellitus, and heart disease. Hmong patients had significantly decreased renal function at the time of presentation; the odds ratio of chronic kidney disease for Hmong patients was 2.33 versus 1.48 for non-Hmong patients ($P < 0.05$), the mean creatinine level was 3.3 mg/dl versus 2.0 mg/dl ($\beta = 1.35$, $P < 0.001$), and the glomerular filtration rate was 44.8 ml/minute versus 49.3 ml/minute ($\beta = -6.95$, $P < 0.001$). Hmong gout patients were more likely to use emergency care versus elective or urgent care, they were less likely to be using medications for the treatment of gout prior to admission (32.3% versus 58.2%), and the length of hospital stay was increased (8.8 versus 5.2 days; $P < 0.05$).

Conclusion. Hmong gout patients who had a tertiary care encounter were 11 years younger than non-Hmong patients with similar rates of comorbidities but had worse renal function despite the age differences. They were more likely to use emergency services, to be insured through Medicaid, and not to use preventive medications for gout prior to their encounter. Intensive efforts are needed in the Hmong population for culturally appropriate preventive care management of gout along with diabetes mellitus, hypertension, heart disease, and kidney disease.

INTRODUCTION

Gout is a major chronic disease in the US, affecting 3.9% of the US population, or 9.2 million people (1). It is an inflammatory arthritis caused by deposition of uric acid crystals in the joint space, leading to pain, swelling, and possible joint deformity. Common risk factors for the development of gout include metabolic syndrome, insulin resistance, renal insufficiency, obesity, hypertension, organ transplantation, and congestive heart failure, as well as the use of thiazide diuretics, low-dose aspirin, and cyclosporine (2). Gout disproportionately impacts men and minority populations such as the Hmong, who may also have a genetic predisposition for developing gout (1,3–5)

The Hmong are an Asian ethnic group from mountainous areas of Laos, Thailand, Vietnam, and China. During the Vietnam War, members of this group acted as covert soldiers on behalf of the US military and were later targeted for their role in aiding the US. As such, many were forced to flee and reside in refugee camps before being evacuated to the US. Large groups of Hmong have settled in Minnesota, Wisconsin, and California. According to the 2016 census, 296,890 Hmong live in the US, with 76,727 Hmong living in Minnesota, predominantly in Hennepin and Ramsey counties. Prior work has shown that Hmong have an earlier onset of gout and more tophaceous gout (6). The Hmong words for gout are "mob ko taw vwm," which translate to "crazy foot pain," and Hmong elders state that while gout

[1]Alison Lerman, MD: Regions Hospital, Saint Paul, Minnesota, and University of Minnesota, Minneapolis; [2]Elie Gertner, MD: Regions Hospital, Saint Paul, Minnesota, University of Minnesota, Minneapolis, and HealthPartners Institute, Bloomington, Minnesota; [3]Terese A. DeFor, MS, Morgan Brown, MS: HealthPartners Institute, Bloomington, Minnesota; [4]Jay Desai, PhD, MPH: HealthPartners Institute, Bloomington, Minnesota, and Minnesota Department of Health, Saint Paul.

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

Address correspondence to Elie Gertner, MD, Section of Rheumatology Regions Hospital, MS-11107E, 640 Jackson Street, Saint Paul, MN 55101. Email: elie.x.gertner@healthpartners.com.

Submitted for publication May 20, 2020; accepted in revised form October 15, 2020.

SIGNIFICANCE & INNOVATIONS

- Hmong patients develop earlier, more severe gout.
- Hmong patients have worse renal function at presentation and longer hospital stays and are less likely to use preventative medications.
- Intensive efforts are needed for culturally appropriate preventative care.

was a recognized illness in Laos, it is more common now among Hmong living in the US (4). However, Hmong patients with gout have been shown to be less likely to have comorbidities such as hypertension, diabetes mellitus, chronic kidney disease, and obesity at initial presentation (6). They are also less likely to be taking diuretics or report heavy alcohol use (6). One prior study showed that Hmong patients are prescribed allopurinol more often than non-Hmong individuals, but they take the medication for a significantly shorter length of time (6).

Because of the markedly increased burden of gout and tophaceous disease in the Hmong and our experience of more treatment-refractory disease in this group, we examined the demographic and clinical characteristics, the medication use, and the length of stays of Hmong versus non-Hmong patients admitted to a large regional hospital.

PATIENTS AND METHODS

HealthPartners Regions Hospital is a level I tertiary care hospital serving Saint Paul, Minnesota and western Wisconsin. We examined all hospital encounters for patients ages ≥18 years from January 1, 2014 through December 31, 2017. Hospital encounters included urgent care, emergency department, and hospitalizations. Gout-related encounters were defined as International Classification of Diseases, Ninth Revision (ICD-9) codes 274.x and ICD-10 codes M10.x and M1A.x listed as a primary or secondary diagnosis on medical or discharge record.

Patients were classified as Hmong if their primary language was Hmong, or their last name was 1 of 18 clans names, and their country of origin was Laos, or their last name was 1 of 18 clan names, their country of origin was the US, and their self-reported race was Asian. The 60 last names associated with the clans were Cha, Chai, Chan, Chang, Cheng, Chieng, Chu, Chue, Fa, Fang, Hang, Her, Herr, Heu, Kang, Kha, Khang, Khue, Kong, Kue, Kwm, Lao, Le, Lee, Li, Lo, Lor, Ly, Moua, Mouacheupao, Mua, Pha, Phang, Shong, Siong, Som, Song, Soung, Tang, Thao, Thoj, Thor, Thorr, Thow, Thsa, Tsue, Va, Vaaj, Vang, Vangh, Vu, Vue, Xiong, Ya, Yang, Pa, Vangsuathao, Lochongvu, Xyaaj, and Van. This classification has been used in prior descriptive studies of the Hmong (4).

Information on patient demographic characteristics, encounter characteristics, gout-related medication use, comorbidities, laboratory clinical values, mortality, and length of hospital stay were extracted from electronic enrollment, medical, and discharge records. Comorbidities examined included diabetes mellitus (ICD-9 250.x; ICD-10 E10.x, E11.x, E13.x), hypertension (ICD-9 401.x-405.x; ICD-10 I10.x-I16.x), ischemic heart disease (ICD-9 410.x-414.x; ICD-10 I20.x-I25.x), heart failure (ICD-9 428.x; ICD-10 I50.x), chronic kidney disease stages 1, 2, 3, and 4 (ICD-9 585.x; ICD-10 N18.x), and end-stage renal disease (ICD-9 585.5x, 585.6x, V42.0x, V45.1x; ICD-10 N18.5x, N18.6x, Z94.0x, Z99.2x). Missing data were minimal, ranging from 0% to 4.2% for marital status.

Descriptive frequencies were conducted on patient demographic characteristics, encounter characteristics, medication use, and length of stay. Univariate and multivariate logistic and linear regressions compared the association of Hmong ethnicity as the main predictor of the outcome of interest. Regression models controlled for age, sex, type of encounter, and principal diagnosis of gout.

RESULTS

During the study period, there were 120,303 encounters at Regions hospital for patients ≥18 years of age. There were 2,839 Hmong encounters, with 92 (3.2%) being gout-related. There were 117,464 non-Hmong encounters, with 1,262 (1.1%) being gout related. These gout-related hospital encounters represent 65 Hmong and 976 non-Hmong patients. On average, Hmong patients had 1.4 gout-related encounters compared to 1.3 encounters for non-Hmong patients during the 4-year study period. Compared to non-Hmong patients, Hmong patients were younger and more likely to be male, married, and have Medicaid insurance (Table 1). Almost 71% of Hmong patients self-reported Hmong as their primary language, with most requiring the use of an interpreter.

Almost 91% of Hmong encounters were inpatient, with 86% via the emergency room (Table 1). This compares to 85% and 56% for the non-Hmong patients. Many non-Hmong patient encounters were elective or through urgent care. Gout was the principal reason for the encounter for 15% of Hmong and 7% of non-Hmong patients. The mean length of a hospital stay was 8.8 days for Hmong patients with gout compared to 5.2 days for non-Hmong patients with gout ($P < 0.05$).

Medication use in the year prior to and during a patient's first documented encounter varied by ethnicity. Hmong patients typically had lower medication use prior to their first encounter compared to non-Hmong patients, but these increased during the encounter stays (Table 1). Only 32% of Hmong patients were taking gout-related medications prior to their first encounter compared to 58% for non-Hmong patients. During an encounter, this increased to 72% for both groups.

In general, Hmong patients had lower rate of diabetes mellitus, hypertension, ischemic heart disease, heart failure, and death

Table 1. Characteristics of Hmong and non-Hmong patients with gout-related encounters (n = 1,041)*

Characteristic	Hmong (n = 65)	Non-Hmong (n = 976)
Patient demographic characteristics		
Age at study entry, mean ± SD years	58.0 ± 13.1	69.1 ± 13.7
Male sex	55 (84.6)	709 (72.6)
Female patient, age <45 years	1 (1.5)	3 (0.3)
Race		
American Indian/Alaskan Native	0 (0)	3 (0.3)
Asian	63 (96.9)	47 (4.8)
Black	0 (0)	144 (14.8)
Native Hawaiian or other Pacific Islander	0 (0)	2 (0.2)
Other	1 (1.5)	10 (1.0)
White	0 (0)	738 (75.6)
Hispanic or Latino	0 (0)	21 (2.2)
Use of an interpreter	43 (66.2)	51 (5.2)
Primary language		
English	19 (29.2)	924 (94.7)
Hmong	46 (70.8)	0 (0)
Other	0 (0)	52 (5.3)
Marital status		
Married	37 (56.9)	477 (48.9)
Divorced	7 (10.8)	84 (8.6)
Legally separated	0 (0)	21 (2.2)
Widowed	6 (9.2)	156 (16.0)
Significant other	1 (1.5)	5 (0.5)
Single	12 (18.5)	189 (19.4)
Country of origin		
US	9 (13.8)	856 (87.7)
Laos	50 (76.9)	1 (0.1)
Other	4 (6.2)	78 (8.0)
Deceased as of October 21, 2018	10 (15.4)	218 (22.3)
Medicaid coverage, ever	46 (70.8)	279 (28.6)
Medicare coverage, ever	31 (47.7)	760 (77.9)
Care encounter characteristics		
Principal diagnosis of gout, any encounter	10 (15.4)	67 (6.9)
First encounter type		
Elective	5 (7.7)	237 (24.3)
Urgent care	4 (6.2)	190 (19.5)
Emergency room	56 (86.2)	547 (56.0)
Trauma care	0	2 (0.2)
First admission type		
Inpatient	59 (90.8)	826 (84.6)
Outpatient, observation	6 (9.2)	150 (15.4)
First length of stay, means ± SD days	8.8 ± 11.7	5.2 ± 5.4
Two or more encounters	16 (24.6)	196 (20.1)
Medication use		
Oral glucocorticoids		
Before first encounter	13 (20)	305 (31.2)
During first encounter	27 (41.5)	244 (25.0)
Injectable glucocorticoids		
Before first encounter	4 (6.2)	191 (19.6)
During first encounter	23 (35.4)	343 (35.1)
Nonsteroidal antiinflammatory drugs		
Before first encounter	5 (7.7)	239 (24.5)
During first encounter	18 (27.7)	217 (22.2)
Gout-related medications		
Before first encounter	21 (32.3)	568 (58.2)
During first encounter	47 (72.3)	700 (71.7)
Anakinra for inflammatory disease like RA		
Before first encounter	0 (0)	1 (0.1)
During first encounter	9 (13.8)	10 (1.0)

* Values are the number (%) unless indicated otherwise. RA = rheumatoid arthritis.

Table 2. Associations of comorbidities and clinical outcomes by Hmong status, unadjusted and adjusted (age, sex, admission type, principal diagnosis of gout)*

	Hmong	Non-Hmong	Unadjusted OR (95% CI)	Adjusted OR (95% CI)	β coefficient unadjusted (95% CI)	β coefficient (95% CI)
Presence of comorbidity†						
Diabetes mellitus (all types)	21 (32.3)	369 (37.8)	0.79 (0.46, 1.34)	0.97 (0.55, 1.67)	–	–
Hypertension	47 (72.3)	777 (79.6)	0.67 (0.38, 1.18)	0.95 (0.53,1.75)	–	–
Ischemic heart disease	15 (23.1)	300 (30.7)	0.68 (0.37, 1.22)	0.93 (0.48, 1.70)	–	–
Heart failure	9 (13.8)	223 (22.8)	0.54 (0.26, 1.11)	0.70 (0.31, 1.42)	–	–
Chronic kidney disease	34 (52.3)	415 (42.5)	1.48 (0.90, 2.45)	2.33 (1.36, 4.04)‡	–	–
End-stage renal disease	10 (15.4)	84 (8.6)	1.93 (0.95, 3.93)	1.76 (0.71, 3.58)	–	–
Mortality	10 (15.4)	218 (22.4)	0.63 (0.32, 1.26)	1.01 (0.46, 2.03)	–	–
Creatinine level, mean ± SD mg/dl§	3.3 ± 3.3	2.0 ± 1.8	–	–	1.37 (0.89, 1.85)	1.35 (0.85, 1.84)¶
GFR, mean ± SD ml/minute§	44.8 ± 18.8	49.3 ± 15.6	–	–	–4.56 (–8.60, 0.53)	–6.95 (–10.95, –2.96)¶

* Values are the number (%) unless indicated otherwise. 95% CI = 95% confidence interval; GFR = glomerular filtration rate; OR = odds ratio.
† Diabetes (ICD-9 250.x; ICD-10 E10.x, E11.x, E13.x), hypertension (ICD-9 401.x-405.x; ICD-10 I10.x-I16.x), ischemic heart disease (ICD-9 410.x-414.x; ICD-10 I20.x-I25.x), heart failure (ICD-9 428.x; ICD-10 I50.x), chronic kidney disease (ICD-9 585.x; ICD-10 N18.x), and end-stage renal disease (ICD-9 585.5x, 585.6x, V42.0x, V45.1x; ICD-10 N18.5x, N18.6x, Z94.0x, Z99.2x).
‡ $P < 0.05$.
§ During a 4-year period, creatinine level based on the highest value, and GFR based on the lowest value.
¶ $P < 0.001$.

(Table 2), but these differences did not meet the level of statistical significance. Additionally, after adjustment for age, sex, encounter type, and gout as a principal diagnosis, comorbidities and deaths were similar between the 2 groups (Table 2). Chronic kidney disease was, however, higher in Hmong patients, and after adjustment, the odds ratio (OR) was 2.33 (95% confidence interval [95% CI] 1.36, 4.04; $P < 0.05$). Importantly, after adjustment, creatinine ratios were significantly higher, and glomerular filtration rates (GFRs) were significantly lower in the Hmong compared to non-Hmong patients (Table 2). A higher percentage of Hmong patients carried a diagnosis of end-stage renal disease (15.4% of Hmong versus 8.6% of non-Hmong; OR 1.76 [95% CI 0.71, 3.58]), but this finding did not reach statistical significance.

DISCUSSION

Our data confirm that Hmong patients with gout visiting a large Midwestern regional tertiary care system are more than a decade younger than non-Hmong patients and yet had similar rates of diabetes mellitus, hypertension, heart disease, and mortality. These comorbidities have been associated with the development of gout in other populations (6). Quite notably, Hmong patients in our study had significantly higher creatinine level and lower GFR at time of presentation. There was also a clear trend toward higher rates of chronic and end-stage renal disease in Hmong versus non-Hmong patients, although this did not reach a level of statistical significance. Hmong patients were much less likely to interact with the health care system in the elective or ambulatory care setting than non-Hmong patients, instead presenting directly to the emergency room. Hmong patients were also significantly less likely to be using medication therapy appropriate for the management of gout, including oral nonsteroidal antiinflammatory drugs, gout-specific medications, and oral glucocorticoids. There was a clear trend for Hmong patients to have a prolonged length of stay during admission.

The reasons for these findings are likely multifactorial, involving both physiologic as well as cultural differences. Multiple studies have previously shown that there are differences in uric acid serum concentration among different ethnic groups. Hmong have higher serum uric acid concentrations (3) and are at higher risk of uric acid stones than control populations (7). Past studies have shown that the Hmong population of Minnesota develop gout much earlier than the non-Hmong population (37.4 versus 55.0 years), have higher uric acid concentrations, and develop more tophaceous gout, with an OR of 4.3; these findings persisted when controlling for age, sex, hypertension, diuretic use, and kidney function (6). Specific alleles that have been associated with increased risk of hyperuricemia are found more frequently in Hmong individuals than individuals of European or Han Chinese descent (3). The specific polymorphisms found in the Hmong people also provide a basis for anecdotal reports suggesting that allopurinol may not be effective when used in Hmong patients, as these alleles code for uric acid transporters in the proximal convoluted tubule, and therefore uric acid levels may not decrease with xanthine oxidase blockade (3).

High dietary purine intake (meat, seafood, eggs), soft drinks, and alcohol use have been strongly associated with increased risk of gout (1,2). The traditional Hmong diet consists of white rice–based dishes with the addition of steamed vegetables and meats/proteins such as chicken, pork, beef, and eggs, which could confer some increased risk depending on quantity consumed. Significantly, the use of tobacco and alcohol in the Hmong population has been characterized as lower than the average non-Hmong individual in prior studies (8). Similarly, sugary beverages were consumed daily by only 10% of Hmong surveyed, less than the average of >30% for Americans in general (9). While traditional Hmong cultural standards value an overweight body type as a sign of health and strength and find being thin to be a sign of weakness and frailty, Hmong people are more likely to be overweight than non-Hmong individuals but less likely to be obese (8). Anecdotally, Hmong individuals in the US report increasing their meat intake due to the greater availability and affordability as compared to their countries of origin (8). Overall, it seems that while diet may play a role, it is not likely a major contributor.

We found that many Hmong patients used the emergency department very frequently rather than utilizing preventative care. This may be due in part to the way that Hmong culture approaches illness differently from Western society. In traditional Hmong teaching, illness is due to either natural causes (exposure to elements, accidents/injuries, bad food) or supernatural causes (offended spirits, malevolent spirits, lost souls). Illnesses due to natural causes are traditionally treated with cupping, massage, herbs, or other nonspiritual methods (10). Illness due to supernatural causes are considered to be more serious and are treated with the help of a shaman (10). A survey done in 2004 showed that 74% of Hmong people identify shamanism as their primary religion (8), and in 2002, another survey showed that 75% of Hmong people used shamans for healing (11). Additionally, in traditional Hmong culture, there is a belief that suffering is a part of life instead of an abnormal occurrence that requires rectifying (10). There is also a belief in the Hmong community that one's lifespan is predetermined, and therefore interventions from the Western medical community may not be considered impactful (10). These factors have been postulated to influence Hmong patients' decisions about whether or not to seek medical attention (10). These considerations in combination with other factors, such as language barrier, lack of access to transportation, financial constraints, and decreased understanding of the health insurance and medical systems, may help to explain why more Hmong patients in our cohort had their initial point of contact with the health care system in the emergency department.

The concept of chronic illness is not present in traditional Hmong culture, and this lack of framework can have important

implications for patient understanding and compliance with treatment (12). Additionally, the Hmong language lacks terms that translate anatomic and physiologic terminology, requiring indirect translation and nonspecific words used to approximate meaning when discussing a Hmong patient's illness with them in their native language (13). There are no Hmong words to describe chronic illnesses such as diabetes mellitus or hypertension (12). Prior studies have shown considerable confusion about the difference between curing and controlling illness, and many Hmong people felt that medication only needed to be taken when one was feeling unwell, or if there are other tangible signs of illness (12). Many Hmong patients who seek care from US physicians also seek care from Hmong herbalists and shamans for the same complaint and will often use all treatment modalities together; a prior study found that 65% of Hmong people polled reported the use of traditional herbal medicines with illness (14). One survey showed that a majority of Hmong patients stated that traditional herbal medicine was easier to get than Western medicine, and that they believed herbal medicine to be just as or more effective than Western medicine because "herbal medicine treats the whole body" not just a specific health problem (14). Further, patients expected Western medicines to have an immediate and powerful effect, and if this was not observed, medications were felt to be unhelpful (10).

Additionally, medication side effects from treatments for conditions that are largely asymptomatic, such as diabetes mellitus and hypertension, were not well accepted and were frequently reported to cause medication nonadherence. Further, due to cultural communication styles, Hmong patients would consider it rude to tell a health care provider that the medications were not being taken as prescribed (13). Some Hmong individuals may be afraid of hospital experiences and mistrust Western medical providers. There is an association between Western medicine and death, as many patients reported knowing people who sought care from a hospital and then died (10). Hmong individuals have indicated fear of multiple blood tests because there is a belief that the body has a limited supply of blood and this might be permanently depleted (11). Some fear surgery due to a concern that this will anger the soul and cause it to vacate the body (10). Concerns that doctors experiment on Hmong people also circulate; there is a perception that the student doctors such as medical students or residents who are there to learn are practicing on Hmong people and that the treatments these doctors offer are not for the benefit of the Hmong patients, but for the doctor who needs to learn and practice skills (13). Further complicating this picture, many Hmong report frank mistrust and lack of confidence in the American health care system, which may stem from concerns that their traditional healing methods are not acknowledged or discussed (11).

Last, members of the Hmong community often make health care decisions as a unit headed by a male family member, which can complicate and delay medical decision-making. The concepts of autonomy and the preference to honor the rights of the individual are not emphasized as they are in Western society (11). All of these factors combined can lead to difficulties with communication, patient education, establishing trust, and formulating and adhering to a treatment plan for management of chronic disease.

It is possible that a language barrier and differences in a cultural framework in which to discuss disease in general, and particularly chronic disease, may negatively impact patients' understanding of their condition; these communication issues might lead to a longer hospital stay, as we observed in our cohort admitted at Regions Hospital. Regions is committed to minimizing these communication barriers, supporting a very robust in-person medical translator department of native Hmong speakers who are always available. Further, the community-based Hmong Gout Coalition (https://www.facebook.com/HmongGoutCoalition) sends representatives to help with education and provides culturally tailored Hmong gout videos (https://www.youtube.com/channel/UC6hhCpJxXdVvW6lOCl7roiA). Prior work has shown that Hmong patients receiving dialysis treatment would have liked to hear from members of their own community regarding the process of receiving treatment (15). A possible future direction of study would be to formally evaluate education interventions such as these as a means to improve medication compliance and trust among Hmong patients. Increased use of preventative medications may result in fewer hospitalizations and would decrease burden on the health care system, especially given the longer expected years of disease involvement in Hmong patients secondary to younger age of onset.

Finally, while this study does have one of the largest populations of Hmong patients admitted to a tertiary care system, the number is still relatively small, which limits evaluation of more subtle findings and determining significant differences with non-Hmong patients. Eleven non-Hmong patients had a last name included on the aforementioned list of traditional Hmong clan last names but did not meet our case definition for Hmong (see the methods section for this definition). Post-analysis medical record review found that 8 of these patients were clearly not of Hmong ancestry. The remaining 3 patients did not meet our strict definition of a Hmong patient but could possibly have been Hmong based on their course of disease being similar to that seen in Hmong patients and their place of residence in a Hmong neighborhood. However, we did not include these patients in our analysis as they did not meet our strict definition. The exclusion of these patients further reduced our sample size and as such is listed as a possible limitation.

In conclusion, this study extends previous findings that a diagnosis of gout in Hmong patients indicates earlier and often more severe disease. While there is not an increase in comorbid conditions, tertiary care system encounters in this population are often longer. Intensive efforts at education, increasing compliance with preventative medications, and identifying the genetic

abnormalities responsible for this increased burden of gout are currently underway.

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

Study conception and design. Lerman, Gertner, Desai.
Acquisition of data. Lerman, Gertner, DeFor, Brown, Desai.
Analysis and interpretation of data. Lerman, Gertner, DeFor, Brown, Desai.

REFERENCES

1. Chen-Xu M, Yokose C, Rai SK, Pillinger MH, Choi HK. Contemporary prevalence of gout and hyperuricemia in the United States and decadal trends: the National Health and Nutrition Examination Survey 2007–2016. Arthritis Rheumatol 2019;71:991–9.
2. Choi H, Mount DB, Reginato AM, American College of Physicians, American Physiological Society. Pathogenesis of gout. Ann Intern Med 2005;143:499–516.
3. Roman YM, Culhane-Pera KA, Menk J, Straka RJ. Assessment of genetic polymorphisms associated with hyperuricemia or gout in the Hmong. Per Med 2016;13:429–40.
4. Portis AJ, Laliberte M, Tatman P, Moua M, Culhane-Pera K, Maalouf NM, et al. High prevalence of gouty arthritis among the Hmong population in Minnesota. Arthritis Care Res (Hoboken) 2010;62:1386–91.
5. Juraschek SP, Miller ER III, Gelber AC. Body mass index, obesity, and prevalent gout in the United States in 1988–1994 and 2007–2010. Arthritis Care Res (Hoboken) 2013;65:127–32.
6. Wahedduddin S, Singh JA, Culhane-Pera KA, Gertner E. Gout in the Hmong in the United States: a retrospective study. J Clin Rheumatol 2010;16:262–6.
7. Portis AJ, Hermans K, Culhane-Pera KA, Curhan GC. Stone disease in the Hmong of Minnesota: initial description of a high-risk population. J Endourol 2004;18:853–7.
8. Yang RC, Mills PK. Dietary and lifestyle practices of Hmong in California. J Health Care Poor Underserved 2008;19:1258–69.
9. Centers for Disease Control and Prevention. National Health and Nutrition Examination Survey, 2009–2010. URL: https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx?BeginYear=2009.
10. Uba L. Cultural barriers to health care for Southeast Asian refugees. Public Health Rep 1992;107:544–8.
11. Cobb TG. Strategies for providing cultural competent health care for Hmong Americans. J Cult Divers 2010;17:79–83.
12. Helsel D, Mochel M, Bauer R. Chronic illness and Hmong shamans. J Transcult Nurs 2005;16:150–4.
13. Johnson SK. Hmong health beliefs and experiences in the Western health care system. J Transcult Nurs 2002;13:126–32.
14. Lor KB, Moua S, Ip EJ. Frequency and perceptions of herbal medicine use among Hmong Americans: a cross sectional survey. J Immigr Minor Health 2016;18:397–401.
15. Kruger L. Experiences of Hmong patients on hemodialysis and the nurses working with them. Nephrol Nurs J 2009;36:379–87.

Arthritis Care & Research
Vol. 74, No. 3, March 2022, pp 468–477
DOI 10.1002/acr.24498

American College of Rheumatology
Empowering Rheumatology Professionals

Factors Associated With Treatment Response in Patients With Idiopathic Inflammatory Myopathies: A Registry-Based Study

Fabricio Espinosa-Ortega, Marie Holmqvist, Maryam Dastmalchi, Ingrid E. Lundberg, and Helene Alexanderson

Objective. To identify predictors of response to immunosuppressive therapy after 1 year, with a focus on autoantibodies, in patients newly diagnosed with idiopathic inflammatory myopathies (IIM) followed longitudinally in an electronic registry.

Methods. We assessed the association between autoantibody-defined groups and improvement according to American College of Rheumatology/European Alliance of Associations for Rheumatology 2016 response criteria.

Results. We identified 156 patients; of those, 111 (71%) were positive for any autoantibody tested, 90% received glucocorticoid treatment at baseline, and 78% received immunosuppressive drugs at some follow-up point. After 1 year from the index date, the overall median improvement score was 27.5 (interquartile range 10–51). No differences were observed in the total improvement score between the autoantibody-defined groups. Overall, 62% of patients (n = 96) showed a minimal response, 38% (n = 60) achieved a moderate response, and 19% (n = 30) achieved a major response. Regarding the different levels of response, dermatomyositis-specific autoantibodies were associated with a moderate response versus the seronegative group (reference), odds ratio 4.12 (95% confidence interval 1.2–16.5). In addition, dysphagia, time from symptom onset to diagnosis, and initial glucocorticoid dose were significant predictors of response after 1 year of follow-up.

Conclusion. Patients with DM-specific autoantibodies achieved better levels of response compared to other autoantibody-defined groups. Dysphagia, a shorter time span from symptom onset to diagnosis, and intensive initial immunosuppressive treatment were associated with a higher response rate after 1 year of pharmacologic treatment from the index date, regardless of autoantibody status.

INTRODUCTION

Idiopathic inflammatory myopathies (IIM) are a group of complex systemic disorders whose main symptoms are muscle weakness, low muscle endurance, and inflammatory infiltrates in muscle tissue biopsies (1). Extramuscular involvement, such as skin rash, arthritis, dysphagia, interstitial lung disease, cardiac disease, and malignancy, are common. Many of these diverse manifestations have been linked to the presence of specific autoantibodies, so-called myositis-specific autoantibodies (MSAs), which are mainly found in patients with IIM, and myositis-associated autoantibodies (MAAs), which are also present in other autoimmune disorders (2,3). The autoantibody profile of each patient often corresponds to a specific clinical phenotype. The frequency of the various clinical manifestations and autoantibodies varies according to both ethnic and genetic background (4). Whether autoantibody status has an impact on treatment response and outcomes has not been studied in detail.

Glucocorticoids are regarded as a first-line therapy in combination with an additional immunosuppressive drug, such as methotrexate, azathioprine, mycophenolate, cyclosporine, or tacrolimus. New biologic drugs have emerged as an alternative for treating patients with refractory disease (5,6), and exercise is an important part of nonmedical treatment (7,8). Despite intense

Supported by the Swedish Research Council (K2014-52X-14045-14-3), the Swedish Rheumatism Association, the King Gustav V's 80-Year Foundation, and the Stockholm County Council (ALF Project).

Fabricio Espinosa-Ortega, MD, Marie Holmqvist, MD, PhD, Maryam Dastmalchi, MD, PhD, Ingrid E. Lundberg, MD, PhD, Helene Alexanderson, PhD: Solna, Karolinska Institutet, and Karolinska University Hospital, Stockholm, Sweden.

Dr. Lundberg has received research grants from Bristol Myers Squibb and AstraZeneca and is a scientific advisor for aTyr and Corbus Pharmaceuticals. No other disclosures relevant to this article were reported.

Address correspondence to Fabricio Espinosa-Ortega, MD, D2:01 Karolinska University Hospital, Solna SE 171 76, Stockholm, Sweden. Email: fabricio.espinosa@ki.se.

Submitted for publication December 26, 2019; accepted in revised form October 22, 2020.

SIGNIFICANCE & INNOVATIONS

- Dermatomyositis-specific autoantibodies were associated with a moderate response after 1 year of pharmacologic treatment from the index date.
- The presence of dysphagia at the index date, a shorter time span from symptom onset to diagnosis, and more-intensive initial glucocorticoid treatment were predictors of response, regardless of autoantibody status.

treatment, many patients have persistent signs of systemic disease activity and do not regain muscle performance. To date, no biomarkers have been identified that predict response to treatment, other than those biomarkers for biologic drugs (9,10). One limitation in addressing this question has been the lack of international consensus as to how to assess improvement after treatment. In 2016 the American College of Rheumatology (ACR)/European Alliance of Associations for Rheumatology (EULAR) proposed response criteria that define improvement in terms of both muscular and nonmuscular measurements, which have since been widely accepted (11). MSAs are an attractive option to test as potential biomarkers for treatment response and outcomes due to their association with distinct clinical phenotypes. Only a few studies have taken this approach so far, and they have been limited to patients with established, treatment-refractory disease (9,12,13). Thus, no information is available regarding MSAs as biomarkers for treatment response in patients newly diagnosed with IIM.

The present study aimed to test the potential of autoantibodies, as well as other clinical features, as predictors of treatment response, applying the new ACR/EULAR response criteria after 1 year of immunosuppressive therapy in a cohort with recent-onset IIM that had been followed longitudinally in an electronic health care registry.

MATERIALS AND METHODS

Study population. Since 2003, patients with IIM have been included and followed in a standardized way using the electronic Swedish Quality of Care Registry, which has a myositis-specific module, SweMyoNet. This registry prospectively collects demographic, clinical, serologic, and treatment data during inpatient and outpatient visits to the rheumatology clinic. Patients with a primary diagnosis of IIM are classified as having dermatomyositis (DM), polymyositis, amyopathic DM, inclusion body myositis, antisynthetase syndrome (ASS), or juvenile DM. For this study, we retrospectively selected patients who fulfilled the EULAR/ACR 2017 classification criteria for definite or probable IIM in any of the above-mentioned subsets, as well as patients who met the criteria for ASS (14,15). All included patients were followed at Karolinska University Hospital and were registered in SweMyoNet within 12 months (range 0.2–11.3 months) of diagnosis between January 1, 2003, and December 31, 2015. Patients with inclusion body myositis and juvenile DM were excluded from this study (16,17). The date of inclusion to SweMyoNet was defined as the index date.

Treatment. Information on pharmacologic treatment was available from the SweMyoNet registry. Treatment of individual patients was based on the treating physician's decision and was in most cases started with high-dose glucocorticoids (0.75–1 mg/kg/day prednisolone, but not >80 mg) for 4–6 weeks, in combination with azathioprine (1.5–2 mg/kg/day), methotrexate (15–20 mg/day), or mycophenolate mofetil (2–2.5 grams/day). Glucocorticoids were tapered approximately every 3 to 4 weeks according to the treating physician's decision, based on the response to treatment (Vårdprogram myosit, Karolinska Universitetssjukhuset [in Swedish]). The use of glucocorticoids, methotrexate, azathioprine, mycophenolate mofetil, cyclophosphamide, and any biologic agent, either abatacept or rituximab (during follow-up), was recorded as dichotomous variables. The glucocorticoid dose at the index date was recorded as a continuous variable.

Autoantibodies. Two assays were applied for autoantibody specificities: RNA- and protein-immunoprecipitation or line blot (Euroline Myositis Antigen Profile 4 [Euroimmun]) as described elsewhere (18). Seventy patients were tested by line blot and 86 patients by immunoprecipitation. 3-hydroxy-3-methylglutaryl-coenzyme A reductase autoantibodies were analyzed at the US National Institutes of Health using a combined protocol of immunoprecipitation followed by an enzyme-linked immunosorbent assay (19).

Autoantibody-defined subgroups. Patients were categorized by the presence of autoantibodies, as follows: 1) ASS (Jo1, PL7, PL12, EJ), 2) DM-associated autoantibodies (MDA5, anti–transcription intermediary factor 1γ [anti-TIF1-γ], Mi2, SAE), 3) autoantibodies associated with immune-mediated necrotizing myopathy (IMNM: SRP and 3-hydroxy-3-methylglutaryl-coenzyme A reductase), 4) MAA without any MSA (PmScl, U1 RNP, Ro52, Ku), and 5) seronegative (negative to any of these autoantibodies).

Comorbidities. Any malignancy within ±3 years of IIM diagnosis was defined as myositis-associated cancer and recorded as such. Interstitial lung disease, cardiac involvement, and dysphagia, as defined elsewhere, were recorded as dichotomous variables (20).

Definition of treatment response. We applied the 2016 ACR/EULAR criteria improvement score to assess clinical response to treatment (11). In short, the International Myositis Assessment and Clinical Studies group 6-item core set measures

of disease activity, all included in the SweMyoNet module, were used: patient global assessment (PtGA) of disease activity and physician global assessment (PhGA) of disease activity, both scored on a 10-cm visual analog scale (VAS); the Manual Muscle Test in 8 muscle groups (MMT8); the Health Assessment Questionnaire (HAQ); levels of the serum muscle enzyme creatine phosphokinase; and global extramuscular disease activity based on the physician's evaluation on a 10-cm VAS, including 6 organ systems (MYOACT tool) (21). Active disease was defined as a value of ≥1.5 on the PhGA (22). According to ACR/EULAR response criteria, the absolute percent change for each core set domain is calculated (final value − baseline value/range × 100) (11). An improvement score is assigned to each measure based on this absolute change, and each individual core set measure is weighted such that those considered more important contribute more to the final score (11). Improvement scores for each of the 6 core set domains are summed to establish a total improvement score. The higher the change, the higher the improvement score. If the patient had <5% improvement or worsened on a particular domain, a score of 0 was assigned to that domain (11). The response thresholds were 20–39 for minimal, 40–59 for moderate, and 60–100 for a major response.

Statistical analysis. Descriptive statistical analyses were performed. The Kruskal-Wallis test was used for continuous variables with >2 independent samples, and Wilcoxon's rank test was used for comparisons of 2 independent samples. Chi-square test or Fisher's exact test were used for categorical variables, when appropriate. A linear regression model was used to test the association between the autoantibody status and the total improvement score adjusted by the baseline values for each core set measure. We also included an interaction parameter to identify whether the effect of the initial values for each core set measure varied among the autoantibody groups. A logistic regression model was used to test the association between autoantibody-defined groups and potential clinical predictors for each category of response, using the nonresponders as the reference category. To test for sensitivity, we excluded those patients who died during the observation period and checked for differences in the proportion of patients meeting the improvement criteria. A value of *P* less than 0.05 was defined as statistically significant. The statistical package employed was R, version 3.5.0 (23).

Ethics. Ethical approval was granted by the Regional Ethics Review Board of the Karolinska Institutet, Stockholm, Sweden. All patients signed an informed consent form before their data were included in the registry.

RESULTS

Patients. A total of 156 cases were identified (see Supplementary Figure 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24498/abstract). Sixty-two patients (40%) had DM, 8 (5%) had amyopathic DM, and 86 (55%) had polymyositis. Of all 156 cases, 39 (25%) met the criteria for ASS. The baseline demographic characteristics across each autoantibody group are summarized in Table 1. In all, 69% were female (n = 107) with a mean ± SD age of 57 ± 14 years. The median duration of symptoms prior to diagnosis was 3 months (IQR 1.0, 8.5). All patients had active disease at the index date: one-third exhibited severe organ involvement (i.e., lung involvement or dysphagia), 6% had cardiac involvement, and 17% had a myositis-associated malignancy. Ninety percent of patients (n = 140) were given glucocorticoids at baseline. The median initial daily dose was 50 mg (IQR 25, 60 mg); 28 patients (18%) started on intravenous pulse steroids (>125 mg methylprednisolone). Besides glucocorticoids, 39% (n = 61) received methotrexate, 18% (n = 29) azathioprine, 22% (n = 35) mycophenolate mofetil, 20% (n = 32) cyclophosphamide, and 1% (n = 2) IVIg. Twelve percent (n = 18) received a biologic drug (abatacept or rituximab) within the first year. In total, 111 patients (71%) were positive to any autoantibody. The number of patients with ASS, DM-specific autoantibodies, IMNM-associated autoantibodies, and MAAs was 39 (25%), 28 (18%), 9 (6%), and 35 (22%), respectively. The number of patients negative for any autoantibody (seronegative) was 45 (29%). For information about each group, see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24498/abstract.

Patients with ASS were more likely to have interstitial lung disease (67%) compared to the rest of the groups (overall $P < 0.001$). Cardiac involvement was more frequent in the IMNM group compared to other groups (overall $P = 0.04$), and the DM-specific autoantibody group had a higher erythrocyte sedimentation rate (ESR) at the index date ($P = 0.008$). A significantly higher number of patients with ASS were treated with mycophenolate mofetil, cyclophosphamide, and biologic drugs during the follow-up compared to the rest of the groups (overall $P < 0.005$, $P < 0.001$, and $P < 0.005$, respectively). Regarding treatment patterns over time, more patients included in SweMyoNet during the registry's first years were given glucocorticoids than those added to the registry in later years, while the latter were more likely to be given biologic drugs (see Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24498/abstract). No differences were observed between the autoantibody-defined groups in terms of dysphagia, cancer, or duration of symptoms prior to diagnosis.

Autoantibodies and treatment response. Table 2 summarizes the levels of response across the autoantibody-defined groups. After 1 year from the index date, the overall median improvement score was 27.5 (IQR 10, 51), with no significant differences among the autoantibody-defined groups. Table 3 and Figure 1 summarize the values for each core set

Table 1. Baseline demographic characteristics of 156 patients with idiopathic inflammatory myopathies by autoantibody group*

	Total	ASS (n = 39)	DM specific (n = 28)	IMNM (n = 9)	MAA (n = 35)	None (n = 45)	P†
Age at diagnosis, mean ± SD years	57 ± 14	54 ± 15	54 ± 16	58 ± 19	58 ± 12	61 ± 13	NS
Female, no. (%)	107 (69)	24 (62)	22 (78)	6 (67)	26 (74)	26 (58)	NS
Diagnosis, no. (%) or no.							
Amyopathic dermatomyositis	8 (5)	3	3	1	0	2	NS
Dermatomyositis	62 (40)	13	23	0	7	18	
Polymyositis	86 (55)	23	2	8	28	25	
Symptom duration before diagnosis, months	3.0 (1.0, 8.5)	3 (1, 9)	2.2 (0, 5.5)	3.2 (1, 12)	3.9 (0, 8.9)	3.5 (1, 8)	NS
Disease duration at index date from diagnosis, months	0.98 (0.2, 2.6)	0.7 (0.2, 2.2)	0.9 (0.1, 1.6)	2.6 (0.7, 7.0)	1.2 (0.5, 3.7)	1.2 (0.1, 3.5)	NS
ESR, mm/hour	20 (12, 31)	22 (16.5, 34)	29 (20, 47)	15 (9, 19)	15.5 (8, 28)	16 (10.5, 26)	0.008
Comorbidities, no. (%)							
ILD	52 (34)	26 (67)	10 (36)	1 (11)	11 (31)	4 (9)	<0.001
Dysphagia	57 (36)	8 (21)	11 (39)	3 (33)	18 (51)	17 (38)	NS
Cancer	26 (17)	5 (13)	8 (29)	2 (22)	3 (9)	8 (18)	NS
Cardiac involvement	9 (6)	1 (3)	1 (4)	3 (33)	2 (6)	2 (4)	0.04
Glucocorticoids, no. (%)	140 (90)	36 (92)	26 (93)	7 (78)	33 (94)	38 (84)	NS
Initial glucocorticoid dose, mg	50 (25, 60)	50 (25, 60)	52 (32, 50)	25 (12, 50)	45 (30, 60)	40 (30, 60)	NS
Methotrexate, no. (%)	61 (39)	11 (28)	8 (28)	4 (44)	19 (54)	19 (42)	NS
Azathioprine, no. (%)	29 (18)	8 (20)	4 (14)	2 (22)	6 (17)	9 (20)	NS
Mycophenolate mofetil, no. (%)	35 (22)	11 (28)	6 (21)	2 (22)	11 (31)	5 (11)	0.004
Cyclophosphamide, no. (%)	32 (20)	19 (48)	7 (25)	0 (0)	5 (14)	1 (0.2)	<0.001
Biologic drug, no. (%)‡	18 (12)	10 (26)	5 (18)	0 (0)	2 (6)	1 (2)	0.005

* Values are the median (interquartile range) unless indicated otherwise. ASS = antisynthetase syndrome group; DM specific = dermatomyositis-specific autoantibodies group; ESR = erythrocyte sedimentation rate; ILD = interstitial lung disease; IMNM = immune-mediated necrotizing myopathy autoantibodies group; MAA = myositis-associated autoantibodies group; NS = not significant.
† *P* value by chi-square test/Fisher's exact test for categorical data and by Kruskal-Wallis test for continuous data. $P < 0.05$ indicates a significant difference between the 5 groups.
‡ Use of a biologic drug (abatacept or rituximab) during the follow-up, i.e., 1 year after the index date.

measure and the absolute percent change at 1 year after the index date across the autoantibody-defined groups. We did not find significant differences in the absolute percent change in any core set measure among the autoantibody groups. At the index date, patients with ASS had a higher MMT8 score ($P = 0.01$), and patients with DM-specific autoantibodies had a higher extramuscular score (overall $P = 0.002$). At 1 year of follow-up, patients with ASS continued to have higher MMT8 scores ($P = 0.038$), and patients with IMNM-associated antibodies had persistently higher creatine phosphokinase levels ($P = 0.001$).

Because initial values for each core set measure at baseline are associated with their values after a period of follow-up, and subsequently with the total improvement score, we tested whether the various autoantibody groups had any effect on the total improvement score that was independent of the initial values for each core set measure and whether there was an effect from the autoantibody groups on the total improvement score depending on the initial values for each core set measure (interaction variable) (see Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24498/abstract). The unadjusted linear regression analysis showed that the DM-specific autoantibodies group was associated with a higher total improvement score compared with the seronegative group (reference group). After adjusting for the initial

Table 2. Total improvement score and number of patients achieving minimal, moderate, and major response after 1 year of treatment between autoantibody-defined groups*

	Total (n = 156)	ASS (n = 39)	DM specific (n = 28)	IMNM (n = 9)	MAA (n = 35)	None (n = 45)	P†
TIS, median (IQR)	27.5 (10, 51)	28 (13, 48)	48 (11, 63)	7.5 (2.5, 35)	35 (11, 58)	21 (7.5, 42)	0.07
Minimal	96 (62)	27 (69)	20 (71)	3 (33)	23 (66)	23 (51)	0.14
Moderate	60 (38)	13 (33)	17 (61)	2 (22)	16 (46)	12 (27)	0.03‡
Major	30 (19)	7 (18)	10 (36)	0 (0)	7 (20)	6 (13)	0.12

* Values are in the number (%) unless indicated otherwise. ASS = antisynthetase syndrome group; DM specific = dermatomyositis-specific autoantibodies group; IMNM = immune-mediated necrotizing myopathy autoantibodies group; IQR = interquartile range; MAA = myositis-associated autoantibodies group; None = patients negative to any antibody; TIS = total improvement score.
† *P* value by chi-square test/Fisher's exact test for categorical data and by Kruskal-Wallis test for continuous data. $P < 0.05$ indicates a significant difference between the 5 autoantibody-defined groups.
‡ Statistically significant.

Table 3. Individual core set measures at index date and follow-up, and absolute percentage change in 156 patients with idiopathic inflammatory myopathies and by autoantibody group*

	Total (n = 156)	ASS (n = 39)	DM specific (n = 28)	IMNM (n = 9)	MAA (n = 35)	Seronegative (n = 45)	P†
PhGA							
Index date	40 (20, 57)	40 (21, 59)	50 (35, 69)	20 (20, 39)	43 (29, 52)	31 (12, 60)	0.18
Follow-up	12 (3, 20)	10 (0, 20)	18 (10, 33)	20 (15, 20)	16 (5, 20)	10 (0, 20)	0.1
Absolute % change	−21(−40, 0)	−22 (−50, 3)	−30 (−45, −5)	0 (0, 0)	−28 (−40, −3)	−20 (−32, 0)	0.12
PtGA							
Index date	44 (25, 71)	47 (25, 61)	50 (32, 75)	35 (17, 46)	35 (24, 69)	46 (28, 71)	0.6
Follow-up	30 (10, 50)	22 (4, 44)	34 (8, 47)	24 (15, 41)	25 (12, 51)	42 (14, 65)	0.5
Absolute % change	−9 (−32, 3)	−12 (−39, 1.0)	−18 (−41, 0)	−3 (−19, 0.5)	−9.5 (−26.8, 4.5)	−2 (−24, 12)	0.29
MMT8							
Index date	73 (63, 78)	78 (69, 78)	69 (57, 78)	76 (68, 78)	71 (63, 76)	71 (63, 78)	0.01
Follow-up	78 (72, 80)	80 (78, 80)	79 (69, 80)	78 (74, 79)	76 (72, 79)	76 (69, 80)	0.038
Absolute % change	3 (0, 13)	1.3 (0, 8.5)	3.2 (0, 14)	2.5 (0.7, 8.2)	3.8 (0, 12.5)	1.3 (−1.6, 11.6)	0.64
HAQ							
Index date	0.8 (0.3, 0.4)	0.6 (0.2, 1.3)	1.0 (0.3, 1.5)	0.6 (0.3, 1.3)	0.8 (0.3, 1.3)	0.9 (0.3, 1.5)	0.7
Follow-up	0.5 (0, 1)	0.2 (0, 0.7)	0.2 (0, 1.7)	0.5 (0.1, 0.8)	0.5 (0.1, 1)	0.8 (0.1, 1.2)	0.19
Absolute % change	0 (−17, 0)	−4 (−21, 0)	−8 (−24, 0)	−4 (−10, −1)	−8 (−23, 3)	0 (−6, 4)	0.37
CK							
Index date	5.2 (1.5, 24)	5.9 (1.5, 5.2)	2.3 (1.2, 8.9)	12.5 (7.4, 9.5)	13.4 (1.6, 40)	4.1 (1.7, 26)	0.12
Follow-up	1.5 (1, 3.6)	1.6 (1.1, 2.1)	1.1 (0.8, 1.8)	7.6 (4.7, 10)	1.4 (0.8, 3.9)	1.7 (1.1, 5)	0.001
Absolute % change	−5 (−32, 0.2)	−5 (−31, 1)	−3 (−16, 0.1)	−8 (−28, −3)	−23 (−49, −1)	−2 (−22, 0.8)	0.27
EM							
Index date	28 (10, 43)	30 (10, 45)	43 (27, 55)	18 (4, 20)	25 (15, 40)	15 (0, 30)	0.002
Follow-up	10 (0, 18)	9.5 (0, 16)	12 (5, 27)	16 (10, 23)	10 (0, 16)	6 (0, 10)	0.08
Absolute % change	−11 (−32, 0)	−15 (−41, 0)	−21 (−37, 0)	0 (0, 0)	−9 (−32, −2)	−3 (−22, 0)	0.12

* Values are the median (interquartile range). ASS = antisynthetase syndrome group; CK = creatine kinase levels; DM specific = dermatomyositis-specific autoantibodies group; EM = extramuscular assessment; HAQ = Health Assessment Questionnaire; IMNM = immune-mediated necrotizing myopathy autoantibodies group; MAA = myositis-associated autoantibodies; MMT8 = Manual Muscle Test in 8 muscle groups; PhGA = physician global assessment; PtGA = patient global assessment.
† *P* value by Kruskal-Wallis test. $P < 0.05$ indicates a significant difference between the 5 autoantibody-defined groups.

values for each individual core set measure separately, the ASS group was associated with a higher total improvement score after adjusting for the initial MMT8 score, but it was associated with a lower improvement score after adjusting for the initial PtGA value ($P = 0.006$ and $P = 0.03$, respectively). The DM-specific autoantibodies group was associated with a lower total improvement score after adjusting for the initial value of PtGA ($P = 0.01$). With respect to this effect of each autoantibody group on the total improvement score depending on the initial values for each core set measure (interaction parameter), we found that besides the independent effect of the ASS group and initial value of the initial MMT8 score on the total improvement score, an even higher total improvement score is expected for the ASS antibody group compared with the reference group, i.e., the lower the initial MMT8 score, the higher the total improvement score, and even higher for the ASS antibody group than for the reference group ($P = 0.01$). Similarly, the effect of a low HAQ score on the total improvement score was higher in the MAA group compared with the reference group. No other significant interactions were observed (see Supplementary Table 3, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24498/abstract).

Of the 156 patients initially identified, 96 (62%) met the criteria for minimal response, 60 (38%) met the criteria for moderate response, and 30 (19%) met the criteria for major response. Regarding the effect of autoantibody status on the different levels of response, only patients with DM-specific autoantibodies were associated with the moderate response level ($\chi^2 = 10.4$, df = 4, $P = 0.034$). No significant associations between the autoantibody-defined groups and minimal or major responses were observed.

Characteristics of nonresponders and responders. The characteristics of nonresponders (improvement score < 20) are summarized in Table 4. Nonresponder patients had a longer duration of symptoms prior to diagnosis and lower disease activity as measured both by PhGA and a lower ESR at the index date. They also received lower mean initial glucocorticoid doses, and significantly fewer patients received cyclophosphamide and biologic drugs than responding patients at each level of response. The nonresponders were less likely to have gastrointestinal involvement, represented by dysphagia, than patients who achieved moderate or major responses. When comparing the autoantibody-defined groups, no differences were found in the number of nonresponders and responding patients at any level of response.

Predictive factors for treatment response. As a next step, we performed a univariate logistic regression analysis to test the predictive value of the antibody-defined groups for each category of response. Table 5 summarizes the results of the

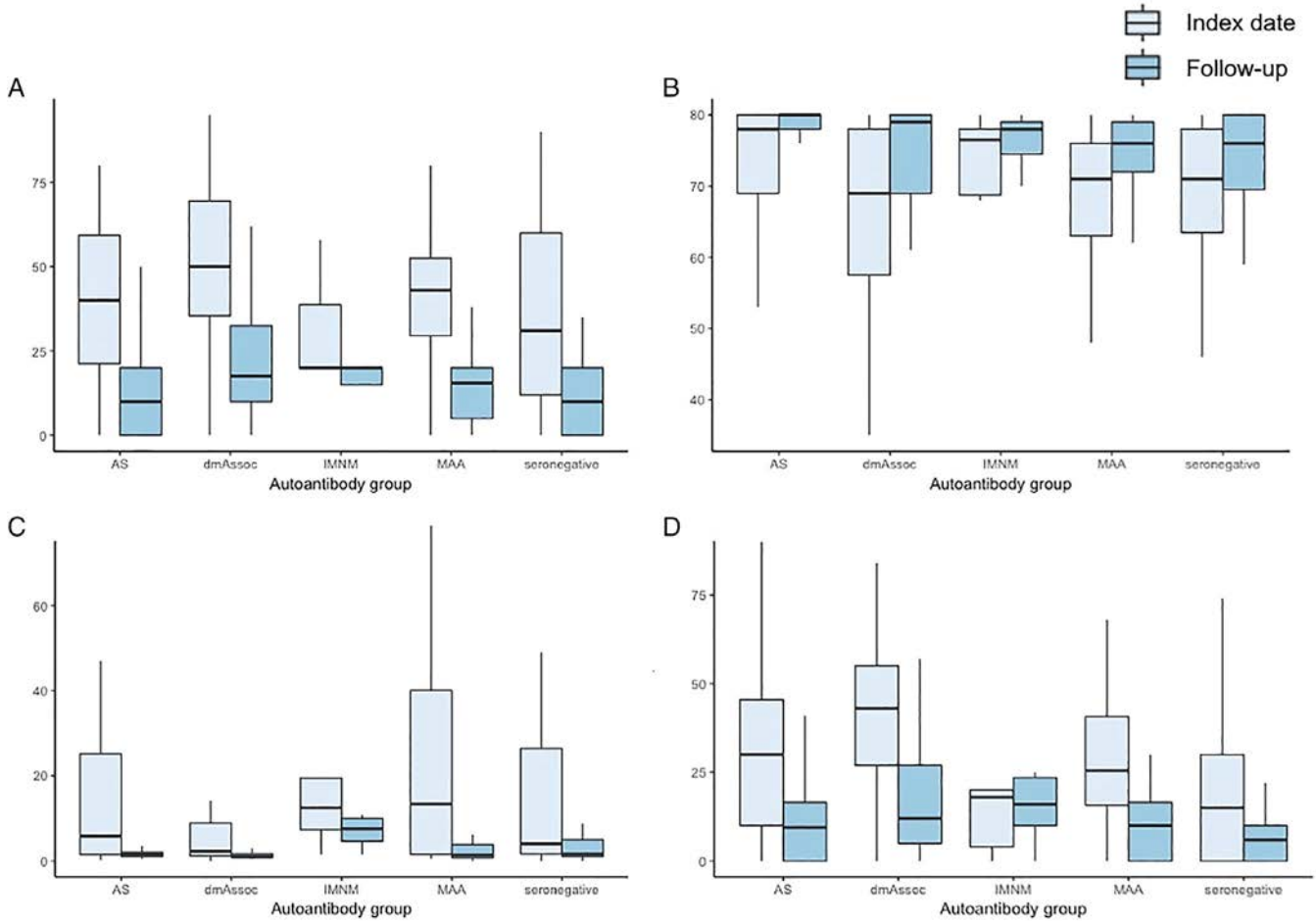


Figure 1. Values of core set measures at baseline and at 1 year after being treated. **A**, Physician global assessment; **B**, Manual Muscle Test in 8 muscle groups score; **C**, Creatine kinase levels, and **D**, Extramuscular visual analog scale. Boxes represent interquartile range. Horizontal lines represent median value. Whiskers show the range. AS = antisynthetase syndrome group; dmAssoc = dermatomyositis-specific autoantibodies group; IMNM = immune-mediated necrotizing myopathy autoantibodies group; MAA = myositis-associated autoantibodies group.

logistic regression models for each category of response. The multivariate logistic regression analysis demonstrated that the DM-specific autoantibody group was associated with moderate response (odds ratio [OR] 4.2 [95% confidence interval (95% CI) 1.2–16.5]). The DM-specific autoantibody group was also associated with minimal and major responses to a degree that did not reach significance. Two independent predictive factors were associated with response to treatment: time from first symptoms to diagnosis (OR 0.86 [95% CI 0.7–0.96] for major response) and dysphagia (OR 3.02 [95% CI 1.3–7.7] for minimal response and OR 3.2 [95% CI 1.2–9.5] for major response). Moreover, because these associations could reflect confounding by indication, we tested the dose of glucocorticoids per se with all 3 levels of response. An increase of 1 milligram of initial glucocorticoid dose was associated with up to a 4% increase in the odds of achieving a response. Finally, as a sensitivity analysis, 9 patients (6%) who died during the observation period were excluded: the associations between predictive factors and levels of response remained similar (see Supplementary Table 4, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24498/abstract).

DISCUSSION

In our single-center cohort of patients newly diagnosed with IIM, we found no significant differences in the total improvement score between the autoantibody-defined groups following the ACR/EULAR 2016 criteria. We did, however, observe that patients who were positive for DM-specific autoantibodies had a higher frequency of achieving a moderate response than the other groups. To our knowledge, this is the first study to analyze the usefulness of autoantibody status as a predictor for treatment response using the ACR/EULAR 2016 criteria.

Earlier studies have shown that DM-specific autoantibodies are markers of response in patients with established or refractory disease after treatment with rituximab, as well as markers of long-

Table 4. Comparison of characteristics between nonresponders and patients achieving minimal, moderate, and major response*

	Nonresponders (ref.) (n = 60)	Minimal (n = 96)	Moderate (n = 60)	Major (n = 30)
Age at diagnosis, mean ± SD years	56 ± 15	58 ± 15	58 ± 14	59 ± 14
Female, no. (%)	38 (64)	66 (68)	46 (76)	21 (70)
Dermatomyositis phenotype, no. (%)	25 (42)	41 (42)	31 (51)	20 (66)
Duration of symptoms, months	5.2 (1.3, 13)	2.2 (0.4, 6.9)†	2.0 (0.0, 4.0)‡	1.0 (0.0, 2.8)§
Physician global assessment	20 (8, 40)	50 (37, 66)†	55 (40, 69) ‡	64 (50, 71)§
Interstitial lung disease, no. (%)	14 (23)	38 (39)	23 (38)	13 (43)
Dysphagia, no. (%)	15 (25)	41 (42)	28 (46)‡	16 (53)£
Cancer, no. (%)	8 (13)	18 (18)	14 (23)	9 (30)
ESR at baseline, mm/hour	14 (6, 26)	22 (15, 34)†	26.5 (20, 41)‡	29 (18, 42)§
ASS, no. (%)	12 (20)	27 (28)	13 (21)	7 (23)
DM specific, no. (%)	8 (13)	20 (20)	17 (28)	10 (33)
IMNM, no. (%)	6 (10)	3 (3)	2 (3)	0 (0)
MAA, no. (%)	12 (20)	23 (20)	16 (26)	7 (23)
Seronegative, no. (%)	21 (35)	23 (24)	12 (20)	6 (20)
Initial glucocorticoid dose, mg	30 (11, 44)	50 (39, 60)†	60 (50, 60)‡	60 (56, 70)§
Methotrexate, no. (%)	20 (33)	41 (43)	29 (43)	13 (43)
Azathioprine, no. (%)	10 (17)	19 (20)	9 (15)	3 (10)
Mycophenolate, no. (%)	14 (24)	21 (22)	14 (23)	9 (30)
Cyclophosphamide, no. (%)	6 (10)	26 (27)†	17 (28)‡	9 (30)§
Biologic drug, no. (%)	4 (7)	13 (13)	10 (16)	5 (16)

* Values are the median (interquartile range) unless indicated otherwise. ASS = antisynthetase syndrome group; DM specific = dermatomyositis-specific autoantibodies group; ESR = erythrocyte sedimentation rate; IMNM = immune-mediated necrotizing myopathy autoantibodies group; MAA = myositis-associated autoantibodies; ref. = reference.
† Comparisons between minimal responders and nonresponders by Wilcoxon's rank test with a *P* value <0.05.
‡ Comparisons between moderate responders and nonresponders by Wilcoxon's rank test with a *P* value <0.05.
§ Comparisons between major responders and nonresponders by Wilcoxon's rank test with a *P* value <0.05.

term remission (9,10,24,25). Similarly, our study suggests that DM-specific autoantibodies are markers of good response after conventional immunosuppressive treatment in patients within the early stages of the disease. There are several reasons that explain these findings. First, compared to the other autoantibody-defined groups, patients who harbored DM-specific autoantibodies had both higher levels of extramuscular involvement and the highest levels of ESR, as measures of disease activity at baseline. Although we were not able to retrieve information regarding the specific extramuscular organs involved, cutaneous manifestations and lung disease in MDA5-positive cases are frequent in these patients. This fact is important, given that one aspect of the ACR/EULAR criteria is that the extramuscular activity, together with the physician's assessment, is the second most

Table 5. Factors associated with clinical response in patients with idiopathic inflammatory myopathies*

	Univariate model			Multivariate model		
	Minimal odds	Moderate odds	Major odds	Minimal odds	Moderate odds	Major odds
Seronegative (ref.)	1.0	1.0	1.0	1.0	1.0	1.0
ASS	2.05 (0.84–5.16)	1.33 (0.5–3.4)	1.38 (0.4–4.7)	2.3 (0.6–8.5)	0.95 (0.26–3.3)	1.6 (0.4–7.3)
DM specific	2.28 (0.85–6.5)	4.12 (1.5–11.6)†	3.5 (1.13–11.8)†	3.9 (0.99–18.3)	4.2 (1.2–16.5)†	3.01 (0.7–13)
IMNM	0.45 (0.1–1.96)	0.76 (0.1–3.7)	0.014 (0.01–2.2)	0.6 (0.1–3.6)	1.19 (0.1–7.4)	2.8 (0.8–5.3)
MAA	1.75 (0.7–4.45)	2.24 (0.9–5.9)	1.6 (0.4–5.4)	1.3 (0.4–3.9)	2.13 (0.7–6.6)	1.25 (0.3–5.2)
Initial GC dose	1.05 (1.03–1.07)†	1.04 (1.02–1.06)†	1.05 (1.03–1.08)†	1.04 (1.02–1.07)†	1.04 (1.02–1.07)†	1.04 (1.01–1.07)†
Time from 1st symptoms to diagnosis, months	0.98 (0.95–1.00)	0.98 (0.94–1.00)	0.86 (0.75–0.95)	0.97 (0.95–1.0)	0.99 (0.96–1.01)	0.86 (0.7–0.96)†
Dysphagia	2.22 (1.1–4.64)	2.1 (1.05–4.0)	2.4 (1.1–5.4)	3.02 (1.3–7.7)†	2.1 (0.9–5.1)	3.2 (1.2–9.5)†
Initial ESR	1.02 (1.0–1.04)	1.04 (1.01–1.06)	1.03 (1.01–1.05)	1.01 (0.98–1.03)	1.03 (1.0–1.05)	1.01 (0.99–1.04)
Use of CFM	3.3 (1.3–9.2)	2.05 (0.9–4.6)	1.9 (0.7–4.5)	1.1 (0.28–4.6)	1.23 (0.38 – 4.15)	6.0 (1.6 – 2.1)

* Values are the odds ratio (95% confidence interval). ASS = antisynthetase syndrome group; CFM = cyclophosphamide; DM specific = dermatomyositis specific autoantibodies group; ESR = erythrocyte sedimentation rate; GC = glucocorticoid; IMNM = immune-mediated necrotizing myopathy autoantibodies; MAA = myositis-associated autoantibodies group; ref. = reference.
† Statistically significant association.

important weighted contributor to the improvement score, after the MMT8 score.

Furthermore, patients with DM-specific autoantibodies had a trend toward higher absolute percentage change in PtGA than the other groups, which could be an effect of improvement in cutaneous signs, which usually correlates with better responses in subjective outcome measures (26). In addition, in patients with anti-TIF1-γ and anti-Mi2 autoantibodies, but not in patients without these specificities, some molecular pathways, such as the interferon signature, are predictors of response to treatment as measured by improvement in muscle strength and PhGA (27).

The autoantibody-defined groups exhibited notable differences in the core set measures at baseline and at follow-up, but not in the absolute percentage change. Patients with ASS usually exhibit a high level of extramuscular disease activity, represented by a high prevalence of interstitial lung disease, and a low level of muscle involvement (15). Indeed, in our study, patients with ASS autoantibodies presented the highest mean MMT8 scores, both at baseline and at follow-up, and a higher frequency of lung involvement (although they did not have a higher mean extramuscular score compared with the other autoantibody-defined groups). Interestingly, when we analyzed the interaction between the autoantibody groups and the initial value of each core set measure, we found that the ASS group and the DM-specific antibodies group were associated with higher total improvement scores than the reference group, an association that was independent of the baseline values for MMT8 and PtGA. Moreover, the total improvement score was higher when initial values for MMT8 were lower at baseline, and even higher in the ASS group compared with the reference group. Together, these findings indicate that the ACR/EULAR response criteria can capture the nature of response in the different autoantibody-defined groups.

In addition to DM-specific autoantibodies, we found other independent factors associated with different levels of response to treatment. The presence of dysphagia was strongly associated with minimal and major responses. In previous reports, dysphagia has been associated with a good response, probably due to more intensive treatment in patients with higher global disease activity and anti-TIF1-γ autoantibodies (28,29). Our findings, however, showed that dysphagia was a predictor independent of the initial dose of glucocorticoids. In this study, time from onset of symptoms to diagnosis and initial glucocorticoid dose were also independent factors associated with response to treatment.

Due to the long observation time of our cohort (>10 years), some concerns may arise about the differences in treatment patterns over the years. In fact, after performing an additional analysis, we found that the use of glucocorticoids was more frequent during the first years of the registry, whereas the use of biologic drugs was more frequent in later years (see Supplementary Table 2, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24498/abstract). However, neither of these differences was associated with any level of response. Nonetheless, our data support the importance of early initial treatment intervention to achieve improvement in patients with IIM according to the response criteria (30–33).

Our study has several limitations. First, autoantibodies were tested using 2 different assays. However, as previously reported, the overall concordance between these 2 assays was 78%, with a moderate agreement. Moreover, the agreement for the most prevalent specificities (e.g., Jo1) was good (18). Second, seronegative patients, particularly those without skin rash, may develop another myopathy beyond the observation period of this study. Third, the small cohort size prevented analyses of individual autoantibodies as predictors, and thus we grouped patients into clinically relevant autoantibody-defined subgroups. Still, the small sample size of autoantibody-defined groups might have limited the ability to detect differences in total improvement scores. Fourth, patients considered to be nonresponders might represent a group with mild disease, irrespective of autoantibody status (22). Fifth, the association of medications with treatment responses could be a potential confounder for indication influencing the degree of response to treatment. Lastly, we cannot rule out the possibility that occasional patients might have received aggressive treatment before the index date or experienced improvement before their inclusion in the SweMyoNet registry. A strength of the SweMyoNet registry is that it includes most patients treated at our clinic, both patients diagnosed in the inpatient ward and those diagnosed in the outpatient clinic, and thus represents various levels of disease severity.

In conclusion, in our retrospective study using prospectively collected data, we found that patients with DM-specific autoantibodies were more likely to have a moderate level of response compared to patients without these autoantibodies. Moreover, the presence of dysphagia, a shorter time from symptom onset to diagnosis, and more-intensive initial glucocorticoid treatment were independently associated with higher rates of clinical improvement after 1 year of pharmacologic treatment, for all subgroups. Our findings highlight the importance of identifying autoantibody-defined subgroups of patients with IIM early on, and of initiating intensive glucocorticoid treatment as soon as possible after diagnosis, as this identification and treatment predict higher rates of clinical response regardless of autoantibody status.

ACKNOWLEDGMENTS

We thank Dr. Andrew Mammen of the NIH for anti-HMGCR autoantibody analysis and Dr. Tsuneyo Mimori, Kyoto, Japan, for performing the immunoprecipitation assays.

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. Espinosa-Ortega 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. Espinosa-Ortega, Holmqvist, Dastmalchi, Lundberg, Alexanderson.

Acquisition of data. Espinosa-Ortega, Holmqvist, Dastmalchi, Lundberg, Alexanderson.

Analysis and interpretation of data. Espinosa-Ortega, Holmqvist, Dastmalchi, Lundberg, Alexanderson.

REFERENCES

1. Zong M, Lundberg IE. Pathogenesis, classification and treatment of inflammatory myopathies. Nat Rev Rheumatol 2011;7: 297–306.
2. Betteridge Z, McHugh N. Myositis-specific autoantibodies: an important tool to support diagnosis of myositis. J Inten Med 2016;280: 8–23.
3. Love LA, Leff RL, Fraser DD, Targoff IN, Dalakas M, Plotz PH, et al. A new approach to the classification of idiopathic inflammatory myopathy: myositis-specific autoantibodies define useful homogeneous patient groups. Medicine (Baltimore) 1991;70:360–74.
4. Rothwell S, Cooper RG, Lundberg IE, Miller FW, Gregersen PK, Bowes J, et al. Dense genotyping of immune-related loci in idiopathic inflammatory myopathies confirms HLA alleles as the strongest genetic risk factor and suggests different genetic background for major clinical subgroups. Ann Rheum Dis 2016;75:1558–66.
5. Tjarnlund A, Tang Q, Wick C, Dastmalchi M, Mann H, Studynkova JT, et al. Abatacept in the treatment of adult dermatomyositis and polymyositis: a randomised, phase IIb treatment delayed-start trial. Ann Rheum Dis 2018;77:55–62.
6. Oddis CV, Reed AM, Aggarwal R, Rider LG, Ascherman DP, Levesque MC, et al. Rituximab in the treatment of refractory adult and juvenile dermatomyositis and adult polymyositis: a randomized, placebo-phase trial. Arthritis Rheumatol 2013;65:314–24.
7. Boehler JF, Hogarth MW, Barberio MD, Novak JS, Ghimbovschi S, Brown KJ, et al. Effect of endurance exercise on microRNAs in myositis skeletal muscle: a randomized controlled study. PloS One 2017; 12:17.
8. Alexanderson H. Physical exercise as a treatment for adult and juvenile myositis. J Intern Med 2016;280:75–96.
9. Aggarwal R, Bandos A, Reed AM, Ascherman DP, Barohn RJ, Feldman BM, et al. Predictors of clinical improvement in rituximab-treated refractory adult and juvenile dermatomyositis and adult polymyositis. Arthritis Rheumatol 2014;66:740–9.
10. Aggarwal R, Oddis CV, Goudeau D, Koontz D, Qi ZB, Reed AM, et al. Autoantibody levels in myositis patients correlate with clinical response during B cell depletion with rituximab. Rheumatology (Oxford) 2016;55:991–9.
11. Aggarwal R, Rider LG, Ruperto N, Bayat N, Erman B, Feldman BM, et al. 2016 American College of Rheumatology/European League Against Rheumatism criteria for minimal, moderate, and major clinical response in adult dermatomyositis and polymyositis: an International Myositis Assessment and Clinical Studies Group/Paediatric Rheumatology International Trials Organisation collaborative initiative. Arthritis Rheumatol 2017;69:898–910.
12. Aggarwal R, Cassidy E, Fertig N, Koontz DC, Lucas M, Ascherman DP, et al. Patients with non-Jo-1 anti-tRNA-synthetase autoantibodies have worse survival than Jo-1 positive patients. Ann Rheum Dis 2014;73:227–32.
13. Vermaak E, Tansley SL, McHugh NJ. The evidence for immunotherapy in dermatomyositis and polymyositis: a systematic review. Clin Rheumatol 2015;34:2089–95.
14. Lundberg IE, Tjarnlund A, Bottai M, Werth VP, Pilkington C, de Visser M, et al. 2017 European League Against Rheumatism/American College of Rheumatology classification criteria for adult and juvenile idiopathic inflammatory myopathies and their major subgroups. Ann Rheum Dis 2017;76:1955–64.
15. Connors GR, Christopher-Stine L, Oddis CV, Danoff SK. Interstitial lung disease associated with the idiopathic inflammatory myopathies: what progress has been made in the past 35 years? Chest 2010;138: 1464–74.
16. Bohan A, Peter JB. Polymyositis and dermatomyositis (first of two parts). N Engl J Med 1975;292:344–7.
17. Griggs RC, Askanas V, Dimauro S, Engel A, Karpati G, Mendell JR, et al. Inclusion-body myositis and myopathies. Ann Neurol 1995;38: 705–13.
18. Espinosa-Ortega F, Holmqvist M, Alexanderson H, Storfors H, Mimori T, Lundberg IE, et al. Comparison of autoantibody specificities tested by a line blot assay and immunoprecipitation-based algorithm in patients with idiopathic inflammatory myopathies. Ann Rheum Dis 2019;78:858–60.
19. Mammen AL, Chung T, Christopher-Stine L, Rosen P, Rosen A, Doering KR, et al. Autoantibodies against 3-hydroxy-3-methylglutaryl-coenzyme A reductase in patients with statin-associated autoimmune myopathy. Arthritis Rheum 2011;63:713–21.
20. Lilleker JB, Vencovsky J, Wang GC, Wedderburn LR, Diederichsen LP, Schmidt J, et al. The EuroMyositis registry: an international collaborative tool to facilitate myositis research. Ann Rheum Dis 2018;77:30–9.
21. Rider LG, Feldman BM, Perez MD, Rennebohm RM, Lindsley CB, Zemel LS, et al. Development of validated disease activity and damage indices for the juvenile idiopathic inflammatory myopathies. I. Physician, parent, and patient global assessments. Arthritis Rheum 1997;40:1976–83.
22. Rider LG, Giannini EH, Harris-Love M, Joe G, Isenberg D, Pilkington C, et al. Defining clinical improvement in adult and juvenile myositis. J Rheumatol 2003;30:603–17.
23. The R Foundation. The R project for statistical computing. 2018. URL: https://www.r-project.org/.
24. Hengstman GJ, Egberts W, Seelig HP, Lundberg IE, Moutsopoulos HM, Doria A, et al. Clinical characteristics of patients with myositis and autoantibodies to different fragments of the Mi-2 beta antigen. Ann Rheum Dis 2006;65:242–5.
25. Mariampillai K, Granger B, Amelin D, Guiguet M, Hachulla E, Maurier F, et al. Development of a new classification system for idiopathic inflammatory myopathies based on clinical manifestations and myositis-specific autoantibodies. JAMA Neurol 2018;75: 1528–37.
26. Robinson ES, Feng R, Okawa J, Werth VP. Improvement in the cutaneous disease activity of patients with dermatomyositis is associated with a better quality of life. Br J Dermatol 2015;172: 169–74.
27. Reed AM, Crowson CS, Hein M, de Padilla CL, Olazagasti JM, Aggarwal R, et al. Biologic predictors of clinical improvement in rituximab-treated refractory myositis. BMC Musculoskelet Disord 2015;16:8.
28. Ogawa-Momohara M, Muro Y, Kono M, Akiyama M. Prognosis of dysphagia in dermatomyositis. Clin Exp Rheumatol 2019;37:165.
29. Mugii N, Hasegawa M, Matsushita T, Hamaguchi Y, Oohata S, Okita H, et al. Oropharyngeal dysphagia in dermatomyositis: associations with clinical and laboratory features including autoantibodies. PloS One 2016;11:e0154746.

30. Joffe MM, Love LA, Leff RL, Fraser DD, Targoff IN, Hicks JE, et al. Drug-therapy of the idiopathic inflammatory myopathies: predictors of response to prednisone, azathioprine, and methotrexate and a comparison of their efficacy. Am J Med 1993;94:379–87.
31. Marie I, Hachulla E, Hatron PY, Hellot MF, Levesque H, Devulder B. Polymyositis and dermatomyositis: Short term and longterm outcome, and predictive factors of prognosis. J Rheumatol 2001;28:2230–7.
32. Fafalak RG, Peterson MG, Kagen LJ. Strength in polymyositis and dermatomyositis: best outcome in patients treated early. J Rheumatol 1994;21:643–8.
33. Rider LG, Lachenbruch PA, Monroe JB, Ravelli A, Cabalar I, Feldman BM, et al. Damage extent and predictors in adult and juvenile dermatomyositis and polymyositis as determined with the Myositis Damage Index. Arthritis Rheum 2009;60:3425–35.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 3, March 2022, pp 478–483
DOI 10.1002/acr.24646

BRIEF REPORT

Prognosis and Treatment of Myositis-Associated Severe Interstitial Lung Disease: A Descriptive Study Using a Nationwide Inpatient Database in Japan

Hiroki Furuya,[1] Mikio Nakajima,[2] Kei Ikeda,[1] Kaito Nakamura,[1] Hiroyuki Ohbe,[3] Shotaro Aso,[3] Ryosuke Kumazawa,[3] Taro Iwamoto,[1] Arifumi Iwata,[1] Shunsuke Furuta,[1] Hiroki Matsui,[3] Kiyohide Fushimi,[4] Hideo Yasunaga,[3] and Hiroshi Nakajima[1]

Objective. The aim of this study was to determine the prognosis, clinical course, and current management of severe interstitial lung disease (ILD) associated with myositis in Japan.

Methods. We conducted a retrospective descriptive study using a nationwide database for inpatient care of acute illness in Japan. Among a total of ~66 million inpatient admissions, we identified patients with severe ILD associated with polymyositis (PM) or dermatomyositis (DM) who required mechanical ventilation and methylprednisolone pulse therapy (≥1 gm/day of methylprednisolone) from July 2010 to March 2018.

Results. We identified 155 patients with PM and 394 with DM who fulfilled the above criteria. The median age of patients was 65 years; DM patients were significantly younger than PM patients (64 versus 68 years; $P < 0.001$). The numbers of patients who were treated with calcineurin inhibitors, intravenous cyclophosphamide, and polymyxin B–immobilized fiber column direct hemoperfusion (PMX-DHP) were 403 (73.4%), 318 (57.9%), and 78 (14.2%), respectively. All these treatments were given significantly more frequently to the patients with DM compared with those with PM. The uses of other treatment options were much less frequent. The median periods after hospitalization when methylprednisolone pulse therapy, calcineurin inhibitors, mechanical ventilation, intravenous cyclophosphamide, and PMX-DHP were initiated and in-hospital death occurred among patients with DM were 2, 4, 7, 8, 17, and 36 days, respectively. In-hospital mortality was significantly higher in patients with DM than in those with PM (76.6% versus 56.8%; $P < 0.001$).

Conclusion. The mortality of patients with myositis-associated severe ILD who require mechanical ventilation is extremely high despite aggressive and prompt interventions.

INTRODUCTION

Polymyositis (PM) and dermatomyositis (DM) are rare systemic autoimmune disorders affecting skeletal muscles and are estimated to affect 10–13 patients per 100,000 population in Japan (1). Interstitial lung disease (ILD) is a life-threatening complication of PM and DM that develops in as much as 50% of cases (2). A certain form of myositis-associated ILD is rapidly progressive and reported to lead 40% of the patients to death within 6–12 months after diagnosis (3).

Due to its rarity and high mortality, there are no established evidence-based treatment strategies for ILD with PM or DM. A study of the past largest cohort of new-onset myositis-associated ILD (n = 497) showed that the most frequently prescribed drugs were glucocorticoids (98%), cyclosporine (48%), tacrolimus (45%), and cyclophosphamide (44%) in

Supported by the Ministry of Health, Labour, and Welfare (grants 19AA2007 and 20AA2005) and the Ministry of Education, Culture, Sports, Science, and Technology (grant 20H03907).

[1]Hiroki Furuya, MD, PhD, Kei Ikeda, MD, PhD, Kaito Nakamura, MD, Taro Iwamoto, MD, PhD, Arifumi Iwata, MD, PhD, Shunsuke Furuta, MD, PhD, Hiroshi Nakajima, MD, PhD: Chiba University Hospital, Chiba, Japan; [2]Mikio Nakajima, MD, PhD: The University of Tokyo and Tokyo Metropolitan Hiroo Hospital, Tokyo, Japan; [3]Hiroyuki Ohbe, MD, PhD, Shotaro Aso, MD, MPH, PhD, Ryosuke Kumazawa, MPharm, Hiroki Matsui, MPH, Hideo Yasunaga, MD, PhD: The University of Tokyo, Tokyo, Japan; [4]Kiyohide Fushimi, MD, PhD: The University of Tokyo and Tokyo Medical and Dental University, Tokyo, Japan.

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

Address correspondence to Kei Ikeda, MD, PhD, Department of Allergy and Clinical Immunology, Chiba University Hospital, 1-8-1 Inohana, Chuo-ku, Chiba, Chiba 260-8677, Japan. Email: K.Ikeda@faculty.chiba-u.jp.

Submitted for publication December 28, 2020; accepted in revised form May 4, 2021.

SIGNIFICANCE & INNOVATIONS

- The largest number of patients with severe myositis-related interstitial lung disease (ILD) who require mechanical ventilation were identified by utilizing a nationwide database for inpatient care of acute illness in Japan.
- The mortality of severe myositis-related ILD in patients who require mechanical ventilation is extremely high despite aggressive and prompt therapeutic interventions.

Japan (4). However, there is limited information on patients with severe cases, which is a subpopulation that desperately needs a new treatment strategy.

Recently, real-world big data are attracting attention because they can provide unique and meaningful insight into rare and/or life-threatening diseases that conventional studies could not capture (5). In this study, we utilize a large real-world database in Japan to determine the prognosis, clinical course, and current management of severe ILD associated with myositis.

PATIENTS AND METHODS

Study design and data source. We conducted a retrospective descriptive study using the Diagnosis Procedure Combination (DPC) database, a nationwide administrative database for inpatient care of acute illness in Japan. Approximately 1,000 acute-care hospitals including all 82 academic hospitals in Japan participate in the database, and the data cover approximately one-half of all inpatient admissions to acute-care hospitals and around 92% of all admissions to tertiary-care emergency hospitals in Japan (6). The DPC database contains age, sex, dates of admission and discharge, content and date of given treatments, and diagnoses that are recorded in both Japanese text and International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10) codes. A previous study showed that sensitivity and specificity of the primary diagnoses were 78.9% and 93.2%, respectively (7). The current study was approved by the Institutional Review Board of The University of Tokyo. The requirement for informed consent was waived due to the anonymized nature of the data.

Patient selection. The data collection period for the study was from July 2010, the earliest date that the DPC data were available, to March 2018. We identified patients with severe ILD associated with PM or DM who required mechanical ventilation. We excluded patients who did not receive methylprednisolone pulse therapy within 30 days after admission in order to exclude patients who were hospitalized for infectious pneumonia as well as mild/nonacute cases. These diagnoses were defined by the following ICD-10 codes: ILD (J841, J849, J990, and J991), DM (M330, M331, and M339), and PM (M332). The methylprednisolone pulse therapy was defined as ≥1 gm/day of methylprednisolone. The patients with ICD-10 codes for both PM and DM were defined as having DM. Patients <16 years of age at admission were excluded (Figure 1).

Clinical information and outcomes. We investigated patients' sex, age, and date of administered medications and interventions and determined in-hospital mortality as an outcome. We extracted information on the following treatments available in Japan that had been reported to be effective for myositis-associated severe ILD (8–12): glucocorticoids; calcineurin inhibitors (CNIs), including tacrolimus and cyclosporine; cyclophosphamide (CYC); rituximab (RTX); methotrexate (MTX); azathioprine (AZA); mycophenolate mofetil (MMF); abatacept; tocilizumab; tumor necrosis factor (TNF) inhibitors, including infliximab, etanercept, adalimumab, and golimumab; JAK inhibitors, including tofacitinib and baricitinib; intravenous immunoglobulin

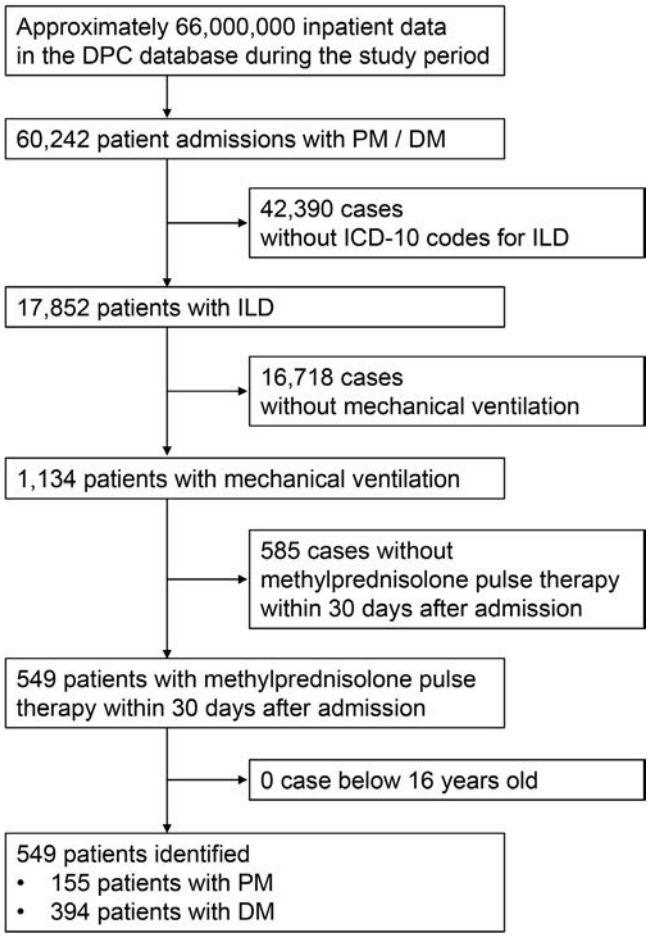


Figure 1. Patient selection flow chart. DM = dermatomyositis; DPC = diagnosis procedure combination; ICD-10 = International Statistical Classification of Diseases and Related Health Problems, Tenth Revision; ILD = interstitial lung disease; PM = polymyositis.

(IVIG); plasmapheresis; polymyxin B–immobilized fiber column direct hemoperfusion (PMX-DHP); and extracorporeal membrane oxygenation.

Statistical analysis. Continuous variables were reported as the median and interquartile range, and categorical variables were reported as count and percentage. Baseline differences between groups were evaluated using chi-square tests or Fisher's exact tests for categorical variables and Wilcoxon's rank sum tests for continuous variables. The threshold for significance was set to be a P value of 0.05. Statistical analyses were conducted using the statistical software package for Stata/MP, version 15.

RESULTS

Patient characteristics. Among a total of ~66 million inpatient admissions in the database during the study period, 155 patients with PM and 394 patients with DM fulfilled the study selection criteria (Figure 1 and Table 1). The median age in patients with DM was significantly younger than that in patients with PM (64 years versus 68 years; $P < 0.001$), while there was no significant difference in sex.

Treatment. The number of overall patients who were treated with CNIs, intravenous CYC, the combination of CNIs and intravenous CYC, and PMX-DHP was 403 (73.4%), 318 (57.9%), 274 (49.9%), and 78 (14.2%), respectively. All of these treatments were used more frequently in patients with DM than those with PM, while the other treatment options were used infrequently (Table 1). The time course of the treatment intervention and outcome after admission is shown in Figure 2. The median numbers of hospital days of initiating methylprednisolone pulse therapy, CNIs, mechanical ventilation, intravenous CYC, and PMX-DHP among patients with DM were 2, 4, 7, 8, and 17, respectively.

In-hospital mortality. The in-hospital mortality of patients with DM was significantly higher than that of patients with PM (76.6% versus 56.8%; $P < 0.001$). The median length of hospital stays among deceased patients with DM was 28 days (Figure 2). Notably, 66 of 78 patients (84.6%) who underwent PMX-DHP, and 221 of 274 patients (80.6%) who were administered a combination of intravenous CYC and CNIs died (see Supplementary Table 1, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.24646/abstract).

DISCUSSION

In the current study, we investigated the real-world practice and in-hospital mortality of 549 patients with myositis-associated

Table 1. Patients' characteristics given treatments and mortalities among patients with severe polymyositis (PM) and dermatomyositis (DM) who required mechanical ventilation and steroid pulse therapy*

Characteristic	Overall (n = 549)	PM (n = 155, 28.2%)	DM (n = 394, 71.8%)	*P*
Demographic information				
Age, median (IQR) years	65 (57–72)	68 (62–76)	64 (56–70)	<0.001
Male sex	256 (46.6)	76 (49.0)	180 (45.7)	0.48
Administered treatment				
Calcineurin inhibitors	403 (73.4)	90 (58.1)	313 (79.4)	<0.001
Intravenous cyclophosphamide	318 (57.9)	59 (38.1)	259 (65.7)	<0.001
Calcineurin inhibitors and intravenous cyclophosphamide	274 (49.9)	47 (30.3)	227 (57.6)	<0.001
PMX-DHP	78 (14.2)	14 (9.0)	64 (16.2)	0.029
Plasmapheresis	42 (7.7)	8 (5.2)	34 (8.6)	0.17
ECMO	15 (2.7)	2 (1.3)	13 (3.3)	0.19
Rituximab	15 (2.7)	4 (2.6)	11 (2.8)	0.89
Azathioprine	13 (2.4)	5 (3.2)	8 (2.0)	0.41
Oral cyclophosphamide	11 (2.0)	3 (1.9)	8 (2.0)	0.94
Mycophenolate mofetil	9 (1.6)	0 (0.0)	9 (2.3)	0.058
Intravenous immunoglobulin	5 (0.9)	0 (0.0)	5 (1.3)	0.16
Methotrexate	3 (0.5)	1 (0.6)	2 (0.5)	0.84
TNF inhibitors	3 (0.5)	0 (0.0)	3 (0.8)	0.28
JAK inhibitors	1 (0.2)	0 (0.0)	1 (0.3)	0.53
Abatacept	0 (0.0)	0 (0.0)	0 (0.0)	–
Tocilizumab	0 (0.0)	0 (0.0)	0 (0.0)	–
In-hospital mortality	390 (71.0)	88 (56.8)	302 (76.6)	<0.001

* Values are the number (%) unless indicated otherwise. ECMO = extracorporeal membrane oxygenation; IQR = interquartile range; PMX-DHP = polymyxin B–immobilized fiber column direct hemoperfusion; TNF = tumor necrosis factor.

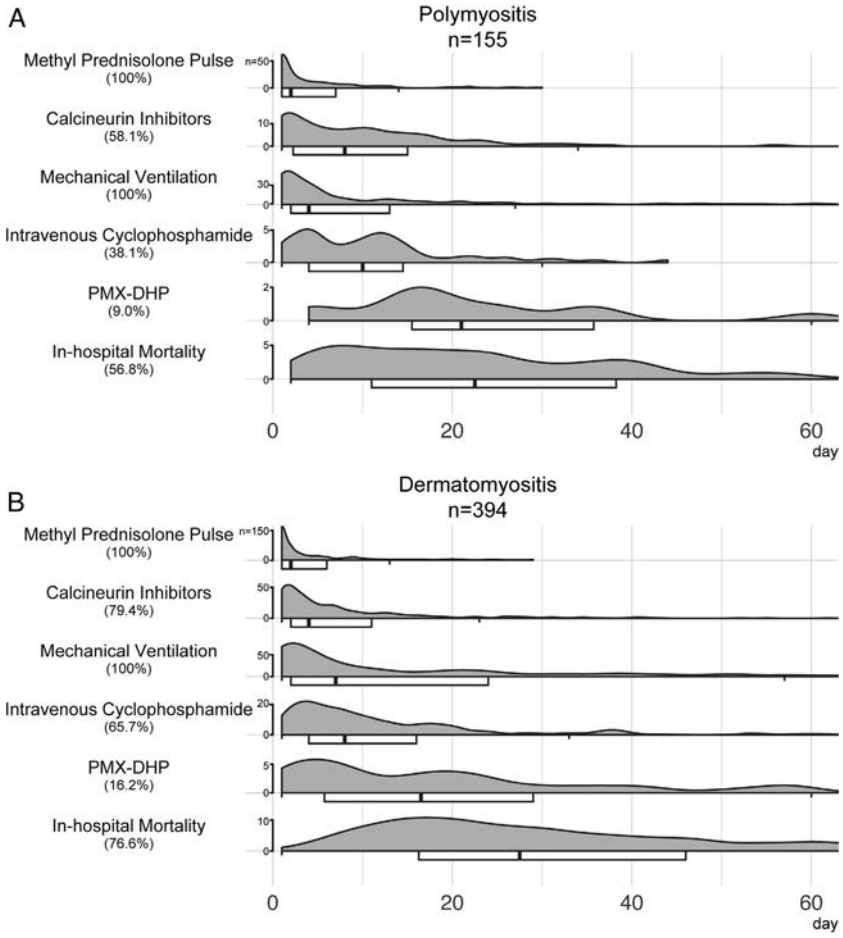


Figure 2. Time course of treatment and in-hospital death for polymyositis (**A**) and dermatomyositis (**B**). The numbers of patients (y-axis) who initiated major medications or therapeutic procedures and who died in each hospital day (x-axis) are shown in kernel density plots. The maximum heights of density plots are standardized to better visualize the time course of all treatments and in-hospital death. The box plots below represent median and lower and upper quantile hospital days. PMX-DHP = polymyxin B–immobilized fiber column direct hemoperfusion.

severe ILD who required mechanical ventilation and methylprednisolone pulse therapy using a nationwide administrative inpatient database in Japan. Although CNIs, intravenous CYC, and PMX-DHP were frequently and promptly used along with methylprednisolone pulse therapy, the in-hospital mortalities of patients with PM and DM were as high as 56.8% and 76.6%, respectively.

A previous study with the largest cohort of new-onset myositis-associated ILD (n = 497) showed that mortalities of patients with PM, clinically amyopathic DM, and classical DM were 2.6%, 23%, and 8%, respectively, during the median observation period of 20 months from diagnosis, most of which were due to progression of ILD (4). Similarly, a previous study in France showed that the overall mortality of myositis-associated ILD was 7.5% among 107 patients (13). However, unlike the current study, which focuses on the mortality of patients with severe cases of myositis-associated ILD requiring both mechanical ventilation and methylprednisolone pulse therapy, the mortalities reported in these previous studies were not stratified by severity. On the other hand, the subanalysis of our previously reported DPC data set (14) revealed that the mortality of acute exacerbation of idiopathic pulmonary fibrosis of similar severity was 57.3% (1,516 of 2,648 patients) (unpublished data). These data indicate that the diagnosis of underlying DM may be an important and independent poor prognostic factor.

Due to its rarity and high mortality, no formal guidelines or randomized controlled studies inform optimal therapy for severe myositis-associated ILD. In addition to high-dose glucocorticoids, which are regarded as first-line treatment, other treatment options include MTX, CNIs, intravenous CYC, RTX, AZA, MMF, IVIG, TNF inhibitors, abatacept, tocilizumab, JAK inhibitors, plasmapheresis, and PMX-DHP (8–12). Among these options, a few studies with a limited number of patients have been conducted to show the effectiveness of CNIs and intravenous CYC (8). Our data demonstrate that CNIs and intravenous CYC were indeed the drugs most commonly chosen in real-world settings for the treatment of severe myositis-associated ILD.

Notably, PMX-DHP was used in 14.2% of our study population. PMX-DHP has been developed and approved in Japan to remove endotoxin in patients with sepsis. However, it has also been used off-label to treat severe idiopathic pulmonary fibrosis, aiming to remove cytokines that contribute to lung injury, and has shown some effectiveness (15). Although the past reports of PMX-DHP in myositis-associated ILD are limited to case reports and case series (9), our data showed that a significant proportion of severe patients receive this treatment possibly due to the lack of better options. Meanwhile, the proportion of patients who underwent PMX-DHP was significantly higher in patients who died than in those who survived (see Supplementary Table 1, available at http://onlinelibrary.wiley.com/doi/10.1002/acr.24646/abstract). However, the impact of treatment with PMX-DHP or medications on mortality cannot be derived from our data due to confounding by indication.

To our knowledge, this is the first descriptive study of myositis-associated severe ILD to demonstrate clinical time course after admission. Our data showed that although methylprednisolone pulse therapy, CNIs, and intravenous CYC were promptly administered after hospitalization, mechanical ventilation was required mostly within a week, and 71.0% died mostly within 1 month after admission. The high mortality in our data indicates the need for a novel treatment strategy to prevent, halt, or reverse the severe lung inflammation or damage in this patient population. Tofacitinib, a candidate treatment option with recent reports on its potential use in myositis-associated ILD (11,12), was administered to only 1 patient captured in our study.

Our study has several limitations. First, the database does not contain either laboratory or imaging data, including known prognostic factors such as autoantibodies, serum biomarkers, and lung computed tomography imaging patterns. Second, the database does not contain the primary objective of hospitalization and the cause of death. In order to mitigate the possibility of enrolling patients who were admitted and underwent mechanical ventilation due to infectious pneumonia, we added methylprednisolone pulse as one of the inclusion criteria. Third, although this study is currently the largest descriptive study of myositis-associated severe ILD, the number of patients who fulfilled inclusion criteria was insufficient to conduct multivariate analyses or further stratifications. Fourth, as this study is based on a database in Japan, its generalizability to other countries is uncertain. Finally, the database does not contain outpatient data, and hence, it is unclear if the patients were newly diagnosed or had been treated with immunosuppressant medications before the hospitalization.

In conclusion, we performed a descriptive study of current practice and in-hospital mortalities of 549 patients with myositis-associated severe ILD who required mechanical ventilation and methylprednisolone pulse therapy using a nationwide administrative inpatient database in Japan. The mortality of these patients was extremely high, although aggressive and prompt interventions were made. Novel treatment strategies to prevent, halt, or reverse the severe lung inflammation and damage in this population with a poor prognosis are greatly needed.

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. Ikeda 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. Furuya, M. Nakajima, Ikeda, Nakamura, Ohbe, Yasunaga, H. Nakajima.

Acquisition of data. Furuya, M. Nakajima, Ikeda, Ohbe, Aso, Kumazawa, Yasunaga, H. Nakajima.

Analysis and interpretation of data. Furuya, M. Nakajima, Ikeda, Nakamura, Ohbe, Aso, Iwamoto, Iwata, Furuta, Matsui, Fushimi, Yasunaga, H. Nakajima.

REFERENCES

1. Ohta A, Nagai M, Nishina M, Tomimitsu H, Kohsaka H. Prevalence and incidence of polymyositis and dermatomyositis in Japan. Mod Rheumatol 2014;24:477–80.
2. Ikeda S, Arita M, Misaki K, Mishima S, Takaiwa T, Nishiyama A, et al. Incidence and impact of interstitial lung disease and malignancy in patients with polymyositis, dermatomyositis, and clinically amyopathic dermatomyositis: a retrospective cohort study. Springerplus 2015; 4:240.
3. Ye S, Chen XX, Lu XY, Wu MF, Deng Y, Huang WQ, et al. Adult clinically amyopathic dermatomyositis with rapid progressive interstitial lung disease: a retrospective cohort study. Clin Rheumatol 2007;26: 1647–54.
4. Sato S, Masui K, Nishina N, Kawaguchi Y, Kawakami A, Tamura M, et al. Initial predictors of poor survival in myositis-associated interstitial lung disease: a multicentre cohort of 497 patients. Rheumatology (Oxford) 2018;57:1212–21.
5. Wu J, Wang C, Toh S, Pisa FE, Bauer L. Use of real-world evidence in regulatory decisions for rare diseases in the United States: current status and future directions. Pharmacoepidemiol Drug Saf 2020;29: 1213–8.
6. Yasunaga H. Real world data in Japan: chapter II. The Diagnosis Procedure Combination Database. Ann Clin Epidemiol 2019;1:76–9.
7. Yamana H, Moriwaki M, Horiguchi H, Kodan M, Fushimi K, Yasunaga H. Validity of diagnoses, procedures, and laboratory data in Japanese administrative data. J Epidemiol 2017;27:476–82.
8. Kawasumi H, Gono T, Kawaguchi Y, Yamanaka H. Recent treatment of interstitial lung disease with idiopathic inflammatory myopathies. Clin Med Insights Circ Respir Pulm Med 2015;9 Suppl 1: 9–17.
9. Ichiyasu H, Horio Y, Tsumura S, Hirosako S, Sakamoto Y, Sakata S, et al. Favorable outcome with hemoperfusion of polymyxin B–immobilized fiber column for rapidly progressive interstitial pneumonia associated with clinically amyopathic dermatomyositis: report of three cases. Mod Rheumatol 2014;24:361–5.
10. Abe Y, Kusaoi M, Tada K, Yamaji K, Tamura N. Successful treatment of anti-MDA5 antibody-positive refractory interstitial lung disease with plasma exchange therapy. Rheumatology (Oxford) 2020;59: 767–71.
11. Chen Z, Wang X, Ye S. Tofacitinib in amyopathic dermatomyositis-associated interstitial lung disease. N Engl J Med 2019;381:291–3.
12. Kato M, Ikeda K, Kageyama T, Kasuya T, Kumagai T, Furuya H, et al. Successful treatment for refractory interstitial lung disease and

pneumomediastinum with multidisciplinary therapy including tofacitinib in a patient with anti-MDA5 antibody-positive dermatomyositis. J Clin Rheumatol 2021;27:S574–7.

13. Marie I, Hatron PY, Dominique S, Cherin P, Mouthon L, Menard JF. Short-term and long-term outcomes of interstitial lung disease in polymyositis and dermatomyositis: a series of 107 patients. Arthritis Rheum 2011;63:3439–47.
14. Aso S, Matsui H, Fushimi K, Yasunaga H. Effect of cyclosporine A on mortality after acute exacerbation of idiopathic pulmonary fibrosis. J Thorac Dis 2018;10:5275–82.
15. Abe S, Azuma A, Mukae H, Ogura T, Taniguchi H, Bando M, et al. Polymyxin B-immobilized fiber column (PMX) treatment for idiopathic pulmonary fibrosis with acute exacerbation: a multicenter retrospective analysis. Intern Med 2012;51:1487–91.

American College of Rheumatology
Empowering Rheumatology Professionals

Arthritis Care & Research
Vol. 74, No. 3, March 2022, pp 484–492
DOI 10.1002/acr.24485

Outcomes and Quality of Care in Rheumatoid Arthritis With or Without Video Telemedicine Follow-Up Visits

Elizabeth D. Ferucci, Gretchen M. Day, Tammy L. Choromanski, and Sarah L. Freeman

Objective. Telemedicine has been proposed to improve access to care in rheumatology, but few studies of telerheumatology have been published. The objective of this study was to evaluate outcomes and quality of care for rheumatoid arthritis (RA) in patients seen by video telemedicine follow-up compared to in-person only.

Methods. Individuals in the Alaska Tribal Health System with a diagnosis of RA were recruited when seeing a rheumatologist either in-person or by video telemedicine, both of which were offered as part of usual follow-up care. At baseline, participants completed the Routine Assessment of Patient Index Data 3 (RAPID3) questionnaire and a telemedicine perception survey and agreed to medical record review. Participants repeated surveys by telephone at 6 and 12 months, and medical record abstraction was performed at 12 months for quality measures.

Results. At the 12-month outcome assessment, 63 of 122 RA patients (52%) had ever used telemedicine for RA. In univariate analysis, functional status improved over 12 months in the telemedicine group. In multivariate analysis, RAPID3 score and functional status were associated with telemedicine group (higher), with no statistically significant change over the 12-month period. The only quality measure that differed between groups at 12 months in univariate analysis was the proportion of visits in which disease activity was documented (higher in the in-person group, 40% versus 25%; $P = 0.02$), but this was not significant after multivariate analysis.

Conclusion. In short-term follow-up, there was no significant difference in most outcome and quality measures in patients with RA who incorporated telemedicine follow-up in their care compared to in-person only.

INTRODUCTION

In rheumatoid arthritis (RA), guidelines highlight the importance of early diagnosis and initiation of disease-modifying antirheumatic drugs (DMARDs), with frequent disease monitoring using a treat-to-target strategy setting a target of remission or low disease activity (1). Studies have supported the importance of rheumatologists in the care and monitoring of patients with RA, with several studies demonstrating improved quality of care when a rheumatologist is included (2–5). However, access to rheumatologists is limited, especially in rural areas (6,7). Leveraging telemedicine has been proposed as a method to address workforce issues in rheumatology, with the potential to improve access to care for underserved communities (8).

Prior to the COVID-19 pandemic, adoption of telemedicine in rheumatology was slower than in other specialties, and few studies of telerheumatology had been published as of a systematic review in 2015 (9). Since 2015 and prior to the pandemic, more rheumatologists were using telemedicine, and a few additional studies were published (10–13). During the COVID-19 pandemic, there has been a dramatic increase in use of telemedicine in all specialties (14). Complicating research in this field, telemedicine encompasses many different methods of using technology to deliver health care, including different communication methods (synchronous or asynchronous), phases of care (initial consult or follow-up visit), types of presenters (physicians, other trained presenters, presenters without specific training, or no presenters), and disease states (9). Synchronous telemedicine can be conducted from clinic to clinic or directly with the patient's home or

The content herein is solely the responsibility of the authors and does not necessarily represent the official views of the Agency for Healthcare Research and Quality.

Supported by the Agency for Healthcare Research and Quality (grant R21HS024540).

Elizabeth D. Ferucci, MD, MPH, Gretchen M. Day, MPH, Tammy L. Choromanski, MPH, Sarah L. Freeman, PharmD: Alaska Native Tribal Health Consortium, Anchorage.

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

Address correspondence to Elizabeth D. Ferucci, MD, MPH, 3900 Ambassador Drive, Suite 201, Anchorage, AK 99508. Email: edferucci@anthc.org.

Submitted for publication April 28, 2020; accepted in revised form October 8, 2020.

SIGNIFICANCE & INNOVATIONS

- This study is one of a few to provide data on disease activity, functional status, and quality of care for rheumatoid arthritis (RA) in patients with video telemedicine visits incorporated in their follow-up care compared to in-person only care.
- Overall, there were few differences between the groups in outcomes and quality of care, although the study was of a relatively small size and short duration.
- The findings of this study suggest that video telemedicine is reasonable to incorporate into usual care for patients with RA.
- This study was conducted prior to the onset of the COVID-19 pandemic, and since that time, the uptake of telemedicine in rheumatology and other specialties has increased dramatically.

mobile device. In the systematic review of telerheumatology, the most commonly studied form of telemedicine was synchronous, with RA or other inflammatory arthritis as the most common disease state (9). The most common presenters were physicians, and telerheumatology was most commonly studied in the initial consultation phase of care (9). Finally, telemedicine can be incorporated into rheumatology practices in many ways, but in most cases it is not used exclusively to replace in-person visits. In this observational study of telerheumatology in the Alaska Tribal Health System (ATHS), we focus on the practice model used in our system during the study, namely synchronous video telemedicine from clinic to clinic, to provide some but not all care in the follow-up phase of RA with the use of a nonphysician presenter who is not specifically trained in rheumatology.

In the ATHS, asynchronous telemedicine (otherwise known as store-and-forward or eConsult) was available for 20 years but not widely used in rheumatology. Improved connectivity allowed for the expansion of synchronous video telemedicine for specialty care in the ATHS. In rheumatology, telemedicine has been available as part of usual care since 2015, as described previously (15). Rheumatologists are based at the Alaska Native Medical Center (ANMC) in Anchorage and provide care to the Alaska Native population statewide. Specialty care is provided using regional field clinics, telemedicine visits, and in-person visits at the ANMC. The cost of travel from rural communities to Anchorage can be significant. Because of the need to provide care to patients residing in >200 small communities, telemedicine is performed with a presenter who is not specifically trained in rheumatology. Telemedicine can be used for any disease state, but our study focused on RA because it is the most common condition seen.

This observational study was designed to evaluate the outcomes of and quality of care for RA in patients seen by video telemedicine at least once for follow-up compared to in-person only over the course of 1 year, when video telemedicine and in-person care were both available options in usual care. A previous publication from this study described the use of telemedicine in rheumatology in the ATHS, as well as factors associated with telemedicine use by RA patients at baseline (15). This analysis focuses on differences in outcomes (patient-reported disease activity and functional status) as well as quality of care (based on medical record abstraction) between individuals ever seen by video telemedicine compared to in-person only over the 1-year study period.

PATIENTS AND METHODS

Inclusion criteria. Individuals age ≥18 years with a diagnosis of RA confirmed by a rheumatologist who were being seen by a rheumatologist at the ANMC, either in person or by telemedicine, were invited to participate in this study. Telemedicine had been available for rheumatology care for ~1 year prior to initial recruitment, with clinical practice described previously (15). Telemedicine was not used exclusively but as an option to supplement in-person care. As an observational study of existing practice, we were unable to randomize patients to receive telemedicine nor control how and when clinic staff offered telemedicine to patients. Enrollment occurred from August 2016 until March 2018. Recruitment relied on clinic staff (to introduce the study to eligible patients) and flyers in clinic, with research staff providing more detailed information to interested patients when clinic staff indicated possible interest. Research staff obtained written or oral informed consent from participants. For enrolled participants, follow-up continued for 1 year, and all follow-up was completed by March 2019. The study was approved by the Alaska Area Institutional Review Board. Tribal approval was obtained from participating tribal health organizations.

Baseline study visit and clinical data. After providing informed consent, participants completed a baseline study assessment at the time of enrollment, which could be conducted by research staff either in-person or by telephone and often coincided with a rheumatology clinic visit. A study procedure flow diagram is provided in Figure 1. The baseline data collection included 2 surveys (telemedicine perception survey and the Routine Assessment of Patient Index Data 3 [RAPID3] questionnaire). The telemedicine perception survey questions and results have been described previously (15). Participants also agreed to medical record abstraction at baseline for demographic information, disease characteristics, comorbidities, and measures of access to and quality of care. Medical record abstraction was performed using a standardized abstraction form, as described previously (15). Health care–related elements included the number of visits with a rheumatologist in the preceding year, whether each visit was conducted by telemedicine or in-person, and which rheumatologist was seen. Quality measures selected for this study

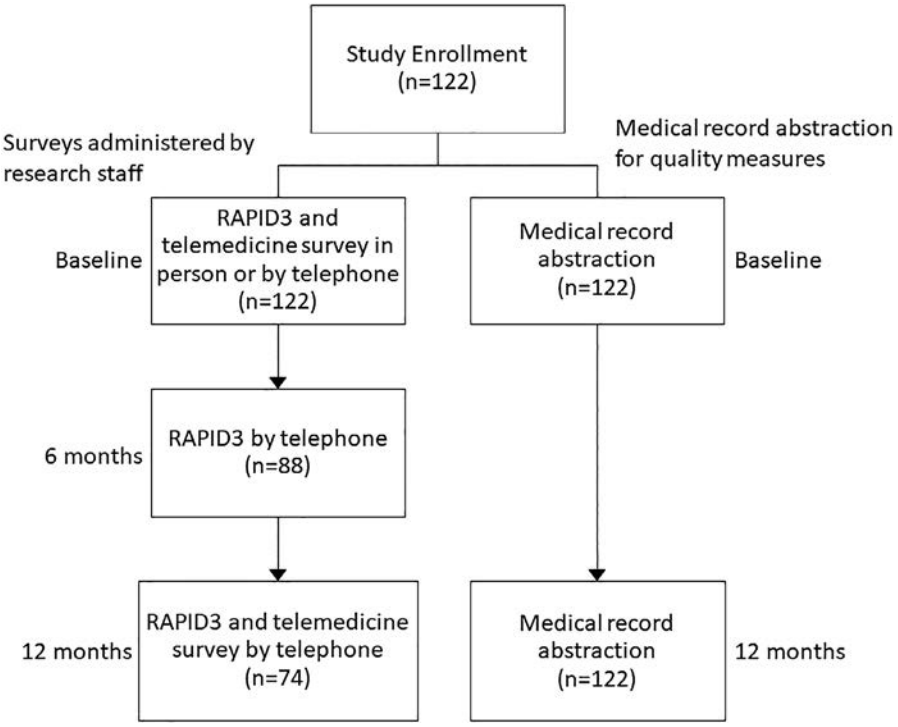


Figure 1. Flow chart for study procedures after study enrollment. The study included 2 forms of data collection: surveys administered by research staff and medical record abstraction for information about clinical care provided. On the left are the procedures for study surveys administered by research staff at baseline, 6 months, and 12 months and the number of participants completing each step. On the right are the procedures for medical record abstraction at baseline and 12 months and the number of participants for whom the abstraction was completed at each time point. RAPID3 = Routine Assessment of Patient Index Data 3.

included those endorsed by the American College of Rheumatology at the time of study design (16,17). Measures abstracted included: 1) whether disease activity was documented at each visit (by composite disease activity score or categorized without a composite score, similar to methods of Desai et al [18]); 2) if the disease activity was moderate or high, whether a change in therapy was prescribed; 3) whether a functional status assessment was documented at each visit; 4) whether a DMARD was prescribed in the past year; 5) whether tuberculosis screening was performed prior to first biologic DMARD initiation; and 6) whether prolonged glucocorticoids were prescribed without a management plan.

Longitudinal data collection. Participants were contacted by research staff by telephone at 6 and 12 months for outcome assessment. This included RAPID3 by telephone at both time points and a repeat telemedicine perception survey at 12 months. Medical record abstraction of quality measures was repeated at 12 months for all participants. Patient-level variables included: 1) the number of visits with a rheumatologist in the year after study enrollment; 2) whether a DMARD was prescribed in the past 12 months; 3) whether the patient had been prescribed prolonged glucocorticoids without a documented glucocorticoid management plan; and 4) whether the first biologic DMARD had been initiated in the past year, and if so, whether tuberculosis screening had been performed within 6 months prior. Visit-level variables included: 1) whether the visit was performed by telemedicine; 2) whether disease activity was documented; 3) whether a medication change was prescribed if disease activity was moderate or high; and 4) whether functional status assessment was documented.

Statistical analysis. Statistical analysis was performed using SAS, version 9.4. A 2-sided *P* value less than 0.05 was considered significant. The study was designed to have 86% power to detect a difference in functional status by group, assuming mean and variability as determined in other populations. Participants were categorized as being in the telemedicine group if they had ever had a telemedicine visit with a rheumatologist. Otherwise, they were categorized in the in-person only group. Factors that differed between groups were analyzed at the 1-year time period, similar to baseline analysis (15). Two composite scores were included in this analysis (the mean rheumatologist telemedicine rate and the telemedicine survey score), as previously described (15). Briefly, the mean rheumatologist telemedicine rate was defined as the weighted visit mean of rheumatologist telemedicine proportions to account for differing proportions of overall visit load of each rheumatologist

Table 1. Characteristics of patients with rheumatoid arthritis seen by telemedicine versus in-person only at 1 year*

Characteristic	Telemedicine (n = 63)	In-person only (n = 59)	*P*
Female sex, no. (%)	52 (83)	50 (85)	0.74
RAPID3 score (0–30 scale; n = 74)	12.1 ± 5.8†	10.0 ± 5.1‡	0.12
No. of rheumatology visits in past year	1.8 ± 1.2	1.7 ± 1.4	0.78
DMARD in past year, no. (%)	61 (97)	58 (98)	0.6
Rheumatologist telemedicine rate in past year	21.7 ± 23.1	8.4 ± 11.2	0.0001
Telemedicine survey score at 1 year (range –2 to 2; n = 74)	0.50 ± 0.41†	0.07 ± 0.41‡	<0.0001
Change in telemedicine survey score from baseline to 1 year (n = 74)	0.02 ± 0.5†	0.05 ± 0.5‡	0.84
Remission or low disease activity by RAPID3 at 1 year, no. (%) (n = 74)	10 (25)	10 (29)	0.57

* Values are the mean ± SD unless indicated otherwise. DMARD = disease-modifying antirheumatic drug; RAPID3 = Routine Assessment of Patient Index Data 3.
† N = 33.
‡ N = 41.

conducted by telemedicine. For this analysis, it was dichotomized into "high" (≥5% of visits conducted by telemedicine) or "low" (<5% of visits conducted by telemedicine) for better fit in multivariate models. There were a total of 5 rheumatologists during the course of the study. The telemedicine survey score is the average response to the 12 survey questions with values ranging from –2 to 2, with a higher score indicating more receptiveness to telemedicine, and a score of 0 indicating a neutral response.

Initial comparisons between groups were performed with univariate methods. For patient-reported measures, the analysis only included patients who completed the RAPID3 questionnaire at 12 months. The change in RAPID3 score, functional status, and proportion of patients in low disease activity or remission by RAPID3 were analyzed from RAPID questionnaires at baseline, 6 months, and 12 months for the telemedicine and in-person groups using repeated-measures analysis of variance for continuous variables and repeated-measures logistic regression for categorical variables. Repeated linear measures mixed models were used to evaluate factors associated with the continuous variables over time. Covariates assessed included age, sex, group (telemedicine versus in-person only), telemedicine survey score, number of rheumatologist visits in the preceding year, number of rheumatologist visits in the study year, rheumatologist telemedicine rate, ever seen by telemedicine by non-rheumatologist, and time point (baseline, 6 months, and 12 months). An interaction term between group and time point was included in the models. A significant interaction would mean that the change in scores over time would be different by group. Models were selected based on the fit of the model and significance of individual covariate estimates in the model. Generalized estimating equations (GEEs) were used for the categorical variable to account for correlation between data from the same patient during multiple rounds of follow-up. The GEE results are based on an unstructured correlation matrix.

The analysis of quality of care for RA by group included data from medical record abstraction at 12 months and included all study participants. Most quality measures were analyzed by visit, and some patients did not have a visit with a rheumatologist during the 12 months after enrollment. Only patients with a visit were included in the analyses conducted by visit. All patients were included in the analysis by patient.

For all quality measures, initial comparisons between groups were performed using univariate methods. We used Poisson regression to examine factors associated with the number of rheumatology visits during the study, controlling for age and sex. We evaluated the following variables in the model: group (telemedicine or in-person only), number of visits in the year prior to

Table 2. Disease activity and functional status over 12 months in patients with rheumatoid arthritis seen by telemedicine versus in-person only*

	Telemedicine				In-person only			
	Baseline (n = 62)	6 months (n = 48)	12 months (n = 39)	*P*	Baseline (n = 60)	6 months (n = 40)	12 months (n = 35)	*P*
RAPID3 score	12.4 ± 5.3	11.5 ± 5.0	12.1 ± 5.7	0.7†	10.4 ± 5.6	9.8 ± 4.7	10.0 ± 5.4	0.50†
Low disease activity or remission, no. (%)	11 (18)	10 (21)	9 (23)	0.5‡	17 (28)	9 (23)	11 (31)	0.42‡
Functional status score	2.9 ± 1.9	2.5 ± 1.7	2.4 ± 1.8	0.02†	1.9 ± 1.9	2.1 ± 2.0	1.8 ± 1.3	0.69†

* Values are the mean ± SD unless indicated otherwise. RAPID3 = Routine Assessment of Patient Index Data 3.
† Repeated-measures analysis of variance.
‡ Repeated-measures logistic regression.

Table 3. Multivariate analyses of disease activity and functional status over 12 months in patients with rheumatoid arthritis seen by telemedicine versus in-person only*

Dependent variable (model type) and independent variables	Estimate (95% CI)	OR (95% CI)	*P*
RAPID3 score (mixed model)			
Age, linear	0.07 (0.01, 0.14)	–	0.04
Male sex	−0.16 (−2.60, 2.27)	–	0.89
Telemedicine group	2.38 (0.47, 4.30)	–	0.02
Telemedicine survey score	−0.86 (−1.58, −0.14)	–	0.02
Visits during study	−1.22 (−2.36, −0.09)	–	0.04
Time 6 months	0.25 (−1.18, 1.68)	–	0.73
Time 12 months	0.47 (−1.26, 2.20)	–	0.59
Group × 6 months interaction	−0.18 (−1.80, 1.43)	–	0.82
Group × 12 months interaction	0.41 (−1.75, 2.57)	–	0.71
Functional status score (mixed model)			
Age, linear	0.04 (0.02, 0.06)	–	0.0007
Male sex	0.26 (−1.09, 0.57)	–	0.39
Telemedicine group	1.02 (0.36, 1.68)	–	0.003
Time 6 months	0.01 (−0.31, 0.33)	–	0.95
Time 12 months	0.02 (−0.34, 0.39)	–	0.91
Group × 6 months interaction	−0.36 (−0.79, 0.08)	–	0.11
Group × 12 months interaction	−0.44 (−0.94, 0.06)	–	0.08
Low disease activity or remission (GEE)			
Age, linear	–	0.96 (0.93, 0.99)	0.009
Male sex	–	2.39 (1.00, 5.70)	0.05
Telemedicine group	–	0.41 (0.17, 1.04)	0.06
Time 6 months	–	0.82 (0.37, 1.82)	0.62
Time 12 months	–	1.15 (0.54, 2.46)	0.73
Group × 6 months interaction	–	1.59 (0.47, 5.37)	0.45
Group × 12 months interaction	–	1.38 (0.42, 4.53)	0.6
Telemedicine survey score	–	1.65 (1.05, 2.57)	0.03

* 95% CI = 95% confidence interval; GEE = generalized estimating equation; OR = odds ratio; RAPID3 = Routine Assessment of Patient Index Data 3.

enrollment, baseline RAPID3 score, baseline telemedicine survey score, and rheumatologist telemedicine rate. Multivariable logistic regression was conducted for each of the remaining quality measures. We controlled for age (linear), sex, and group in all models and evaluated the covariates listed above in addition to whether the visit was conducted by telemedicine or not for visit-level analyses. Models were selected based on the fit of the model and of individual covariate estimates in the model.

RESULTS

The characteristics of study participants at 1 year after study enrollment are presented in Table 1. Baseline characteristics, including disease characteristics, were described previously (15). As demonstrated in Table 1 and consistent with baseline, participants in the telemedicine group had a higher mean rheumatologist telemedicine rate and more positive perceptions of telemedicine. Although the telemedicine group had a higher mean RAPID3 score than the in-person group at 1 year (12.1 versus 10.0), the difference was not statistically significant, while the difference had been significant at baseline (12.6 versus 10.4; $P = 0.037$) (15). Not all participants completed the final RAPID3 questionnaire (n = 74 of 122), so the power to detect differences was lower. In addition, variability in RAPID3 score and functional status were higher than expected. Demographic and disease characteristics of those lost to follow-up compared to those completing follow-up surveys were similar,

Table 4. Quality measures from medical record abstraction at 1 year*

Quality measure	Telemedicine†	In-person only‡	*P*
No. of rheumatologist visits in the year after study enrollment, mean ± SD	1.8 ± 1.2	1.7 ± 1.4	0.67
At least 1 visit to a rheumatologist in the study year	56 (89)	45 (76)	0.06
Proportion of visits in which disease activity is documented	28 (25)	41 (40)	0.02
Proportion of visits with moderate or high disease activity documented in which a change in medications is prescribed, no./total no. (%)	19/23 (83)	17/23 (74)	0.47
Proportion of visits in which functional status assessment is documented	28 (25)	30 (29)	0.45
Patients prescribed a DMARD in past year	61 (97)	58 (98)	0.6

* Values are the number (%) unless indicated otherwise. DMARD = disease-modifying antirheumatic drug.
† N = 63 patients with 114 visits.
‡ N = 59 patients with 103 visits.

except that a higher proportion of those completing follow-up were female (89% versus 76%; $P = 0.05$). Six participants who had never used telemedicine prior to study enrollment had telemedicine visits during the study. These participants were categorized in the telemedicine group.

Patient-reported outcomes, including RAPID3 scores and functional status, are presented from baseline, 6-month, and 12-month surveys in Table 2. In the telemedicine group, there was no change in RAPID3 score or the proportion of patients in low disease activity over time in univariate analyses, but there was a statistically significant improvement in functional status over time. In the in-person group, there was no significant change in any measures over time. Multivariate analyses of patient-reported outcomes are presented in Table 3, including details about the models and variables included. In the repeated linear measures mixed model for RAPID3, telemedicine group, increasing age, a lower number of visits during the year, and a lower telemedicine survey score were significantly associated with a higher RAPID3 score. There was no significant association of RAPID3 with time or group over time. The multivariate model of functional status score over time had significant associations with telemedicine group and age, but the decline in functional status score over time that was evident for the telemedicine group in univariate analysis was no longer significant in multivariate analysis. Being in low disease activity or remission was significantly associated with male sex and telemedicine survey score and negatively associated with age, but not associated with group or time.

Univariate analysis of differences in quality measures between groups is presented in Table 4. There were 63 patients with 114 visits in the telemedicine group, and 59 patients with 103 visits in the in-person group. Although the mean number of visits was similar for each group, the in-person group was somewhat more likely to have had no visits during the study period ($P = 0.06$). There was no difference between groups in the measures analyzed at the patient level (number of rheumatologist visits and proportion prescribed a DMARD). For visit-level analysis, a higher proportion of visits had disease activity documented in the in-person group compared to the telemedicine group. Table 4 does not include glucocorticoid management plan and tuberculosis screening prior to first biologic because few patients were eligible for these measures (<10 patients for each, with at least 80% of those patients meeting the target for each measure).

Table 5. Multivariate analysis of factors associated with quality measures from medical record abstraction at 1 year*

Dependent variable (model type) and independent variables	Estimate, beta (95% CI)	OR (95% CI)	*P*
No. of rheumatologist visits in past year (Poisson regression) (n = 122 patients)			
Age, linear	–0.002 (–0.012, 0.009)	–	0.73
Male sex	0.01 (–0.357, 0.379)	–	0.95
Telemedicine group	–0.02 (–0.293, 0.249)	–	0.87
No. of visits in year prior to study	0.11 (0.016, 0.201)	–	0.02
Documentation of disease activity at the visit (multivariate logistic regression) (n = 217 visits)			
Age, linear	–	0.99 (0.97, 1.02)	0.66
Male sex	–	0.52 (0.21, 1.30)	0.16
Telemedicine group	–	0.54 (0.29, 1.02)	0.06
RAPID3 at baseline	–	0.89 (0.83, 0.95)	0.0004
No. of visits in year prior to study	–	0.76 (0.60, 0.97)	0.02
Change in medications is prescribed when disease activity is moderate or high (multivariate logistic regression) (n = 46 visits)			
Age, linear	–	1.03 (0.96, 1.11)	0.38
Male sex	–	1.92 (0.09, 42.5)	0.68
Telemedicine group	–	2.44 (0.45, 13.2)	0.3
No. of visits in year prior to study	–	0.42 (0.18, 0.98)	0.04
Functional status assessment documentation (multivariate logistic regression) (n = 217 visits)			
Age, linear	–	1.00 (0.97, 1.03)	0.85
Male sex	–	1.40 (0.60, 3.2)	0.44
Telemedicine group	–	1.48 (0.75, 2.90)	0.26
High rheumatologist telemedicine rate	–	0.04 (0.01, 0.20)	<0.001
No. of visits in year prior to study	–	1.30 (1.04, 1.62)	0.02
DMARD prescribed in past year (multivariate logistic regression) (n = 122 patients)			
Age, linear	–	1.27 (1.03, 1.55)	0.02
Male sex	–	0.06 (0.002, 1.72)	0.1
Telemedicine group	–	1.04 (0.05, 21.3)	0.98

* 95% CI = 95% confidence interval; DMARD = disease-modifying antirheumatic drug; OR = odds ratio; RAPID3 = Routine Assessment of Patient Index Data 3.

Multivariate analyses of quality measures are presented in Table 5. The number of visits in the past year was associated with number of visits in the year prior to the study, but not with group or other variables. Documentation of disease activity was less likely for those with a higher RAPID3 score at baseline and with a higher number of visits in the past year but was not associated with group. A change in medications when disease activity was moderate or high was negatively associated with the number of visits in the year prior to the study but was not associated with group or other variables. Having a functional status assessment was associated with number of visits in the year prior to the study and negatively associated with a high rheumatologist telemedicine rate, but not with group. Having a DMARD prescribed in the past year was associated with increased age, but not group or other factors. As shown in Table 4, 97% of patients in the telemedicine group and 98% of patients in the in-person only group were prescribed a DMARD in the past year.

DISCUSSION

In this observational study comparing telemedicine incorporated into usual follow-up care for patients with RA to in-person only care, few differences were detected in patient-reported outcomes or quality of care. In univariate analysis, functional status improved over the study period in the telemedicine group. In multivariate analysis, RAPID3 score and functional status were higher in the telemedicine group, but this did not change over time. The only quality measure that differed between groups was the proportion of visits in which disease activity was documented, but this was not significant in multivariate analysis. This study was limited in size and duration.

Few studies have evaluated the outcomes of telehealth in rheumatology, and many recently published studies in RA focus on modalities other than synchronous video telemedicine, such as text messaging (19,20) or telephone-based interventions (21). A study evaluating synchronous video telemedicine for inflammatory arthritis in rural veterans, similar to our study, found no difference in RAPID3 score between the telemedicine and usual care groups at baseline and no change in RAPID3 score in longitudinal follow-up (10). There were significant savings in visit costs and distance traveled, leading to the conclusion that factors important to patients should be considered in structuring health care delivery (10). A survey-based study in pediatric rheumatology similarly found that families reported fewer financial burdens when telemedicine was used (22). In Australia, a study of the patient perspective on a program providing synchronous telerheumatology follow-up of stable patients found that patients reported a significant reduction in travel (23). Although patients expressed some reservations about the quality of telemedicine visits initially, they viewed telerheumatology as equivalent to in-person care (23).

In this study, we found a higher RAPID3 score in the telemedicine group at baseline (15), and although functional status improved over time in the telemedicine group, multivariate analysis found higher RAPID3 and functional status scores in the telemedicine group. Given that our study was observational, we hypothesize that higher disease activity factored into the patient's decision to have a telemedicine visit, when the alternative was likely to wait longer for an in-person visit. Although we attempted to control for patient factors in multivariate analyses, not all differences were measured by our study instruments. Patient views on telemedicine and more frequent use of telemedicine by the rheumatologist were the main factors associated with its use. We have ongoing qualitative studies to better inform our understanding of how patients and providers make decisions about using telemedicine. Overall, we concluded that there was no statistically significant improvement in RAPID3 or functional status scores over the study period in either group, and no difference in the change over time by group.

Several measures of quality of care were abstracted from the medical record. The in-person group was slightly more likely to have no visits with a rheumatologist in the study period. This may indicate more availability of telemedicine visits or that people seen in-person only may not have needed as frequent follow-up visits. Disease activity documentation was more likely in the in-person group in univariate analysis. The logistics of documenting disease activity differ in a telemedicine visit compared to an in-person visit, which may make it less likely to be documented. With a presenter not trained in rheumatology, only some aspects of a joint examination can be performed such as visual inspection and range of motion, not a formal tender joint count and swollen joint count. Therefore, disease activity cannot be measured using disease activity scores commonly used in clinical practice such as the Clinical Disease Activity Index and the Disease Activity Score in 28 joints. A RAPID3 score does not require a joint count and can be performed by a presenter not trained in rheumatology. However, if it is not collected prior to the visit, it may be time consuming for the rheumatologist to collect during the visit itself. Because the process of documenting disease activity in a telemedicine visit is more complex than in an in-person visit, it might be expected that it would be less commonly performed. Other quality measures did not differ between groups. It is likely that quality measures captured by medical record abstraction do not adequately capture all aspects of the quality of a visit, especially in the case of visits conducted by telemedicine, and alternate measures designed for telemedicine visits should be considered. For example, because less time is typically spent on physical examination in a telemedicine visit, there is more time available for education about RA and medications, including shared decision-making.

This study had some limitations. First, this was an observational study in a setting where telemedicine was already in use, incorporated with in-person care, and we were unable to randomize patients to telemedicine or in-person only care nor control how the option of telemedicine was presented to patients by clinic staff. We found that RAPID3 scores were higher in patients selecting telemedicine at baseline, but there are likely

other unmeasured factors associated with the choice of telemedicine that could influence outcomes. We controlled for covariates when possible but may not have accounted for all differences between groups in our analysis. A randomized controlled trial of telemedicine is no longer feasible in most practice settings given the increase in use during the COVID-19 pandemic. Second, the study was small, with some attrition over time with respect to patient-reported outcome data and more variability than expected, restricting the conclusions that can be drawn from this preliminary study due to power limitations. Despite survey attrition, data on quality measures from medical record abstraction were available for all patients. In addition, the characteristics of patients lost to follow-up were similar to those who remained in the study. Because the patient-reported data were collected by telephone rather than in the clinic, there is a lower risk of bias. Third, the study provides a comparison of telemedicine incorporated into practice along with in-person care, not telemedicine alone, because we were constrained by the existing practice. However, this practice model is one we believe to be most appropriate for rheumatology. Finally, the duration of this study was short, and gains that may occur with telemedicine likely require a longer duration of follow-up.

In conclusion, in this preliminary study, we found few differences in outcomes and quality of care for RA in the short term in patients seen by telemedicine at least once compared to in-person only, when telemedicine was incorporated into usual care. These findings indicate that telemedicine can reasonably be offered as a component of care for RA and may provide the ability for patients to be seen more often. There are benefits of telemedicine, such as reducing costs and improving communication, which may make it attractive even if outcomes are not improved compared to in-person only care. Longer-term studies are needed to evaluate telemedicine in RA. Although randomized controlled trials would be ideal, telemedicine has already been incorporated into many rheumatology practices, with a dramatic increase due to the COVID-19 pandemic, and such studies are no longer feasible. Ongoing research by our group will expand on this study using mixed methods to evaluate patient and provider perceptions of telemedicine, as well as assessing outcomes and costs of telemedicine. If the costs of providing clinical care incorporating telemedicine are much lower than for in-person only care or there are other benefits with similar outcomes, then it becomes an increasingly viable option.

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. Ferucci 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. Ferucci, Freeman.

Acquisition of data. Ferucci, Choromanski.

Analysis and interpretation of data. Ferucci, Day, Choromanski.

REFERENCES

1. Singh JA, Saag KG, Bridges SL Jr, Akl EA, Bannuru RR, Sullivan MC, et al. 2015 American College of Rheumatology guideline for the treatment of rheumatoid arthritis. Arthritis Rheumatol 2016; 68:1–26.
2. Smolen JS, Aletaha D, Bijlsma JW, Breedveld FC, Boumpas D, Burmester G, et al. Treating rheumatoid arthritis to target: recommendations of an international task force. Ann Rheum Dis 2010; 69:631–7.
3. MacLean CH, Louie R, Leake B, McCaffrey DF, Paulus HE, Brook RH, et al. Quality of care for patients with rheumatoid arthritis. JAMA 2000; 284:984–92.
4. Lacaille D, Anis AH, Guh DP, Esdaile JM. Gaps in care for rheumatoid arthritis: a population study. Arthritis Rheum 2005;53:241–8.
5. Ferucci ED, Donnithorne KJ, Koller KR, Swango-Wilson A, Pflaum J, Lanier AP. Performance on rheumatoid arthritis quality indicators in an Alaska Native healthcare system. BMJ Qual Saf 2010;19: 387–91.
6. Yazdany J, Tonner C, Schmajuk G, Lin GA, Trivedi AN. Receipt of glucocorticoid monotherapy among Medicare beneficiaries with rheumatoid arthritis. Arthritis Care Res (Hoboken) 2014;66: 1447–55.
7. Battafarano DF, Ditmyer M, Bolster MB, Fitzgerald JD, Deal C, Bass AR, 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.
8. Schmajuk G, Tonner C, Yazdany J. Factors associated with access to rheumatologists for Medicare patients. Semin Arthritis Rheum 2016; 45:511–8.
9. Ward IM, Schmidt TW, Lappan C, Battafarano DF. How critical is telemedicine to the rheumatology workforce? Arthritis Care Res (Hoboken) 2016;68:1387–9.
10. McDougall JA, Ferucci ED, Glover J, Fraenkel L. Telerheumatology: a systematic review. Arthritis Care Res (Hoboken) 2017;69:1546–57.
11. Wood PR, Caplan L. Outcomes, satisfaction, and costs of a rheumatology telemedicine program: a longitudinal evaluation. J Clin Rheumatol 2019;25:41–4.
12. Taylor-Gjevre R, Nair B, Bath B, Okpalauwaekwe U, Sharma M, Penz E, et al. Addressing rural and remote access disparities for patients with inflammatory arthritis through video-conferencing and innovative inter-professional care models. Musculoskeletal Care 2018;16:90–5.
13. Nguyen-Oghalai TU, Hunter K, Lyon M. Telerheumatology: the VA experience. South Med J 2018;111:359–62.
14. Rostom K, Smith CD, Liddy C, Afkham A, Keely E. Improving access to rheumatologists: use and benefits of an electronic consultation service. J Rheumatol 2018;45:137–40.
15. Ferucci ED, Holck P, Day GM, Choromanski TL, Freeman SL. Factors associated with use of telemedicine for follow-up of rheumatoid arthritis. Arthritis Care Res (Hoboken) 2020;72:1404–9.
16. American College of Rheumatology. ACR endorsed measures. URL: http://www.rheumatology.org/Practice-Quality/Clinical-Support/Quality-Measurement/ACR-Endorsed-Measures.
17. Yazdany J, Robbins M, Schmajuk G, Desai S, Lacaille D, Neogi T, et al. Development of the American College of Rheumatology's rheumatoid arthritis electronic clinical quality measures. Arthritis Care Res (Hoboken) 2016;68:1579–90.
18. Desai SP, Liu CC, Tory H, Norton T, Frits M, Lillegraven S, et al. Rheumatoid arthritis quality measures and radiographic progression. Semin Arthritis Rheum 2014;44:9–13.
19. Mary A, Boursier A, Desailly Henry I, Grados F, Séjourné A, Salomon S, et al. Mobile phone text messages and effect on treatment adherence

in patients taking methotrexate for rheumatoid arthritis: a randomized pilot study. Arthritis Care Res (Hoboken) 2019;71:1344–52.

20. Kuusalo L, Sokka-Isler T, Kautiainen H, Ekman P, Kauppi MJ, Pirilä L, et al. Automated text message–enhanced monitoring versus routine monitoring in early rheumatoid arthritis: a randomized trial. Arthritis Care Res (Hoboken) 2020;72:319–25.
21. De Thurah A, Stengaard-Pedersen K, Axelsen M, Fredberg U, Schougaard LM, Hjollund NH, et al. Tele-health followup strategy for tight control of disease activity in rheumatoid arthritis: results of a randomized controlled trial. Arthritis Care Res (Hoboken) 2018; 70:353–60.
22. Kessler EA, Sherman AK, Becker ML. Decreasing patient cost and travel time through pediatric rheumatology telemedicine visits. Pediatr Rheumatol Online J 2016;14:54.
23. Devadula S, Langbecker D, Vecchio P, Tesiram J, Meiklejohn J, Benham H. Tele-rheumatology to regional hospital outpatient clinics: patient perspectives on a new model of care. Telemed J E Health 2020;26:912–9.

Arthritis Care & Research
Vol. 74, No. 3, March 2022, pp 493–500
DOI 10.1002/acr.24438

American College of Rheumatology
Empowering Rheumatology Professionals

Physical Activity and Sedentary Behavior in People With Inflammatory Joint Disease: A Cross-Sectional Study

Kirsty Bell,[1] Gordon Hendry,[2] and Martijn Steultjens[2]

Objective. To determine whether patients with inflammatory joint disease (IJD) meet current guidelines on physical activity, and to determine which factors influence physical activity levels and sedentary behavior (SB) in patients with IJD.

Methods. This was a cross-sectional study of 137 patients with a medical diagnosis of an IJD prior to commencing an NHS-run inflammatory arthritis exercise program. Physical activity and SB were measured objectively using a thigh-worn physical activity monitor for 7 consecutive days. Activity levels were subdivided into low physical activity (LPA) and moderate-to-vigorous physical activity (MVPA). First, activity levels were analyzed against current guidelines of 150 minutes of MVPA per week. Second, time spent in SB, LPA, and MVPA was analyzed against possible determinants.

Results. In total, 29% of patients with IJD met current physical activity guidelines. Patients on average spent 10 hours per day in SB. Poor physical fitness measured by the 6-minute walk test was the only significant predictor ($P = 0.019$) of high SB ($R^2 = 4.7\%$). Attending an exercise facility in the community ($P = 0.034$) and low role limitations due to physical health ($P = 0.008$) predicted high levels of LPA, following a backward multiple regression ($R^2 = 8.0\%$). Low role limitations due to emotional problems ($P = 0.031$), higher physical fitness ($P = 0.002$), and healthier exercise attitudes and beliefs ($P = 0.021$) predicted meeting current physical activity guidelines, following a backward conditional logistic regression, explaining between 22.2% and 31.7% of variance.

Conclusion. Patients with IJD are inactive and spent much time in SB. Good general health predicts high activity levels. No disease-specific factors were found to determine SB, LPA, or MVPA.

INTRODUCTION

Physical activity, defined as "any bodily movement produced by skeletal muscles that results in energy expenditure" (1), has been shown to be of great benefit to people with inflammatory joint disease (IJD), such as rheumatoid arthritis (RA), inflammatory arthritis, and spondyloarthritis (2). Physical activity can help improve joint range of movement, muscle strength, aerobic capacity, and overall function (3). Evidence also exists that regular physical activity does not have any harmful effects such as an increase in joint pain or radiologic joint damage or an increase in disease activity (3,4). However, people with IJD are generally less active compared to healthy controls (5,6). A significant proportion of people with RA have been shown to be physically inactive, characterized by a failure to participate in bouts of moderate-to-vigorous physical activity (MVPA) of ≥10 minutes over 1 week (7). People with RA have also been shown not to meet physical activity guidelines for healthy physical activity levels, but instead demonstrate reduced physical activity and increased sedentary time relative to healthy controls (5,8). Lack of motivation to exercise, lack of belief in its benefits, and beliefs about negative side effects of exercise have been reported as important barriers to exercise in people with RA (9,10).

People with IJD have an increased risk of developing cardiovascular disease (CVD) compared to the general population (11–13). Cardiorespiratory fitness is low in people with RA and this condition is likely to be associated with the increased incidence of CVD-related deaths in RA (14). Cardiorespiratory fitness is important, as emerging evidence suggests that more time spent in sedentary behavior (SB), defined as an energy expenditure of ≤1.5 metabolic equivalents while in a sitting or reclining posture (15), is independently associated with greater risk of developing CVD, cancer, and diabetes mellitus (16). Conversely, people with RA who have higher cardiovascular fitness have a better CVD risk profile and a lower 10-year CVD events risk compared to those with lower cardiovascular fitness (14). There are also several

Supported by Versus Arthritis (grant 20498) and by PAL Technologies Ltd.
[1]Kirsty Bell, MRes, BSc (Hons): Crieff Community Hospital, Crieff, and Glasgow Caledonian University, Glasgow, UK; [2]Gordon Hendry, PhD, BSc (Hons), Martijn Steultjens, PhD, MSc: Glasgow Caledonian University, Glasgow, UK.
No potential conflicts of interest relevant to this article were reported.

Address correspondence to Kirsty Bell, MRes, BSc (Hons), Physiotherapy Department, Crieff Community Hospital, Crieff, Scotland, UK, PH7 3HR. Email: kirstybell2@nhs.net.
Submitted for publication January 8, 2020; accepted in revised form August 27, 2020.

SIGNIFICANCE & INNOVATIONS

- The majority of people with inflammatory joint disease do not meet current physical activity guidelines. People on average spend >10 hours of their waking time in sedentary behavior.
- People who attend an exercise facility in the community are more physically active.
- No disease-specific factors could be found to determine sedentary behavior, low physical activity, or moderate-to-vigorous physical activity in people with inflammatory joint diseases.

additional health benefits of physical activity for people with IJD beyond reducing health risks, such as reduced levels of fatigue, reduced disease activity, reduced stiffness, and increased joint health (5,17–19).

Current knowledge of the determinants of physical activity levels in people with IJD is limited. Studies investigating the determinants are largely confined to people with RA and do not extend to those with other common and clinically important IJDs (5,17,19). Methods of physical activity monitoring across these studies also differ, limiting the scope for comparison (7,8,17,19). The majority of studies have used subjective self-report methods to measure physical activity levels, which have been suggested to be subject to recall bias and to be less valid than objective methods (20). Studies that have adopted objective measures of physical activity in RA appear to lack internal validity due to monitor removal during activities such as swimming, and external validity due to variable definitions of low and high physical activity levels that do not follow current guidelines (5,15,17,21–23). Several putative factors that have been identified in other adult populations (24) that could influence physical activity levels and time spent sedentary, such as social derivation and exercise self-efficacy, have not been investigated in people with IJD.

The scarce evidence on determinants of physical activity and SB poses a significant challenge to clinicians who seek to address physical inactivity and SB in this patient group. A greater understanding of the determinants of physical activity levels and SB in people with IJD may facilitate a move toward alternative and enhanced approaches to physical activity interventions in the future.

Therefore, the main objectives of this study were to determine whether patients with IJD meet the current guidelines on physical activity, to determine which factors influence physical activity levels in patients with IJD, and to determine which factors influence SB in patients with IJD.

PATIENTS AND METHODS

Design. This was a cross-sectional study approved by the NHS Health Research Authority, NRES Committee South West–Exeter, UK (Ref: 14/SW/1183). All participants provided written informed consent according to the Declaration of Helsinki.

Participants. Patients were recruited from referrals into the NHS-run Inflammatory Arthritis Exercise Programme (IAEP) across the Greater Glasgow & Clyde (GG&C) Health Board. The NHS is a nation-wide universal health care system in Britain that is free at point of provision. The GG&C Health Board is the largest health board in Scotland, serving 1.2 million people with wide and variable socioeconomic characteristics. The IAEP is a 12-week exercise program run by rheumatology physiotherapists across the GG&C Health Board. Any adult within the health board who has a clinician-confirmed IJD and is under the care of the rheumatology department can be referred into the program.

Inclusion and exclusion criteria. Patients were included in the study if they met all of the following inclusion criteria: physician-confirmed diagnosis of an IJD such as RA, psoriatic arthritis, ankylosing spondylitis, or any other type of inflammatory arthritis/polyarthritis, and age ≥18 years. Patients were excluded from the study if they met any of the following criteria: they did not provide informed consent to be part of the study, they were unable to complete the study within the designated data collection period, or the presence of comorbidity severely limited the patient's ability to participate in an exercise program, such as unstable angina, heart failure, uncontrolled heart arrhythmias, uncontrolled hypertension, severe respiratory condition, uncontrolled epilepsy, or uncontrolled diabetes mellitus, or if the patient had a recent medical instability, such as a stroke, wheelchair use, or pregnancy.

Recruitment strategy. The study population of interest comprised patients who were under the care of the rheumatology department of the GG&C Health Board and who were referred into the IAEP between March 2015 and July 2017. Referrals into this program came from rheumatology consultants, rheumatology nurse specialists, rheumatology allied health professionals, and patient self-referrals. Every patient who was referred into this program and met the inclusion/exclusion criteria for the study was informed in writing and verbally of the research project by their rheumatology physical therapist, who they saw prior to attending the program. If the patient was interested in being part of the study, they were then contacted by the researcher to discuss the study in more depth and to gain written informed consent to become part of the study sample.

Data collection. Data were collected by the researcher prior to the patient commencing the IAEP. Physical activity and SB were objectively measured by wearing an ActivPAL (PAL Technologies Ltd) physical activity monitor permanently for 7 consecutive days prior to commencing the IAEP. This device measures body motion, which is defined by an energy expenditure

classification and a postural classification, enabling free-living behavior to be more accurately quantified (25). The device records acceleration counts used to determine energy expenditure, which can be converted into physical activity levels (26). It also records body position, which enables true SB to be recorded as classified by the Sedentary Behaviour Research Network (15). The ActivPAL was programmed to collect data for 7 consecutive days, as this collection provides a reliable measure of adult activity behaviors (27). The device was waterproofed as per the manufacturer's guidelines and worn centrally on the anterior aspect of the left or right thigh. The ActivPAL was fitted by the researcher on the day of data collection and removed by the patient at the beginning of day 8 and posted back to the researcher in a self-addressed stamped envelope. The device was programmed to commence data collection from midnight on the day that the device was fitted. Participants were also asked to self-monitor their physical activity via a hard copy activity diary while wearing the ActivPAL. Physical activity was specifically to record rise time and bedtime on each day of monitoring so that sleep time could be deducted from the data, to enable analysis on just the waking-time data. ActivPAL software was used for physical activity monitor programming, data processing, and data analysis. Low physical activity (LPA) was defined as <100 steps/minute and MVPA was defined as ≥100 steps/minute (26). True SB as defined by the Sedentary Behaviour Research Network (15) was calculated from the ActivPAL data.

Health-related quality of life was measured using the Short Form 36 (SF-36) and Health Assessment Questionnaire disability index (HAQ DI) (28). Self-perceived levels of control were measured using the Arthritis Self-Efficacy Scale (ASES), attitudes and beliefs toward physical activity were measured using the Exercise Attitudes and Beliefs Questionnaire for patients with RA (RA-EAQ), and mental health was measured using the Hospital Anxiety and Depression Scale (HADS). All of these measures have good psychometric properties that have been verified in populations with IJD (29–32). The Scottish Index of Multiple Deprivation (SIMD) measures across 7 domains: current income, employment, health, education, skills and training, housing, geographic access, and crime. These 7 domains are calculated and weighted for small areas, called data zones, with roughly equal population and can be obtained using the participant's postcode (33).

The Disease Activity Score in 28 joints (DAS28) was recorded as a marker of disease activity by the researcher who was trained in undertaking the DAS28 (KB). Acute-phase reactants from blood test results (within 3 months of each data collection session) were obtained from the patient's medical records to complete the DAS28 score. Disease duration was measured from the date of physician-confirmed diagnosis, which was obtained from the participant's medical records. Drug therapy was obtained from the patient's medical records and clarified with the patient in case of any recent changes; the level of pain on average over the past week was measured using a visual analog scale (VAS), and the level of fatigue was measured in the same way using the same 0–100-mm line as the pain VAS (18,19).

To evaluate whether there were any physical condition–related and/or environmental factors that could determine physical activity levels and SB, the following measurements were undertaken. Body mass index (BMI), calculated from the patient's height and weight on the day of data collection; the 6-minute walk test (34), using the American Thoracic Society and current clinical practice protocol (35,36), which measures fitness levels and is well established in IJD research (37); grip strength, using a Jamar grip dynamometer using the Southampton protocol for adult grip strength measurement (38), which has also been well established in IJD research (34,37); and a custom-made environmental questionnaire that was developed with assistance from the study advisory board, which consisted of rheumatology clinicians, NHS health improvement officers, patients, and academics. The questionnaire asked about cost, affordability, transportation to/from, and the variety of activities on offer at the community exercise facilities.

Statistical analysis. Descriptive statistics were used to summarize the variables. All variables were then assessed for normality of distribution using the Kolmogorov–Smirnov test. A Kruskal-Wallis test was carried out between the different diagnostic groups that showed no difference in activity levels between the groups; therefore, these were grouped together for analysis. Time spent in SB, LPA, and MVPA were analyzed against the possible determinants: HAQ DI, SF-36, age, disease duration, DAS28, pain, fatigue, medication, ASES, RA-EAQ, HADS, SIMD, BMI, general fitness, and grip strength using Pearson's (rp) or Spearman's (rs) correlation; and whether participants attended an exercise facility in the community using a Mann–Whitney test. Associations found to have a *P* value less than 0.2 were taken forward to multiple linear regression modeling. Due to MVPA not being normally distributed, MVPA was dichotomized into those patients meeting and not meeting 150 minutes of MVPA per week following the updated physical activity recommendations published by the American College of Sports Medicine (ACSM) (23), which have removed the requirement of activity taking place in bouts of ≥10 minutes. The groups were then analyzed against the possible determinants listed above using Mann–Whitney or chi-square tests. Associations found to have a *P* value less than 0.2 (39) were taken forward to multiple logistic regression modeling. Data analysis was undertaken using SPSS software, version 25, and a statistical significance level was set at a *P* value less than 0.05 for all multivariate tests.

RESULTS

A total of 137 participants provided sociodemographic information (Table 1). A Kruskal-Wallis test revealed that diagnosis was not associated with SB or physical activity levels (SB:

Table 1. Participant sociodemographic characteristics*

Characteristic	Value
Sex, no. (%)	
Female	112 (82)
Male	25 (18)
Age, years	57.8 ± 11.9
Presenting condition, no. (%)	
Rheumatoid arthritis (RA)	73 (53.3)
Inflammatory arthritis excluding RA	37 (27)
Spondyloarthritis	27 (19.7)
Disease duration, years	8.5 ± 11.9
Body mass index (BMI)	31.61 ± 7.37
BMI category, no. (%)	
Underweight	1 (1)
Healthy	24 (18)
Overweight	34 (25)
Obese	78 (57)
Scottish Index of Multiple Deprivation, no. (%)	
1	35 (25.5)
2	28 (20.4)
3	17 (12.4)
4	28 (20.4)
5	29 (21.2)

* Values are the mean ± SD unless indicated otherwise.

$P = 0.50$; LPA: $P = 0.36$; MVPA: $P = 0.89$); therefore, all participants were grouped together for analysis. The total number of patients providing MVPA and LPA data was 122, as some participants were unable to wear the activity monitor due to being allergic to the tape used to attach the device, and some monitors were also not returned. The total number of patients providing SB data was 115 due to the previous reasons, plus incomplete sleep diaries, so that we could not extract true SB during waking hours (Table 2).

Meeting current activity guidelines. In total, 2% of participants (n = 3) met the older ACSM guidelines and European Alliance of Associations for Rheumatology recommendations on physical activity, which are 150 minutes of MVPA in bouts of ≥10 minutes in a week. A total of 29% of participants (n = 35) met the recently updated ACSM guidelines on physical activity, which are 150 minutes of MVPA per week with no requirement of bouts of activity lasting at least 10 minutes. A strong association was found between more time spent in LPA and less in SB (rp = –0.651, $P = 0.000$), a moderate association with more time spent in LPA and more time spent in MVPA (rs = 0.342, $P = 0.000$), and a moderate association with more time spent in MVPA and less time spent in SB (rs = –0.252, $P = 0.007$).

Determinants of SB. A backward multiple regression was run to predict SB from associations found to have a P value of <0.2 (Tables 3–5). The 6-minute walk test ($P = 0.019$) was the only variable left in the model that statistically predicted SB ($F[1, 113] = 5.632$, $P = 0.019$, $R^2 = 4.7\%$). The model indicates (b = –1.787) that for every meter walked on the 6-minute walk test, SB reduces by 1.8 minutes per week.

Determinants of LPA. A backward multiple regression was run to predict LPA from associations found to have a P value of <0.2 (Tables 3–5). Whether or not participants attended an exercise facility in the community and the SF-36 domain of role limitations due to physical health (SF-36 [PH]) were retained in the final model ($F[2, 119] = 5.724$, $P = 0.004$, $R^2 = 8\%$). SF-36 (PH) was statistically significant ($P = 0.008$), as was attending an exercise facility in the community ($P = 0.034$). The model indicated that for every 25% increase in the SF-36 (PH) scale, LPA increased by 5.9 minutes per week, and if the participant attended an exercise facility in the community, LPA increased by 356.7 minutes per week (5.94 hours [5 hours and 57 minutes]).

Determinants of participants meeting 150 minutes of MVPA per week. A backward conditional logistic regression was performed to assess the impact of associations found to have a P value of <0.2 on the likelihood of participants meeting 150 minutes of MVPA per week (Table 6). The final model was statistically significant ($X^2[3, N = 122] = 30.571$, $P < 0.001$), which consisted of the SF-36 domain of role limitations due to emotional problems, the 6-minute walk test, and the RA-EAQ. The model as a whole explained between 22.2% (Cox and Snell R^2) and 31.7% (Nagelkerke R^2) of the variance in meeting 150 minutes of MVPA per week and correctly classified 78.7% of cases. Participants with lower role limitations due to emotional problems ($P = 0.031$), better fitness ($P = 0.002$), and healthier exercise attitudes and beliefs ($P = 0.021$) were more likely to meet the 150 minutes of MVPA per week.

DISCUSSION

Despite evidence for the effectiveness, feasibility, and safety of the physical activity guidelines in people with IJD (2), in the results of this study, only 2% of participants met previous physical activity guidelines (40), and 29% met the updated physical activity guidelines based on the ACSM guidelines of 150 minutes of

Table 2. Waking-time activity levels and sedentary behavior across 7 days monitoring*

	No.	Minimum	Maximum	Mean	SD
SB, minutes (hours)	115	2,095.80 (34.930)	6,016.80 (100.280)	4,100.94 (68.349)	766.68 (12.778)
LPA, minutes (hours)	122	7.90 (<1)	4,122.47 (68.71)	1,902.65 (31.71)	812.34 (13.54)
MVPA, minutes (hours)	122	0.34 (<1)	586.76 (9.78)	120.06 (2)	111.11 (1.85)

* LPA = low physical activity; MVPA = moderate-to-vigorous physical activity; SB = sedentary behavior.

Table 3. Possible determinants of physical activity and sedentary behavior: bivariate analysis, health-related quality of life*

	HAQ DI	SF-36 (PF)	SF-36 (PH)	SF-36 (EP)	SF-36 (EF)	SF-36 (EWB)	SF-36 (SF)	SF-36 (P)	SF-36 (GH)
95% CI	1.21–1.43	32.80–40.63	13.52–24.44	36.32–51.27	29.51–36.40	58.67–66.29	49.77–58.99	36.44–43.74	35.42–42.56
Time in SB									
P	0.41	0.18	0.04	0.14	0.56	0.93	0.32	0.06	0.16
rs	0.08	−0.13	−0.19	−0.14	−0.06	−0.01	−0.09	−0.18	−0.13
Time in LPA									
P	0.47	0.33	0.06	0.14	0.06	0.13	0.02	0.02	0.02
rs	−0.07	0.09	0.17	0.14	0.17	0.14	0.21	0.21	0.21
Meeting 150 minutes of MVPA/week, *P*†	0.05	0.01	0.01	0.00	0.07	0.09	0.01	0.01	0.02

* 95% CI = 95% confidence interval; EF = energy/fatigue; EP = role limitations due to emotional problems; EWB = emotional well-being; GH = general health; HAQ DI = Health Assessment Questionnaire disability index; LPA = low physical activity; MVPA = moderate-to-vigorous physical activity; P = pain; PF = physical functioning; PH = role limitations due to physical health; rs = Spearman's correlation; SB = sedentary behavior; SF = social functioning; SF-36 = Short Form 36.
† *P* value by Mann–Whitney test.

MVPA per week (23). This finding means that only a minority of people with IJD are undertaking the recommended amount of physical activity per week to keep themselves healthy and to decrease their risk of developing noncommunicable diseases (41). The results suggest that on average 10 hours per day are spent in SB during waking hours, and only 17 minutes per day in MVPA. This finding does, however, correlate with the findings of Hernandez-Hernandez et al and Paul et al (5,8) that patients with RA spend more time in SB and less time in MVPA compared to healthy controls. This finding also correlates with the findings of Swinnen et al (6) that people with spondyloarthritis exhibit lower physical activity levels compared to healthy controls. This lack of activity is a major health concern, as increased time spent in SB is independently associated with a greater risk of developing CVD, cancer, and diabetes mellitus, and people with an IJD already have an increased risk of developing CVD compared to healthy controls (11,12)

A lack of time spent in LPA and MVPA found in this study suggests possible reasons why cardiorespiratory fitness has been found to be low in people with RA (14). RA patients appear to spend long periods during waking hours in SB and only short amounts of time undertaking physical activity. A strong correlation in this study has been found between more time spent in LPA and less time spent in SB. Also, a moderate correlation has been found between more time spent in LPA and more time spent in MVPA, therefore indicating an important and significant public health message to try and break up SB by sitting less and moving more. This action may result in an increase in physical activity levels, improving cardiorespiratory fitness and reducing the health risk of developing noncommunicable diseases (41). As previously stated, people with IJD have a higher CVD risk compared to the general population, and the exact reasons for this risk are debatable (11). People with an IJD who have a higher cardiovascular fitness, however, have a better CVD risk profile and a lower 10-year CVD events risk (14). SB is a modifiable CVD risk factor that clinicians should be aiming to address as a high priority in people with IJD.

Limited determinants of SB in people with IJD have been found in this study following bivariate analysis and when taken forward to multivariate regression analysis. Independent determinants of SB were found to be total drug burden, with the more medications a person was prescribed, the more time spent in SB; the more role limitations a person self-reported to have due to physical health, the more time spent in SB; and the lower a

Table 4. Possible determinants of physical activity and sedentary behavior: bivariate analysis, with disease-specific factors*

	Age, years	Disease duration, years	DAS28	VAS pain	Fatigue	Total drug burden
95% CI	55.63–59.75	6.76–10.05	3.58–4.08	4.78–5.69	6.09–6.94	6.27–7.29
Time in SB						
P	0.49	0.25	0.83	0.28	0.80	0.01
rs/rp	rs = 0.07	rs = 0.11	rp = −0.02	rs = 0.10	rs = 0.02	rs = 0.23
Time in LPA						
P	0.61	0.69	0.99	0.14	0.12	0.08
rs/rp	rs = 0.05	rs = −0.04	rp = −0.00	rs = −0.13	rs = −0.14	rs = −0.16
Meeting 150 minutes of MVPA/week, *P*†	0.39	0.51	0.36	0.44	0.18	0.03

* 95% CI = 95% confidence interval; DAS28 = Disease Activity Score in 28 joints; LPA = low physical activity; MVPA = moderate-to-vigorous physical activity; rp = Pearson's correlation; rs = Spearman's correlation; SB = sedentary behavior; VAS = visual analog scale (VAS for pain and the level of fatigue were measured in the same way, using the 0–100-mm line).
† *P* value by Mann–Whitney test.

Table 5. Possible determinants of physical activity and sedentary behavior: bivariate analysis, with personal, physical condition, and environmental factors*

	ASES	RA-EAQ	HADS	SIMD	BMI	MWT6	Grip strength	Enviro 1
95% CI	41.98–47.42	32.55–34.67	13.50–16.35	2.65–3.17	30.31–32.81	298.51–330.92	15.38–18.53	0.16–0.31
Time in SB								
P	0.09	0.29	0.751	0.97	0.46	0.02	0.79	0.1†
rp/rs	rp = –0.16	rs = –0.10	rs = 0.03	rs = 0.01	rs = 0.07	rp = –0.22	rs = 0.03	–
Time in LPA								
P	0.01	0.04	0.12	0.36	0.16	0.26	0.35	0.05†
rp/rs	rp = 0.23	rs = 0.18	rs = –0.14	rs = 0.08	rs = –0.13	rp = 0.10	rs = –0.09	–
Meeting 150 minutes of MVPA/week, *P*	0.1†	<0.01†	0.03†	0.76‡	0.02†	<0.01†	0.98†	0.05‡

* 95% CI = 95% confidence interval; ASES = Arthritis Self-Efficacy Scale; BMI = body mass index; Enviro 1 = attending an exercise facility in the community (yes/no); HADS = Hospital Anxiety and Depression Scale; LPA = low physical activity; MVPA = moderate-to-vigorous physical activity; MWT6 = 6-minute walk test to measure fitness; RA-EAQ = Exercise Attitudes and Beliefs Questionnaire for patients with rheumatoid arthritis; rp = Pearson's correlation; rs = Spearman's correlation; SB = sedentary behavior; SIMD = Scottish Index of Multiple Deprivation.
† *P* value by Mann–Whitney test.
‡ *P* value by chi-square test.

person's fitness, the more time spent in SB. When taken forward to multivariate analysis, no health-related quality of life, disease specific, psychological, personal, or physical conditioning factors apart from the 6-minute walk test, which measures general fitness and endurance, could be found to determine SB in this patient sample. The 6-minute walk test only explained 4.7% of the variance, indicating that either everyone in the study had high amounts of SB, which Table 2 does suggest, therefore resulting in too little variation to be able to explain differences between patients, or indicating that there are other possible determinants of SB in people with IJD that have not been investigated in this study.

More time spent in LPA was associated with attending an exercise facility in the community and having less self-reported role limitations due to physical health. However, these determinants only explained 8% of the variance; therefore again indicating that other possible determinants of LPA exist in people with IJD that have not been investigated in this study. These study findings correspond with the findings of Larkin and Kennedy (19) that an increase in physical health rating increases physical activity levels and of Rongen-van Dartel et al (17) that the level of activity is not associated with pain, disability, coping, or cognition. However, these findings do not agree with their findings that there is an association between increased physical activity and decreased fatigue. These contrasting findings may, however, be explained by the heterogeneity of the study designs and the fatigue measurement tools used. Nonetheless, this study does demonstrate that people with an IJD who attend an exercise facility in the community are more likely to gain the health benefits that activity can bring as their overall activity levels are increased.

People with IJD who have lower role limitations due to emotional problems, better fitness levels, and better exercise attitudes and beliefs were more likely to meet the current ACSM physical activity guidelines (23) of 150 minutes of MVPA per week. The percentage of variance was low (31.7%); therefore other determinants probably exist that were not investigated in this study. These findings appear to inversely correspond with the findings of Larkin and Kennedy (19) that an increase in physical activity increases motivation to exercise, increases metal health, and increases beliefs about the benefits of physical activity.

These findings indicate the probability that if people with IJD meet the physical activity guidelines (23), they will have better fitness levels, which will decrease their CVD risk profile and lower their 10-year CVD events risk (5,14). They may also improve their mental health and well-being, as depression has been found to be more common in patients with RA than in healthy individuals (32,42).

A limitation of this study may be the wearing of an activity monitor. Although limited standardized information was given about the device, participants may have been more active due to wearing the device. If this possibility is the case, genuine activity levels could be over-recorded and SB under-recorded. This limitation could make the overall findings with regards to time spent

Table 6. Model for participants meeting 150 minutes of moderate-to-vigorous physical activity per week*

	B	SE	Wald	df	Sig	Exp(B) (95% CI)
SF-36 (EP)	0.011	0.005	4.679	1	0.031	1.011 (1.001–1.022)
MWT6	0.009	0.003	9.903	1	0.002	1.009 (1.003–1.015)
RA-EAQ	0.108	0.047	5.334	1	0.021	1.114 (1.017–1.221)

* 95% CI = 95% confidence interval; MWT6 = 6-minute walk test to measure fitness; RA-EAQ = Exercise Attitudes and Beliefs Questionnaire for patients with rheumatoid arthritis; SF-36 (EP) = Short Form 36 health survey role limitations due to emotional problems.

in SB, LPA, and MVPA even more alarming. Another limitation could be that the study participants were recruited from referrals made into an NHS-run IAEP, therefore already showing an interest and willingness to becoming more active. The participants may have also received a consultation from a health professional on the benefits of exercise and been given advice and information prior to being recruited into the study. If so, generalizability to the wider IJD population may be reduced, as this study may not have recruited the most inactive of participants. However, this possibility would essentially mean that the issues described in this article are even more pronounced in that wider population.

In conclusion, the majority of people with IJD in this study did not meet the current guidelines on physical activity. Those who did appeared to have increased fitness, better mental health, and better exercise attitudes and beliefs. However, many hours per day were spent in SB. Few determinants of SB and physical activity could be found when factors such as health-related quality of life, or disease-specific, psychological, personal, or physical conditioning were investigated. There was a strong correlation with regard to more time spent in LPA and less time spent in SB, with a moderate correlation with more time spent in LPA and more time spent in MVPA. This finding therefore may mean that if SB can be broken up, then more LPA will be undertaken, which may result in more MVPA. Further research looking into physical activity levels over time is required to fully address this issue. Further research is also needed into other possible determinants of physical activity and SB that have not been investigated in this study.

ROLE OF THE STUDY SPONSOR

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

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. Ms. Bell 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. Bell, Hendry, Steultjens.
Acquisition of data. Bell, Hendry, Steultjens.
Analysis and interpretation of data. Bell, Hendry, Steultjens.

REFERENCES

1. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public Health Rep 1985;100:126–31.
2. Osthoff AR, Niedermann K, Braun J, Adams J, Brodin N, Dagfinrud H, et al. 2018 EULAR recommendations for physical activity in people with inflammatory arthritis and osteoarthritis. Ann Rheum Dis 2018;77:1251–60.
3. Hurkmans E, Van Der Giesen FJ, Vliet Vlieland TP, Schoones J, Van Den Ende EC. Dynamic exercise programs (aerobic capacity and/or muscle strength training) in patients with rheumatoid arthritis. Cochrane Database Syst Rev 2009;4:CD006853.
4. Baillet A, Vaillant M, Guinot M, Juvin R, Gaudin P. Efficacy of resistance exercises in rheumatoid arthritis: meta-analysis of randomized controlled trials. Rheumatology (Oxford) 2012;51:519–27.
5. Hernandez-Hernandez V, Ferraz-Amaro I, Diaz-Gonzalez F. Influence of disease activity on the physical activity of rheumatoid arthritis patients. Rheumatology (Oxford) 2014;53:722–31.
6. Swinnen TW, Scheers T, Lefevre J, Dankaerts W, Westhovens R, de Vlam K. Physical activity assessment in patients with axial spondyloarthritis compared to healthy controls: a technology-based approach. PLoS One 2014;9:e85309
7. Lee J, Dunlop D, Ehrlich-Jones L, Semanik P, Song J, Manheim L, et al. Public health impact of risk factors for physical inactivity in adults with rheumatoid arthritis. Arthritis Care Res (Hoboken) 2012;64:488–93.
8. Paul L, Rafferty D, Marshall-McKenna R, Gill JM, McInnes I, Porter D, et al. Oxygen cost of walking, physical activity, and sedentary behaviours in rheumatoid arthritis. Scand J Rheumatol 2014;43:28–34.
9. Henchoz Y, Zufferey P, So A. Stages of change, barriers, benefits, and preferences for exercise in RA patients: a cross-sectional study. Scand J Rheumatol 2013;42:136–45.
10. Neuberger GB, Aaronson LS, Gajewski B, Embretson SE, Cagle PE, Loudon JK, et al. Predictors of exercise and effects of exercise on symptoms, function, aerobic fitness, and disease outcomes of rheumatoid arthritis. Arthritis Rheum 2007;57:943–52.
11. Agca R, Heslinga SC, Rollefstad S, Heslinga M, McInnes IB, Peters MJ, et al. EULAR recommendations for cardiovascular disease risk management in patients with rheumatoid arthritis and other forms of inflammatory joint disorders: 2015/2016 update. Ann Rheum Dis 2017;76:17–28.
12. Avina-Zubieta JA, Thomas J, Sadatsafavi M, Lehman AJ, Lacaille D. Risk of incident cardiovascular events in patients with rheumatoid arthritis: a meta-analysis of observational studies. Ann Rheum Dis 2012;71:1524–9.
13. Koivuniemi R, Paimela L, Suomalainen R, Leirisalo-Repo M. Cardiovascular diseases in patients with rheumatoid arthritis. Scand J Rheumatol 2013;42:131–5.
14. Metsios GS, Koutedakis Y, Veldhuijzen van Zanten JJ, Stavropoulos-Kalinoglou A, Vitalis P, et al. Cardiorespiratory fitness levels and their association with cardiovascular profile in patients with rheumatoid arthritis: a cross-sectional study. Rheumatology (Oxford) 2015;54:2215–20.
15. Sedentary Behaviour Research Network. What is sedentary behaviour? 2016. URL: http://www.sedentarybehaviour.org/what-is-sedentary-behaviour/.
16. Biswas A, Oh PI, Faulkner GE, Bajaj RR, Silver MA, Mitchell MS, et al. Sedentary time and its association with risk for disease incidence, mortality, and hospitalization in adults: a systemic review and meta-analysis. Ann Intern Med 2015;162:123–32.
17. Rongen-van Dartel SA, Repping-Wuts H, van Hoogmoed D, Knoop H, Bleijenberg G, van Riel PL, et al. Relationship between objectively assessed physical activity and fatigue in patients with rheumatoid arthritis: inverse correlation of activity and fatigue. Arthritis Care Res (Hoboken) 2014;66:852–60.
18. Sveaas SH, Smedslund G, Hagen KB, Dagfinrud H. Effect of cardiorespiratory and strength exercises on disease activity in patients with inflammatory rheumatic diseases: a systematic review and meta-analysis. Br J Sports Med 2017;51:1065–72.
19. Larkin L, Kennedy N. Correlates of physical activity in adults with rheumatoid arthritis: a systematic review. J Phys Act Health 2014;11:1248–61.
20. Chastin SF, Culhane B, Dall PM. Comparison of self-reported measure of sitting time (IPAQ) with objective measurement (activPAL). Physiol Meas 2014;35:2319–28.
21. Garber CE, Blissmer B, Deschenes MR, Franklin BA, Lamonte MJ, Lee IM, et al. American College of Sports Medicine position stand.

Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise. Med Sci Sports Exerc 2011;43:1334–59.

22. American College of Sports Medicine. Resource manual for guidelines for exercise testing and prescription. 6th ed. Philadelphia: Lippincott Williams and Wilkins, 2010.
23. American College of Sports Medicine. Updated physical activity guidelines now available. 2018. URL: https://www.acsm.org/read-research/newsroom/news-releases/news-detail/2018/11/12/updated-physical-activity-guidelines-now-available.
24. Trost SG, Owen N, Bauman AE, Sallis JF, Brown W. Correlates of adults' participation in physical activity: review and update. Med Sci Sports Exerc 2002;34:1996–2001.
25. Granat MH. Event-based analysis of free-living behaviour. Physiol Meas 2012;33:1785–800.
26. Tudor-Locke C, Rowe DA. Using cadence to study free-living ambulatory behaviour. J Sports Med 2012;42:381–98.
27. Mathews CE, Ainsworth B, Thompson RW, Basset DR. Sources of variance in daily physical activity levels as measured by an accelerometer. Med Sci Sports Exerc 2002;34:1376–81.
28. Hurst NP, Ruta DA, Kind P. Comparison of the MOS short form-12 (SF12) health status questionnaire with the SF36 in patients with rheumatoid arthritis. Br J Rheumatol 1998;37:862–9.
29. Lorig K, Chastain RL, Ung E, Shoor S, Holman HR. Development and evaluation of a scale to measure perceived self-efficacy in people with arthritis. Arthritis Rheum 1989;32:37–44.
30. Desai NK. Psychometric properties of an exercise attitude and belief questionnaire for patients with rheumatoid arthritis. Boston: MGH Institute of Health Professions; 2009.
31. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatrica Scandinavica 1983;67:361–70.
32. Dickens C, McGowan L, Clark-Carter D, Creed F. Depression in rheumatoid arthritis: a systematic review of the literature with meta-analysis. Psychosom Med 2002;64:52–60.
33. Scottish Government. Scottish Index of Multiple Deprivation 2000 . 2012; URL: http://simd.scotland.gov.uk/publication-2012/.
34. Tveter AT, Dagfinrud H, Moseng T, Holm I. Measuring health-related physical fitness in physiotherapy practice: reliability, validity, and feasibility of clinical field tests and a patient-reported measure. J Orthop Sports Phys Ther 2014;44:206–16.
35. American Thoracic Society Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: guidelines for the six-minute walk test. Am J Respir Crit Care Med 2002;166:111–7.
36. Beekman E, Mesters I, Hendriks EJ, Klaassen MP, Gosselink R, Van Schayck OC, et al. Course length of 30 metres versus 10 metres has a significant influence on six-minute walk distance in patients with COPD: an experimental crossover study. J Physiother 2013;59:169–76.
37. Metsios GS, Stavropoulos-Kalinoglou A, Veldhuijzen Van Zanten JJ, Treharne GJ, Panoulas VF, Douglas KM, et al. Rheumatoid arthritis, cardiovascular disease and physical exercise: a systematic review. Rheumatology (Oxford) 2008;47:239–48.
38. Roberts HC, Denison HJ, Martin HJ, Patel HP, Syddall H, Cooper C, et al. A review of the measurement of grip strength in clinical and epidemiological studies: towards a standardised approach. Age Ageing 2011;40:423–9.
39. Holla JF, Steultjens MP, Roorda LD, Heymans MW, ten Wolde S, Dekker J. Prognostic factors for the two-year course of activity limitations in early osteoarthritis of the hip and/or knee. Arthritis Care Res (Hoboken) 2010;62:1415–25.
40. Haskell WL, Lee IM, Pate RR, Powell KE, Blair SN, Franklin BA, et al. Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. Med Sci Sports Exerc 2007;39:1423–34.
41. Guthold R, Stevens GA, Riley LM, Bull FC. Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1·9 million participants. Lancet Glob Health 2018;6:1077–86.
42. Margaretten M, Julian L, Katz P, Yelin E. Depression in patients with rheumatoid arthritis: description, causes and mechanisms. Int J Clin Rheumatol 2011;6:617–23.

Arthritis Care & Research
Vol. 74, No. 3, March 2022, pp 501–503

American College of Rheumatology
Empowering Rheumatology Professionals

LETTERS

DOI 10.1002/acr.24843

Reexamining remission definitions in rheumatoid arthritis: considering the 28-joint Disease Activity Score, C-reactive protein level, and patient global assessment: comment on the article by Felson et al

To the Editor:

We read with great interest the editorial by Felson et al on definitions of remission in rheumatoid arthritis (RA), recently published in *Arthritis Care & Research* (1). The article gives a comprehensive and historical overview of the development of remission criteria and provides a well-founded critique of remission criteria based on the 28-joint Disease Activity Score (DAS28). The DAS28 has been primarily developed and validated for evaluations at the group level, i.e., for measuring effects in clinical trials. However, in almost forgotten earlier times, when patient remission was rarely achieved, there was a need for a single index, expressing disease activity of the individual patient, and the only instrument available was the 44-joint Disease Activity Score (2). When biologics became available in many countries of Europe, the use of the DAS28 as a single index of disease activity was also stimulated by health authorities and insurance companies, requiring DAS28 proof of active RA and documented previous treatment failure (or contraindication) of conventional synthetic disease-modifying antirheumatic drugs, before allowing reimbursement of an expensive biologic drug. Since then, remission has proved to be an achievable goal, and for clinical trials and for individual patients, DAS28 cutoffs have been used for this purpose, especially in Europe, although their limitations for evaluations at the individual patient level have indeed been recognized (3).

Moreover, we agree with Felson et al that patient global assessment (PtGA) is a valuable assessment. However, we feel compelled to clarify the misunderstanding that seems to persist regarding our relatively simple proposal. We do not suggest merely eliminating PtGA from the definitions of remission; we suggest that a second target, based on valid and discriminative patient-reported measures of disease impact, be adopted, in parallel but separated from the existing target for inflammatory disease activity, which, we believe, could be refined by the exclusion of PtGA. Although Felson et al cite our article (4), they do not depict our proposal for this dual-target strategy and its conceptual framework, summarized in the conclusions of that article. Following our proposal, the patient's perspective would become more valued, rather than being ignored.

We disagree with the interpretation of the evidence provided by Felson et al to support the concept that PtGA should be kept as a component of the American College of Rheumatology (ACR)/European Alliance of Associations for Rheumatology (EULAR) definitions of remission. Although PtGA and measures of clinical disease activity are correlated at high levels of disease activity, contributing to the ability of PtGA to distinguish active treatment from placebo in the context of clinical trials, they are only poorly, if at all, correlated at low levels of disease activity (5,6), precisely when the practicing clinician needs to make difficult decisions regarding escalating or maintaining immunosuppressive/immunomodulatory therapy. Thus, while the inclusion of PtGA may facilitate the distinction between treatments in clinical trials, we are concerned regarding the implications of including PtGA as an element of composite definitions of remission used to tailor immunosuppressive/immunomodulatory therapy in clinical practice and the potential risk of overtreatment that this practice entails. As many as 45–61% of all patients with RA (in clinical trials [4] and cohort studies [7]) who are otherwise in remission fail to meet the Boolean definition of remission solely because of a too high PtGA score. These patients, in so-called PtGA near-remission, are exposed to the risk of overtreatment, because their disease cannot be improved by additional immunosuppression/immunomodulation. However, they still endure a significant impact of nondisease activity manifestations and outcomes of the disease (8), which were recently touched upon in the EULAR points to consider for the management of difficult-to-treat RA (9). The use of the ACR/EULAR remission definitions in clinical practice was explicitly predicted in the original 2011 report (10), and the definitions have been extensively adopted as part of the treat-to-target strategy. Thus, the implications of these definitions are more extensive than those for clinical trials only.

The assertion that PtGA reflects subclinical inflammation is, in our view, unsupported by evidence. We, and in fact, some of the authors of the editorial themselves, have shown no correlation between PtGA and joint damage accrual (11). We have also demonstrated that in patients who are in PtGA near-remission there is no evidence of inflammation in other joints or synovial structures, through extensive ultrasonography assessment (12). It is difficult to envisage what room is left for the consideration in the editorial that "…the patient global assessment reflects components of disease activity that are otherwise not captured, …as inflammation in joints not included in a 28-joint count, such as the feet and ankles." This is, therefore, not the reason "why high patient global assessment scores, even when 28-joint counts are low, identify patients at high risk of later functional loss" (1). This may be simply and better explained by the fact that function is a major

determinant of PtGA, irrespective of inflammatory disease activity, as repeatedly reported (5,6,8,13). These publications are the basis of our dual-target strategy proposal, which, we hypothesize, may result in more accurate and comprehensive definitions of remission. We proposed the dual target to comprise 1) biologic remission, which will be sharper and more sensitive to help guide immunosuppressive/immunomodulatory therapy in individual patients in clinical practice, and 2) patient remission, also addressing all other important aspects of nondisease activity manifestations, of outcomes of the disease, and of medication adverse effects (disease impact), and will thus be more informative than the current 1-item PtGA. Surely, this approach highlights the importance of patients' perspective, as it ensures that clinicians address both the disease activity and the disease impact aspects accordingly.

In summary, we agree with many of the points made in the editorial by Felson et al, but we feel that it distorts our proposal by omitting to mention the patient remission aspect, which is what makes it a dual target: a holistic strategy that empowers patients and promotes health by allowing patients to gain greater control over decisions and actions affecting their health, a World Health Organization recommendation since the Ottawa conference in 1986.

Author disclosures are available at https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1002%2Facr.24843&file=acr24843-sup-0001-Disclosureform.pdf.

Ricardo J. O. Ferreira, PhD
Centro Hospitalar e Universitário de Coimbra and Nursing School of Coimbra
Coimbra, Portugal
Paco M. J. Welsing, PhD
Johannes W. G. Jacobs, PhD
University Medical Center
Utrecht, The Netherlands
Laure Gossec, PhD
Sorbonne Université–INSERM and Pitié Salpêtrière hospital, AP-HP
Paris, France
Mwidimi Ndosi, PhD
University of the West of England
Bristol, UK
Pedro M. Machado, PhD
University College London, University College London Hospitals NHS Foundation Trust, and Northwick Park Hospital, London North West University Healthcare NHS Trust
London, UK
Désirée van der Heijde, PhD
Leiden University Medical Centre
Leiden, The Netherlands
José A. P. da Silva, PhD
Centro Hospitalar e Universitário de Coimbra and University of Coimbra
Coimbra, Portugal

1. Felson DT, Lacaille D, LaValley MP, Aletaha D. Reexamining remission definitions in rheumatoid arthritis: considering the twenty-eight–joint Disease Activity Score, C-reactive protein level, and patient global assessment [editorial]. Arthritis Care Res (Hoboken) 2022;74:1–5.
2. Van der Heijde DM, van 't Hof MA, van Riel PL, Theunisse LA, Lubberts EW, van Leeuwen MA, et al. Judging disease activity in clinical practice in rheumatoid arthritis: first step in the development of a disease activity score. Ann Rheum Dis 1990;49:916–20.
3. Jacobs JW, Ten Cate DF, van Laar JM. Monitoring of rheumatoid arthritis disease activity in individual patients: still a hurdle when implementing the treat-to-target principle in daily clinical practice [editorial]. Rheumatology (Oxford) 2015;54:959–61.
4. Ferreira RJ, Welsing PM, Jacobs JW, Gossec L, Ndosi M, Machado PM, et al. Revisiting the use of remission criteria for rheumatoid arthritis by excluding patient global assessment: an individual meta-analysis of 5792 patients. Ann Rheum Dis 2020;80:293–303.
5. Ferreira RJ, Duarte C, Ndosi M, de Wit M, Gossec L, da Silva JA. Suppressing inflammation in rheumatoid arthritis: does patient global assessment blur the target? A practice-based call for a paradigm change. Arthritis Care Res (Hoboken) 2018;70:369–78.
6. Ferreira RJ, Carvalho PD, Ndosi M, Duarte C, Chopra A, Murphy E, et al. Impact of patient's global assessment on achieving remission in patients with rheumatoid arthritis: a multinational study using the METEOR database. Arthritis Care Res (Hoboken) 2019;71:1317–25.
7. Ferreira RJ, Santos E, Gossec L, da Silva JA. The patient global assessment in RA precludes the majority of patients otherwise in remission to reach this status in clinical practice. Should we continue to ignore this? Semin Arthritis Rheum 2020;50:583–5.
8. Ferreira RJ, Dougados M, Kirwan JR, Duarte C, de Wit M, Soubrier M, et al. Drivers of patient global assessment in patients with rheumatoid arthritis who are close to remission: an analysis of 1588 patients. Rheumatology (Oxford) 2017;56:1573–8.
9. Nagy G, Roodenrijs NM, Welsing PM, Kedves M, Hamar A, van der Goes MC, et al. EULAR points to consider for the management of difficult-to-treat rheumatoid arthritis. Ann Rheum Dis 2022;80:20–33.
10. Felson DT, Smolen JS, Wells G, Zhang B, van Tuyl LH, Funovits J, et al. American College of Rheumatology/European League Against Rheumatism provisional definition of remission in rheumatoid arthritis for clinical trials. Arthritis Rheum 2011;63:573–86.
11. Studenic P, Felson D, de Wit M, Alasti F, Stamm TA, Smolen JS, et al. Testing different thresholds for patient global assessment in defining remission for rheumatoid arthritis: are the current ACR/EULAR Boolean criteria optimal? Ann Rheum Dis 2020;79:445–52.
12. Brites L, Rovisco J, Costa F, Freitas JPD, Jesus D, Eugénio G, et al. High patient global assessment scores in patients with rheumatoid arthritis otherwise in remission do not reflect subclinical inflammation. Joint Bone Spine 2021;88:105242.
13. Craig ET, Perin J, Zeger S, Curtis JR, Bykerk VP, Bingham CO III, et al. What does the patient global health assessment in rheumatoid arthritis really tell us? Contribution of specific dimensions of health-related quality of life. Arthritis Care Res (Hoboken) 2020;72:1571–8.

DOI 10.1002/acr.24842

Reply

To the Editor:

We read with interest the letter by Ferreira and colleagues in response to our editorial about the measurement of remission in

RA. These authors agree with large parts of our editorial, particularly the importance of including patient-reported outcomes in the evaluation of RA disease activity. Their main argument is that a patient-reported outcome such as the PtGA does not capture active RA when the DAS28 suggests quiescent disease. For example, in one article, the authors suggested that the main predictors of PtGA in patients with a small number of swollen and tender joints in the DAS28 are pain and fatigue (1).

The DAS28 was adopted as an RA outcome measure based on data suggesting that this reduced count (ignoring feet/ankles/hips/neck) could adequately capture response to treatment. It was never intended to comprehensively assess all active joints. The pain experienced by patients with persistently elevated PtGA scores who have low DAS28 levels may well be due to RA disease activity affecting joints not counted but that are often affected by RA. In fact, one of Dr. Ferreira's coauthors championed the assessment of structural disease in the feet, showing that including this site enhanced our ability to comprehensively assess structural damage in RA. We believe that PtGA gives us a window into overall disease activity, including disease in these other joints. We further note that 1 active joint among the 28 joints counted can cause patients considerable pain and can serve as the basis for elevated PtGA scores. In psoriatic arthritis, investigators recently agreed to readopt a 66/68-joint count to evaluate outcomes, including remission, for the reasons outlined above (2).

Ferreira and colleagues argue that they are concerned about overtreatment that may lead to unnecessary exposure to potentially harmful medicines. They assert that there is a sizable group of patients in near-remission (3,4) because the patients fail to reach the PtGA cut point of ≤1, and that additional treatment of these patients would constitute overtreatment. However, we disagree with the concern about overtreatment; no single study has to date shown that overtreatment is a problem. In a recent publication using the 44-joint DAS of <1.6 as the treat-to-target goal, 38% of RA patients not on target were found not to have treatments increased (i.e., were undertreated), while only 9% of those with treatment increased were actually at the 44-joint DAS target of <1.6. So, while the concern of our colleagues for patients is appreciated, evidence suggests that the main problem is under- and not overtreatment in patients in RA. With all the therapies available to patients with RA, all efforts by the scientific community should be taken to achieve remission for at least the majority of our patients in the third decade of the 21st century.

Our colleagues argue in favor of using C-reactive protein (CRP) level as a main outcome, since it reflects the inflammatory response. However, as noted in our editorial, effective RA treatments have variable effects on CRP level, with some, such as anti–interleukin 6 agents and JAK inhibitors, causing a drop in CRP level, whereas others do not affect CRP level. Considering also the possibility that CRP level may reflect infection and not inflammation, an exclusion of CRP level from composite instruments would therefore follow an analogous logic as PtGA exclusion. All this would be a dramatic step back in time, before composite measures had been introduced to cover the activity of a complex systemic disease such as RA.

Lastly, we are concerned that a dual target as proposed by Ferreira et al will make it easier for sponsors of new treatments to focus on the easier to achieve target for approval (as has already occurred with DAS28 remission thresholds). If so, the dual-target approach would lead to ignoring patient assessments entirely. Although the authors argue that separating patient-reported outcomes from objective markers will serve patient interests, the reality will likely be the opposite, with the separation leading to patient-reported outcomes being measured on the side, as a secondary outcome. So in our view, the only guarantee that patient-reported outcomes will not be "put aside" is if they continue to remain an integral, and thus required, part of disease activity instruments.

David T. Felson, MD, MPH
Boston University School of Medicine
Boston, Massachusetts
and University of Manchester and the NIHR Manchester
BRC, Manchester University NHS Trust
Manchester, UK
Diane Lacaille, MD, MHSc
University of British Columbia
and Arthritis Research Canada
Vancouver, British Columbia, Canada
Michael P. LaValley, PhD
Boston University School of Public Health
Boston, Massachusetts
Daniel Aletaha, MD, MS, MBA
Medical University Vienna
Vienna, Austria

1. Ferreira RJ, Dougados M, Kirwan JR, Duarte C, de Wit M, Soubrier M, et al. Drivers of patient global assessment in patients with rheumatoid arthritis who are close to remission: an analysis of 1588 patients. Rheumatology (Oxford) 2017;56:1573–8.
2. Duarte-García A, Leung YY, Coates LC, Beaton D, Christensen R, Craig ET, et al. Endorsement of the 66/68 joint count for the measurement of musculoskeletal disease activity: OmeRACT 2018 psoriatic arthritis workshop report. J Rheumatol 2019;46:996–1005.
3. Ferreira RJ, Duarte C, Ndosi M, de Wit M, Gossec L, da Silva JA. Suppressing inflammation in rheumatoid arthritis: does patient global assessment blur the target? A practice-based call for a paradigm change. Arthritis Care Res (Hoboken) 2018;70:369–78.
4. Studenic P, Smolen JS, Aletaha D. Near misses of ACR/EULAR criteria for remission: effects of patient global assessment in Boolean and index-based definitions. Ann Rheum Dis 2012;71:1702–5.

ARP Announcements

Association of Rheumatology Professionals
2200 Lake Boulevard NE, Atlanta, Georgia 30319
www.rheumatology.org

Submissions are invited for the AC&R 2023 Themed Issue: Health Disparities and Health Equity in the Rheumatic Diseases

Arthritis Care & Research is soliciting manuscripts for a Themed Issue addressing health disparities and health equity, as related to outcomes and issues in the rheumatic diseases. This theme issue is part of an initiative across the American College of Rheumatology journals to better address these important issues in our field. Manuscripts may focus on health disparities or health equity related to race and ethnicity, sex and gender, socioeconomic status, sexual orientation, or other characteristics. Exploration of health disparities among understudied patient groups, as well as the intersection of multiple patient characteristics, are encouraged.

Manuscripts covering a broad range of topics related to the major theme are invited. Examples include observational studies that elucidate factors underlying disparities in health care quality or access; intervention studies that address health disparities; studies of differential impacts of treatments or behavioral interventions; studies describing mechanisms underlying disparities in key outcomes in rheumatic diseases (e.g., pain, function). Manuscripts addressing research related to disparities in rheumatology training and work force are also of interest. Both Original Research and Review articles will be considered.

The 2023 Themed Issue will include regular submissions as well, but a certain number of pages will be reserved for manuscripts accepted in response to this solicitation. All manuscripts will be peer reviewed. The Editor will select papers for publication in the Themed Issue based on reviewer ratings and the balance of subject matter. It is possible that manuscripts submitted for the themed issue may be accepted for publication in a regular issue of *Arthritis Care & Research,* rather than the themed issue.

Please follow the formatting requirements found in the Author Guidelines section at https://onlinelibrary.wiley.com/page/journal/21514658/homepage/ForAuthors.html. The deadline for submission is March 31, 2022. For further information, contact the Editor of Arthritis Care & Research, Dr. Kelli D. Allen; email: kdallen@email.unc.edu.

Download the ACR Publications Mobile App

The ACR Publications app can be downloaded for free from the Apple store or Google Play. ACR members can log in for full-text access to all articles in *Arthritis Care & Research* and *Arthritis & Rheumatology*. Nonmembers can access abstracts of all *AC&R* and *A&R* articles, the full text of articles published more than one year ago, and select open-access articles published recently, as well as the full text of all articles from *ACR Open Rheumatology* and *The Rheumatologist*.

ARP Membership

The Association of Rheumatology Professionals (ARP), a division of the American College of Rheumatology, appreciates your continued membership and looks forward to serving you another year. Membership costs range from $30 to $140. ARP welcomes nurse practitioners, nurses, physician assistants, office staff, researchers, physical therapists, occupational therapists, assistants, and students. Student membership is complimentary; the Annual Meeting registration fee is waived for students who submit the required student verification letter. For information, go to www.rheumatology.org and select "Membership" or call 404-633-3777 and ask for an ARP staff member.

ACR Journal Twitter Account (@ACR_Journals) and Social Media Editor

The ACR journals are heightening our focus on social media, to benefit authors and readers. Among our first activities is the introduction of an official ACR Journals Twitter account: @ACR_Journals. Followers will enjoy special features and the opportunity to engage with authors and other fellow professionals about studies published in *Arthritis Care & Research, Arthritis & Rheumatology,* and *ACR Open Rheumatology*. Authors of published articles will have the opportunity to use @ACR_Journals to share their work and engage in dialogue with others interested in the research. The journals welcome Dr. Paul Sufka of Minneapolis as our first Social Media Editor.